9-2018

Phage Therapy and Development of Delivery Systems for Gram-Positive Phage Endolysins

Jude Ajuebor
Department of Biological Sciences, Cork Institute of Technology

Phage therapy and development of delivery systems for Gram-positive phage endolysins

A thesis submitted to Cork Institute of Technology

for the degree of

Doctor of Philosophy

By

Jude Ajuebor, B.Sc.

Supervisor: Prof. Aidan Coffey

Department of Biological Sciences, Cork Institute of Technology

September 2018

Acknowledgements

Foremost, I would like to thank my supervisor Prof. Aidan Coffey for his support, guidance and devotion throughout the duration of my PhD.

I would also like to thank the Lecturing and Technical staff in the Department of Biological Sciences for their support and advice. My sincere appreciation to those scientists that graciously contributed invaluable resource and tools aiding me throughout this work, Dr. Daniel Burke, Dr. Andrey Shkoporov, Dr. Angela Back, Dr. Mark J. Van Raaij, Dr. Antonio Pichel Beleiro, Dr. Horst Neve, Dr. Olivia McAuliffe, Prof. Jim O'Mahony and Prof. Colin Hill.

A special gratitude to all the postgrad past and present for their friendship and support during the duration of my PhD, in particular Marcel de Kruij, Rodney Govender, Colin Buttimer and Caiomhe Lynch. A special thanks also to my friends Adonai Djankah and Carlota Marquez Grana for their advice and support.

Finally, to my family, special thanks for your understanding and continued support during the duration of my academic pursuit.

Thesis Abstract

This thesis focussed on Gram positive phages and their endolysins. Here, two similar *kay-like* staphylococcal phages B1 (vB_SauM_B1) and JA1 (vB_SauM_JA1) were isolated from a commercial therapeutic phage mix. Their host range was established on the Irish National MRSA bank, which included twenty one sequence types in addition relevant control strains. Based on this, distinct phages were identified and subjected to genome sequencing. The sequences were compared with the sequence of phage K (vB_SauM_K), which was also determined in this work. All three phages had a genome size of at least 139 kb, although some key differences were identified between each. The new phages B1 and JA1 possessed double stranded DNA and generally had a broader host range than phage K. A comparative genomic analysis on the phage genomes identified several (open reading frames) ORFs that were absent in the genome of phage K but present in genomes of phages B1 and JA1. One of the cloned genes from phage K was shown to encode a protein for the receptor-binding-protein and this protein was demonstrated to slightly inhibit phage adsorption. The other cloned gene encoded the phage endolysin and this peptidoglycan hydrolase were identical across all three phages and thus, the CHAPk endolysin of phage K was chosen to demonstrate the application of the endolysin for the control of staphylococci in milk. A two-log reduction in staphylococcal numbers in milk was observed. When the endolysin was introduced into a lactococcal secretion system using the pNZ8048 vector, detectable secretion was successfully demonstrated. Simultaneously, a *Clostridium difficile* phage endolysin, an amidase, was also cloned into the same secretion system with successful secretion also being demonstrated. In addition, this latter endolysin was also secreted from a recombinant *E. coli* strain, suggesting potential applications for delivery of the endolysin to the intestine from a hypothetical probiotic *E. coli* strain.

List of publications

- Ajuebor, J.; McAuliffe, O.; O'Mahony, J.; Ross, R. P.; Hill, C.; Coffey, A. Bacteriophage endolysins and their applications. *Sci. Prog.* **2016**, *99*, 183–199, doi:10.3184/003685016X14627913637705.

- Hathaway, H.; Ajuebor, J.; Stephens, L.; Coffey, A.; Potter, U.; Sutton, J. M.; Jenkins, A. T. A. Thermally triggered release of the bacteriophage endolysin CHAPKand the bacteriocin lysostaphin for the control of methicillin resistant *Staphylococcus aureus* (MRSA). *J. Control. Release* **2017**, *245*, 108–115, doi:10.1016/j.jconrel.2016.11.030.

- Ajuebor, J.; Buttimer, C.; Arroyo-moreno, S.; Chanishvili, N.; Gabriel, E. M.; Mahony, J. O.; Mcauliffe, O.; Neve, H.; Franz, C.; Coffey, A. Comparison of *Staphylococcus* phage K with close phage relatives commonly employed in phage therapeutics. *Antibiotics* **2018**, *7*, 37, doi:10.3390/antibiotics7020037.

List of conferences attented

- Ajuebor, J., McAuliffe, O., O'Mahony, J., Ross, R.P., Hill, C., Coffey, A. (2015). Expression and secretion of staphylococcal phage endolysin CHAP in a *Lactococcus lactis* cheese starter. *44th Annual Food Research Conference. Cork, Ireland. Dec 14 2015.*

- Ajuebor, J., McAuliffe, O., O'Mahony, J., Ross, R.P., Hill, C., Coffey, A. (2016). Engineering a phage endolysin against *Clostridium difficile* for delivery into the gastrointestinal tract. *EMBO conference on the Viruses of Microbes. Liverpool, U.K. July 18-22 2016.*

- Ajuebor, J., Keating, A., Djankah, A., McAuliffe, O., O'Mahony, J., Ross, R.P., Hill, C., Coffey, A. (2017). Host range and comparative genomics analysis on three similar *Myoviruses*. *Phages 2017: Bacteriophage in medicine, food and biotechnology. Oxford, U.K. September 13-14 2017.*

List of Abbreviations

ORF	Open Reading Frames
CHAP	Cysteine/Histidine-Dependent Amidohydrolase/Peptidase
DNA	Deoxyribonucleic Acid
RNA	Ribonucleic Acid
CWBD	Cell Wall Binding Domain
EDTA	Ethylenediaminetetraacetic Acid
TMD	Transmembrane Domain
LAB	Lactic Acid Bacteria
GIT	Gastrointestinal Tract
PCR	Polymerase Chain Reaction
SPR	Surface Plasmon Resonance
MRSA	Methicilin resistant *Staphylococcu aureus*
BHI	Brain Heart Infusion
CFU	Colony Forming Unit
PFU	Plaque Forming Unit
EOP	Efficiency of Plaquing
MOI	Multiplicity of Infection
SDS	Sodium Dodecyl Sulphate
LTR	Long Terminal Repeats
NCBI	National Centre for Biotechnology Information
BLAST	Basic Local Alignment Search Tool
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrilamide Gel Electrophoris
OD	Optical Density
RBP	Receptor Bnding Protein

PBS	Phoshate Buffered Saline
VICTOR	Virus Classification and Tree Building Online Resource
SOE	Splicing by Overlap Extention
GRAS	Generally Regarded as Safe
IPTG	Isopropyl β-D_1-thiogalactopyranoside
SLPA	Surface Layer Protein A
BRP	Bacteriocin Release Protein
LB	Luria-Bertani
MRS	De Man, Rogosa and Sharpe
MLST	Multilocus Sequence Typing
TEM	Transmission Electron Microscopy

Table of Contents

Chapter I

Bacteriophage endolysins and their applications

A manuscript based on this chapter has been published in the journal *Science Progress*

1.1Abstract

Endolysins (lysins) are bacteriophage-encoded enzymes that have evolved to degrade specific bonds within the bacterial cell wall. These enzymes represent a novel class of antibacterial agents against infectious pathogens, especially in light of multidrug-resistant bacteria, which have made antibiotic therapy increasingly difficult. Lysins have been used successfully to eliminate/control bacterial pathogens in various anatomical locations in mouse and other animal models. Engineering tactics have also been successfully applied to improve lysin function. This review discusses the structure and function of lysins. It highlights protein-engineering tactics utilised to improve lysin activity. It also reviews the applications of lysins towards food biopreservation, therapeutics, biofilm elimination and diagnostics.

1.2 Introduction

Bacteriophages are viruses that can specifically target and infect bacterial cells without causing damage to cell lines from other organisms. These viruses have been employed in the treatment of bacterial infections for nearly a century (Wittebole *et al.,* 2014). Only recently did research into the use of phage-encoded recombinant endolysins (lysins), as potential therapeutic candidates, begin (Nelson *et al.,* 2001).

During the phage lytic replication cycle, progeny phages (Figure 1) are released from their host by the action of the lysin enzyme that degrades the host's peptidoglycan cell wall layer, subsequently leading to cell lysis and death. Lysin accumulates in the cytoplasm of the host (Fischetti, 2008) but can also cross the cytoplasmic membrane due to the action of another phage-encoded protein, designated holin, at a specific time point. A holin-independent secretory lysin containing a signal sequence has also been reported (Sao-Jose *et al.*, 2000). It was suggested that this lysin crosses the cytoplasmic membrane to the cell wall compartment using a specific regulatory system preventing premature cell lysis (Sao-Jose *et al.*, 2000).

Double-stranded DNA phages typically use lysin/holin for the lysis of host cells, whereas single-stranded RNA and DNA phages generally employ the expression of a single lysis gene encoding a small membrane protein (Young *et al.*, 2000). An example of this is the ϕX174 phage lysis protein E, a 91 amino acid membrane protein, which causes lysis by inhibiting the MraY enzyme involved in lipid I synthesis (Zheng *et al.,* 2009). Phage lysins possess a typical modular domain architecture, consisting of catalytic domain(s) and a cell wall binding domain (CWBD). Most lysins targeting the cell walls of Gram-negative bacteria do possess a single domain architecture (Oliveira *et al.*, 2014). The CWBD is thought to be responsible for targeting lysins to their bacterial cell wall substrate (Schmelcher *et al.*, 2010). This binding property has been exploited in various applications ranging from pathogen detection to the

isolation and differentiation of pathogenic bacteria from food sources (Schmelcher *et al.*, 2010). The therapeutic application of recombinant lysins in eliminating bacterial infections has also been achieved (Nelson *et al*., 2001). Lysins are generally active against the bacterial genera associated with the phage, *i.e.* a lysin originating from a streptococcal phage will specifically target streptococci (Fischetti, 2008). However, in some cases, phage lysins with broad lytic activity have been reported (Yoong *et al.*, 2004). For example, the enterococcal lysin PlyV12 is capable of lysing streptococci and staphylococci in addition to *Enterococcus faecalis* and *Enterococcus faecium* (Yoong *et al.*, 2004). In this case, it was suggested that the lysin recognises a common receptor across the different bacterial targets (Yoong *et al.*, 2004). Due to antibiotic resistance in many key pathogens, there is increased pressure for novel antimicrobials to replace the increasingly redundant traditional antibiotics. Lysins possess the potential to satisfy this role. Unlike antibiotics, bacterial resistance to this enzymes are rare (Loeffler *et al.*, 2001; Schuch *et al.*, 2002; Rodríguez-Rubio *et al.*, 2013), making these agents interesting therapeutic candidates for biocontrol of pathogenic bacteria. Lysins targeting many well-known infectious bacteria have been reported to-date including *Streptococcus* (Nelson *et al.*, 2001), *Staphylococcus* (O'Flaherty *et al.*, 2005), *Listeria* (Gaeng *et al.*, 2000), *Clostridium* (Mayer *et al.*, 2008) and *Bacillus* (Schuch *et al.*, 2002).

This review will focus on the current knowledge gained from the study of phage lysins, which includes: their structure and function; engineering tactics adopted to improve enzymatic function; lysin applications; and other phage encoded proteins associated with lysin activity.

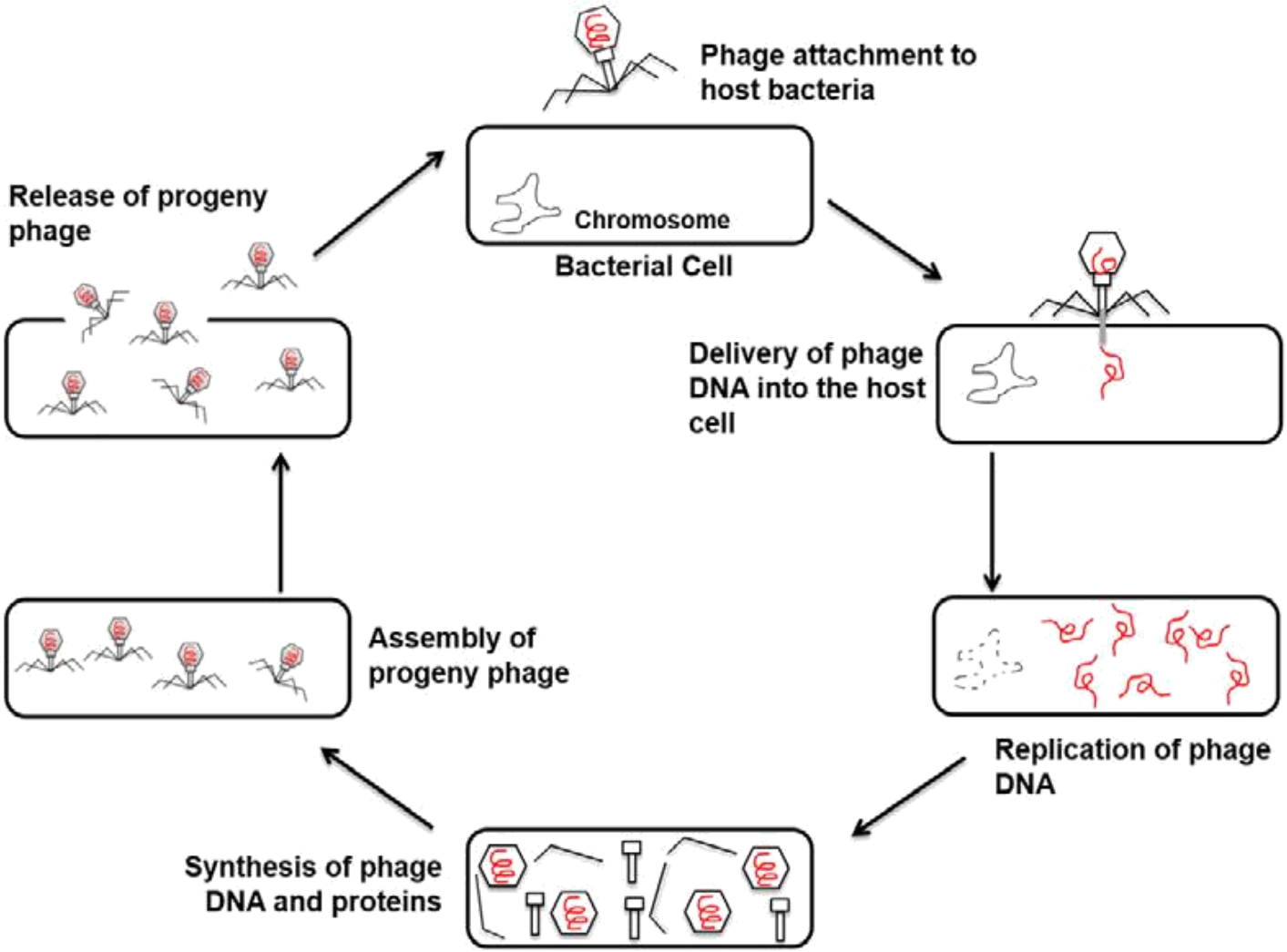


Figure 1: Bacteriophage replication cycle (Lytic phage).

1.3 Structure and function of phage lysins

Phage lysins in Gram-positive bacteria are generally comprised of multiple domains: typically one or more N-terminal catalytic domains and a C-terminal cell wall binding domain (CWBD). In contrast, the majority of lysins acting against Gram-negative bacteria usually have a globular structure, comprising of just the catalytic domain (Oliveira *et al.*, 2014). Although lysins of Gram-negative origin have been identified with more than one domain, this is uncommon (Walmagh *et al.*, 2012).

The CWBD of lysins serves as a binding function to specialised ligands within the bacterial cell wall, and is often linked with substrate recognition. The catalytic domain is responsible for the enzymatic hydrolysis of the peptidoglycan after recognition. Lysins with multiple domains are known to display linker, which bridge the catalytic and cell wall binding domains (Proença *et al.*, 2012; Tišáková *et al.*, 2014; Pohane, Patidar *et al.*, 2015). This linker contains an amino acid cleavage residue allowing for autoproteolytic cleavage of the C-terminal CWBD as reported for the clostridia lysin CTP1L (Dunne *et al.*, 2014) and CD27L (Mayer *et al.*, 2008).

1.3.1 Cell wall binding domain (CWBD)

The CWBD is responsible for recognising and binding to conserved modules within the bacterial cell wall, conferring specificity towards the lysin target. These targets include molecular structures like N-acetylglucosamine (Eugster and Loessner, 2012), choline (Hermoso *et al.*, 2003) and polyrhamnose (Lood *et al.*, 2014) as well as many other bacterial cell wall subunits. These components attach non-covalently to CWBDs with high affinity and specificity (Tišáková *et al.*, 2014).

Further demonstrating CWBD specificity, the C-terminal of the *Lactobacillus casei* lysin Lc-Lys could specifically target bacterial strains of peptidoglycan containing an amidated D-Asn cross bridge, eventually leading to cell lysis but when tested against cell mutants with modified cell wall, lytic activity was completely abolished (Regulski *et al.*, 2013). This supports the suggestion that the lytic activity of lysins acting against strains of related species is due to binding of CWBD to a specifically conserved epitope in the cell wall (Proença *et al.*, 2012). As such, lysins without a CWBD tend to have a broad antibacterial host range in contrast to those containing a CWBD, which exhibit a narrow host range (Plotka *et al.*, 2014). In some cases, the CWBD is crucial for full enzymatic activity against the lysin substrate (Korndörfer *et al.*, 2006; Kikkawa *et al.*, 2008), as removal of such domain resulted in loss of lysin catalytic activity as reported with the *Bacillus anthracis* lysin PlyG (Kikkawa *et al.*, 2008).

X-ray crystallography has been used to determine the 3D structure of several CWBDs. These included the *Listeria monocytogenes* lysin PlyPSA, whose CWBD revealed a unique fold with its structural motif displaying a pronounced hydrophobic cleft consisting of aromatic side chain residues at the interface of the lysin's two subdomains, which was suggested to be involved in substrate recognition (Korndörfer *et al.*, 2006). A similar structure based determination of the CWBD of pneumococcal lysin Cpl-1 revealed a choline binding motif that facilitates anchoring onto choline-containing teichoic acid of the pneumococcal cell wall (Hermoso *et al.*, 2003).

1.3.2 Endolysin catalytic domain

The catalytic domain brings about the hydrolytic degradation of the peptidoglycan cell wall, specifically targeting its conserved bonds (Schmelcher *et al.*, 2012). Lysins can be classified into five major groups depending on the cell wall peptidoglycan bonds they cleave (Figure 2).

These groups include: (a) N-acetyl-β-D-acetylmuramoyl-L-alanine amidases that cleaves the amide bonds between N-actylmuramic acid and the first L-alanine (Schmelcher *et al.*, 2012); (b) N-acetyl-β-D-muramidases; (c) Lytic transglycosylases, both of which are involved in the cleavage of glycosidic linkage between N-acetylmuramic acid and N-acetyl-D-glucosamine (Vasala *et al.*, 1995); (d) N-acetyl-glucosaminidases which cleaves the other glycosidic bonds (Schmelcher *et al.*, 2012); and (e) endopeptidases, involved in the cleavage of peptide bonds at the D-alanyl-glycyl moieties (Donovan *et al.*, 2006).

Several lysins are known to contain two catalytic domains. An example is the staphylococcal phage lysin LysK whose catalytic domain harbours a cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) domain as well as an amidase counterpart (O'Flaherty *et al.*, 2005). CHAP domains typically contain a cysteine and histidine residues at the active site (Bateman and Rawlings, 2003; Rigden *et al.*, 2003). Other examples with two catalytic domains include the mycobacteriophage lysins, which often contain a central catalytic domain, an N-terminal domain predicted to possess peptidase activity and an associate cell wall recognition motif (Payne and Hatfull, 2012).

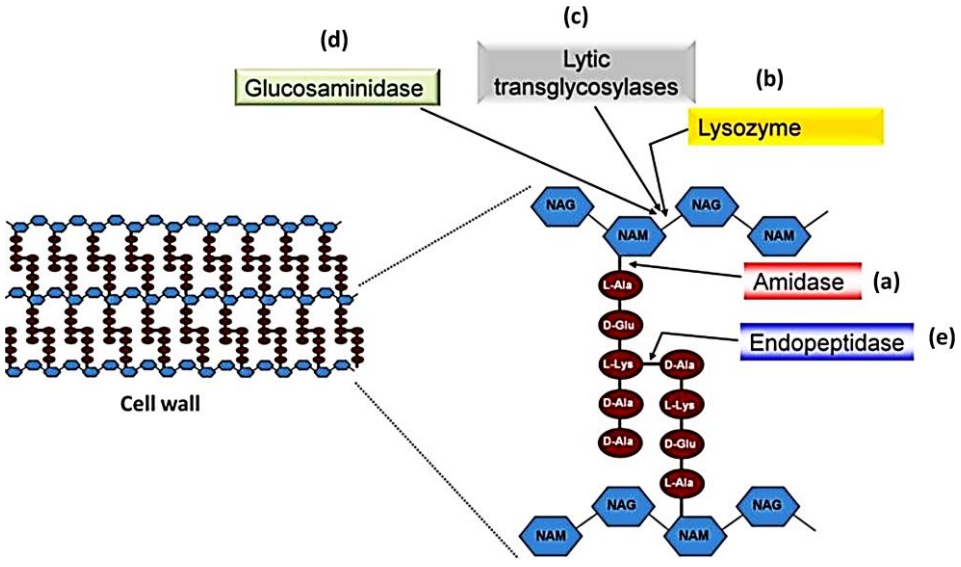


Figure 2: Schematic representation of a Gram-positive bacterial cell wall structure depicting endolysin cleavage sites within the peptidoglycan

1.3.3 Lysin activity

Lysins have the potential for use as therapeutic agents due to their antibacterial properties and this has been exploited in a variety of studies ranging from the elimination/control of drug resistant bacteria (Horgan *et al.*, 2009; Jun *et al.*, 2013; Briers *et al.*, 2014; Huang *et al.*, 2014) to the elimination of bacterial biofilms (Sass and Bierbaum, 2007; Gutiérrez *et al.*, 2014; Díez-Martínez *et al.*, 2015)**.** Cell lysis by the exogenous application of lysins is more easily performed in the case of Gram-positive bacteria compared to Gram negatives (due to the presence of the outer membrane in Gram-negative bacteria). Pretreatment of Gram-negative cells with EDTA significantly increases the permeability of the outer membrane, thereby exposing the cell wall to the hydrolytic effect of lysins (Son *et al.*, 2012)**.** However, some lysins have been shown to possess lytic activity against Gram-negative cells without the need of an osmotic permeabiliser (Lim *et al.*, 2014). It is suggested that the C-terminal region of such lysins could be responsible, as it enhances the permeability of the bacterial outer membrane aiding the N-terminal enzymatic domain in reaching its peptidoglycan target (Lim *et al.*, 2014). The modified so-called artilysins are other examples of lysins with the ability to penetrate the outer membrane of Gram-negative bacteria without the need of an osmotic permeabiliser (Briers *et al.*, 2014). These enzymes constitute a novel class of antibacterial enzyme (Briers *et al.*, 2014).

Recently, the staphylococcal lysin 2638A was reported with an unusual activity (Abaev *et al.*, 2013), as the amidase domain was more active than its peptidase counterpart. This was reported by Abaev and co-workers (Abaev *et al.*, 2013) to be in direct contrast with lysin possessing similar domain architecture such as the staphylococcal lysins LysK (Horgan *et al.*, 2009) and phi11 (Sass and Bierbaum, 2007), where both CHAP domains were reported to have higher lytic activities than their amidase counterparts (Donovan *et al.*, 2006; Sass and Bierbaum, 2007; Horgan *et al.*, 2009; Abaev *et al.*, 2013). A *Salmonella* phage lysin

designated SPN1S with superior lytic activity to the non-phage-derived cell wall-degrading enzyme lysozyme has also been reported. This enzyme, containing a lysozyme-like catalytic domain, had a 30-fold increase in lytic activity over the chicken egg white lysozyme (Lim *et al.*, 2012). Interestingly, two individual lysins with only a 3 amino acid difference between their protein sequences exhibited a significant difference in their cell wall hydrolysing activities despite their high degree of similarities (Jun *et al.*, 2011). This indicates that certain amino acid residues play a key role in the overall catalytic function of lysins (Sanz-Gaitero *et al.*, 2013). Some lysins can have broad spectrum lytic activity, as demonstrated by Lai and co-workers (Lai *et al.*, 2011) who reported that the *Acinetobacter baumannii* phage lysin LysAB2 was capable of effectively lysing seven different bacterial genera including *Staphylococcus aureus*, *Bacillus subtilis*, *Streptococcus sanguis*, *Acinetobacter baumannii*, *Escherichia coli*, *Citrobacter freundii* and *Salmonella enterica* (Lai *et al.*, 2011). Lysins with different cleavage specificities have been shown to demonstrate synergy with each other against infectious bacteria, both *in vitro* and in mouse models (Jado *et al.*, 2003; Loeffler and Fischetti, 2003).

Synergistic effects have also been demonstrated with other antibacterial agents such as nisin (García *et al.*, 2010), lysostaphin (Becker *et al.*, 2008) and antibiotics (Rashel *et al.*, 2007; Rodríguez-Cerrato *et al.*, 2007). It is believed that such synergy arises from the cleavage of peptidoglycan at two different recognition sites leading to an increased overall activity (Linden *et al.*, 2015). Microscopic visualisation of lysin lytic activity has been demonstrated to portray physical changes experienced by bacterial cells in response to the lysin treatment. Treating a streptomycin-resistant *B. cereus* strain RSVF1 with the *B. anthracis* phage lysin PlyG resulted in normal filamentous RSVF1 being converted to short rod and mini-cell like forms, after 30 seconds of exposure (Schuch *et al.*, 2002).

1.3.4 Resistance to endolysins

Most bacteria have the capacity to develop resistance mechanisms to protect themselves against the action of antibacterial agents. These mechanisms include the modification of cell wall components, efflux pump overexpression, enzyme modification and porins (Martins *et al.,* 2013; Cullen and McClean, 2015). However, no resistance mechanism has been reported for phage lysins to date. Repeated exposure of *Streptococcus pneumononiae* grown on agar plates containing low concentrations of the pneumococcal lysin Pal did not lead to resistant strains (Schuch *et al.,* 2002). Neither did the successive exposure of *S. aureus* to subinhibitory concentrations of the staphylococcal lysin LysH5 (Rodríguez-Rubio *et al.,* 2013). However, similar exposures to other antibacterial agents resulted in the generation of mutants resistant to lysostaphin, novobiocin and streptomycin (Schuch *et al.,* 2002; Rodríguez-Rubio *et al.*, 2013). This suggests that bacteriophage evolved their lysins over the millennia by targeting highly conserved, central modules in the bacterial cell wall, thereby making bacterial resistance to lysins a very rare event (Fischetti, 2008).

1.4 Phage-encoded proteins associated with lysins

The principal phage-encoded proteins associated with lysins include the holins, signal peptides and spanins.

1.4.1 Holins

During phage replication, lysins move across the cytoplasmic membrane to degrade the peptidoglycan target aided by a membrane protein designated holin. These proteins accumulate in the cytoplasmic membrane of the host bacteria leading to lesion formation in the cytoplasmic membrane, thereby controlling lysins' access to the peptidoglycan (Young, 1992). Depending on their membrane topology, holins fall into one of three different classifications based on the number of transmembrane domains (TMD) they possess (Young,

2002). The most studied holin is the λ holin gene designated S, which encodes two distinct proteins termed S105 and S107, differing in their protein sequence by the first two amino acids (Bläsi *et al.,* 1989). The C-terminal domain of the λ holin was reported not to be necessary for formation of inner membrane lesions but rather has a regulatory role in the proper scheduling of hole-forming events (Bläsi *et al.*, 1999).

The first *in vitro* study of a phage encoded holin was reported for the λ holin and this involved using purified holin to form lesions in artificial liposome (Smith *et al.,* 1998). It was reported that the λ holin also directly interacted with the antiholin in the bacterial membrane by forming heterodimer under oxidative condition (Gründling *et al.,* 2000). Green fluorescence protein (GFP) fusion study revealed that these proteins accumulate uniformly in the cytoplasmic membrane forming aggregates or rafts in the membrane (White *et al.*, 2011). Antiholins were also revealed to block lysis by preventing this raft formation (White *et al.*, 2011).

1.4.2 Signal sequences

Evidence of a signal sequence in the N-terminal of lysins has been reported (Sao-Jose *et al.*, 2000; Kakikawa *et al.,* 2002). The first experimental evidence of these secretory lysins was from Sao-Jose and co-workers (Sao-Jose *et al.,* 2000). In this case, expression of the oenococcal lysin Lys44 resulted in the generation of two polypeptides revealed as precursor and mature forms of the enzyme. Supporting evidence was also reported for the *Lactobacillus fermentum* phage lysin Lyb5 (Guo *et al.,* 2015). Here, chimeric linkage between the N-terminal of the lysin and the nucB gene from *S. aureus* resulted in the export of NucB protein into the surrounding environment following gene expression in *L. lactis*. Moreover, expression of Lyb5 secretory lysin in *E. coli* also resulted in morphological changes as the normal rod-shaped *E. coli* adopted a spherical shape 20mins post induction. It was thus

suggested that the morphological change was due to export of lysin to the cell wall (Guo *et al.*, 2015).

An experimental assay to examine the production of secretory lysin during phage infection was reported for the fOg44 phage (Sao-Jose *et al.*, 2000). Immunoblot analysis revealed that the mature lysin was first detected at 80mins post-infection. It was thus suggested that a regulatory mechanism must be operational to down-regulate lytic activity of secreted lysin during the latent period, which was determined in this case to be 150mins postinfection. It was also suggested that maturation of the secretory lysin Lys44 was dependent on the SecA general secretion pathway (Sao-Jose *et al.*, 2000).

1.4.3 Spanins

A third class of lysis proteins, designated spanins, was also identified (Summer *et al.*, 2007). These proteins were composed of an outer membrane lipoprotein with a C-terminal transmembrane domain capable of integrating into the inner membrane. The best-characterised spanins are the lambda Rz and Rz1 proteins. These were suggested to interact forming a complex, which spans the entire periplasm (Berry *et al.*, 2010). It was also reported that the Rz protein was unstable *in vivo* in the absence of Rz1 and required complex formation with Rz1 to prevent proteolysis. It was recently reported that the spanin complex was essential for lambda lysis, as expression of lysogens carrying the lambda holin and endolysin genes as well as a null mutant spanin did not result in cell lysis, but rather led to the development of fragile spherical cells. It was thus suggested that spanins carried out an essential step in outer membrane disruption, in a manner regulated by the state of the peptidoglycan layer (Berry *et al.*, 2012).

1.5 Protein engineering

Various protein-engineering techniques have been utilised on lysins to modify their activity. These include domain swapping and shuffling, lysin mutagenesis and other modifications leading to active translocation of lysins.

1.5.1 Domain swapping and shuffling

The modular structure of lysins endows them with the potential for domain swapping and shuffling, which has been exploited in engineering lysins in several reports. For example, an improved version of a pneumococcal lysin was developed, following chimeric linkage between the catalytic domain of an engineered variant of pneumococcal lysin (Díez-Martínez *et al.*, 2013) and the CWBD of another pneumococcal lysin (Díez-Martínez *et al.*, 2015). The resulting chimeric lysin showed increased bactericidal activity over the parent enzyme. In contrast, another chimeric lysin resulting from the fusion of a *Clostridium sporogenes* catalytic domain and *Clostridium difficile* CWBD showed reduced lytic efficiency against *Clostridium tyrobutyricum* compared with the parent *C. sporogenes* lysin (Mayer *et al.*, 2012).

The chimeric linkage between the catalytic domain of phage lysin and the CWBD of the antibacterial peptidase enzyme lysostaphin has also been reported (Schmelcher *et al.*, 2012). Here, the resulting chimeric enzyme was capable of controlling *S. aureus* mastitis and could also reduce the bacterial load in mouse models in addition to possessing a synergistic effect with the parent lysostaphin, thus demonstrating the potential of chimeric lysins as potential antimicrobials. Protein engineering studies have also been utilised to improve the thermostability of lysins. For example, by replacing the CWBD of a *Clostridium perfringens* lysin with that of another lysin originating from a thermophilic phage, an engineered lysin with improved thermostability was created (Swift *et al.*, 2015). In addition, the poor

solubility of a staphylococcal phage lysin resulting in inadequate large-scale production and purification of such lysin was improved by protein engineering studies (Fernandes *et al.*, 2012). The resulting chimeric enzyme, composed of CWBD of a staphylococcal phage lysin together with the highly soluble catalytic domain of an enterococcal phage (Proença *et al.*, 2012), did not only possess improved solubility but also had a broad lytic activity against a range of staphylococcal strains including streptococci and enterococci (Fernandes *et al.*, 2012).

Not only does domain swapping improve the lytic activity of lysin catalytic domains, experimental evidence also suggests that lysin-binding properties can also be affected by domain shuffling (Schmelcher *et al.*, 2011). Supporting data on domain shuffling showed that substituting the CWBD of the *Listeria* lysin Ply118 with that of PlyPSA resulted in an abolished lytic activity towards *Listeria* strains of serovar 1/2, while enhancing its lytic activity towards serovar 4. This is an interesting finding, as the native Ply118 lysin could only target the cell wall of *Listeria* serovar strains 1/2, while PlyPSA could also target those of serovar 4 (Schmelcher *et al.*, 2011).

1.5.2 Mutagenesis

Mutagenesis studies have also been employed in an effort to improve lysin activity. These studies usually employ amino acid substitution(s) and/or deletions. For example, an enhanced bactericidal activity of the pneumococcal phage lysin Cpl-7 was achieved following a 15 amino acid substitution in its CWBD (Díez-Martínez *et al.*, 2013). This substitution also resulted in an inversion of the lysin's net charge at neutral pH from – 14.93 to +3. Using a similar approach, a CWBD-dependent catalytic domain was also converted to a CWBD-independent enzyme (Low *et al.*, 2011). This study suggested that a positive net charge was a requirement for the lytic activity of lysins without its cognate CWBD. Moreover, it was

suggested that altering the net charge on the catalytic domain could bring about a refinement or increase in the host range of lysins (Low *et al.*, 2011).

The influence of deleting the CWBD on the lytic activity of lysin was studied. Interestingly, this was associated with variable effects. While CWBD deletion dramatically improved lysis in some cases, it either reduced or abolished activity in others. These effects are most likely due to the change in charge of the truncated lysin (Low *et al.*, 2011) as it is known that many Gram-positive bacteria do possess a negatively charged surface component, facilitating the action of small cationic antibacterial in the disruption of the bacterial cell (Oyston *et al.*, 2009).

1.5.3 Lysin translocation

Protein engineering studies involving the active translocation of lysins across the bacterial membrane have been undertaken. As protein secretion involves the attachment of a signal peptide (containing a positively charged N-terminal region, a hydrophobic core and a C-terminal cleavage site) to the protein under secretion (Auclair *et al.*, 2012). This is vital for the active translocation of the attached protein through the cell membrane of the host following expression. Gaeng and co-workers revealed that by attaching the *Lactobacillus brevis* S-layer protein signal peptide to the *Listeria monocytogenes* phage lysin A511, active translocation of the lysin from the *Lactococcus lactis* host cells to the surrounding environment was possible (Gaeng *et al.*, 2000). This was demonstrated experimentally as the lysin-secreting *L. lactis* brought about a zone of inhibition around the recombinant *L. lactis* in agar medium embedded with heat-inactivated *L. monocytogenes.* However, recombinant lysin-secreting lactic acid bacteria (LAB) showed poor antimicrobial activity against viable bacterial cells in an *in vitro* coculture assay (Turner *et al.*, 2007). This was suggested to be related to the growth rate of the LAB strain affecting the production rate of the secreted lysin.

A similar approach in bringing about the active translocation of the *Clostridium perfringens* lysin CP25L to its surrounding environment has also been performed (Gervasi *et al.*, 2014). Here, the CP25L lysin was capable of lysing *C. perfringens* cells in complex media designed to simulate the conditions of the gastrointestinal (GI) tract. Given that, the CP25L lysin did not lyse other members of the gut microflora tested; this suggested that lysins could have the potential to control specific pathogenic strains of bacteria residing in the gut, assuming secretion was adequate by the relevant recombinant bacterial delivery system. Codon optimisation could be an interesting avenue in bringing about increased secretion efficiency leading to higher bactericidal activity of secreted lysins (Rodríguez-Rubio *et al.*, 2012). This was demonstrated by Rodríguez-Rubio and co-workers, where codon optimisation of a gene encoding a signal peptide and lysin based on an *L. lactis* codon usage resulted in an increased activity of the secreted lysin (Rodríguez-Rubio *et al.*, 2012).

1.6 Applications of lysins

The lytic capacity of phage lysins in the control of bacteria endows them with various potential applications. These applications ranging from food preservation to pathogen detection ultimately utilise either the lysin's peptidoglycan hydrolytic action or its (CWBD) binding function to achieve its end goal.

1.6.1 Food biopreservation

Numerous studies have demonstrated the ability of phage lysins to function as a preservative agent in the control of foodborne pathogens posing a major threat to the health and wellbeing of individuals, especially the elderly and the immunocompromised. Zhang and co-workers provided experimental evidence for the ability of the *Listeria monocytogenes* phage lysin LysZ5 to successfully control *L. monocytogenes* to undetectable levels in soya milk (Zhang *et al.*, 2012). The lysin was also capable of controlling *L. monocytogenes* at refrigeration

temperature. It was also reported that 45U/ml of the staphylococcal lysin LysH5 was sufficient in eliminating *S. aureus* in milk at a contamination level of 10^3 CFU/ml (Obeso *et al.*, 2008). The enzyme also exhibited synergy with the bacteriocin nisin at low concentrations resulting in complete elimination of *S. aureus* in milk (García *et al.*, 2010). This combination presents a potential food preservative in the control of food pathogens. Also, investigations of a lysin, formulated with silica nanoparticle with the ability to control bacterial growth in lettuce have been performed (Solanki *et al.*, 2013).

1.6.2 Lysins as therapeutics

Lysin technology represents an alternative therapeutic approach for the control of pathogenic bacteria involved in a variety of animal and human infections. Lysins differ from antibiotics as there is little to no chance of the development of bacterial resistance. This is because lysins generally target conserved bonds within the peptidoglycan structure as mentioned earlier. Lysins' ability to combat pathogenic bacteria *in vitro* and *in vivo* (mouse models) has been demonstrated in several laboratories, with the first *in vivo* experiments reported by the group of Fishetti (Nelson *et al.*, 2001). Here, the streptococcal lysin PlyC was capable of providing protection against *Streptococcus pyogenes* colonisation following bacterial challenge in a mouse model. Several other *in vivo* experiments utilising lysins in the control of infectious bacterial pathogens residing in the nasal cavity have also been reported. These include studies performed by Loeffler and co-workers (Loeffler *et al.*, 2001), Rashel and co-workers (Rashel *et al.*, 2007), Daniel and co-workers (Daniel *et al.*, 2010) as well as Fenton and co-workers (Fenton *et al.*, 2010).

Studies involving *in vivo* applications of lysins in the control of infectious bacteria in other anatomical locations of mouse models have also been reported. One such study reports the treatment of an *S. aureus* induced endophthalmitis by the lysin ply187 (Singh *et al.*, 2014). In

this work, a single intravitreal injection of the enzyme at 6hrs post infection drastically reduced bacterial load in the mice's eyes. This also provided a protective effect on the retina at the tissue level (Singh *et al*,, 2014). In another study, an intraperitoneal injection of *Enterococcus faecalis* in mouse sepsis model also revealed that the lysin IME-EF1 was capable of providing better protection against infectious *E. faecalis* compared to its producing phage (Zhang *et al.*, 2013). Topical skin application of lysin has also been reported, where a chimeric lysin ClyS (Daniel *et al.*, 2010) was found to be effective for bacterial decolonisation from mice infected skin. In this case, the lysin formulated in ointment had a better decolonisation effect compared to the standard topical antibacterial agent mupirocin (Pastagia *et al.*, 2011). For respiratory infections, it has also been shown that the pneumococcal lysins Cpl-1 could be delivered to the respiratory airway in aerosolised format to combat pneumococcal lung infections (Doehn *et al.*, 2013). This was demonstrated in a mouse model, where aerosolised Cpl-1 significantly reduced bacterial load in the lung, thus protecting the mice from pneumococcal bacteraemia. Other *in vivo* studies focussed on the zebrafish embryo infection model (Díez-Martínez *et al.*, 2013). Here, the engineered pneumococcal lysins Cpl-7s improved the survival rate of zebrafish embryo.

Applications of endolysins in animals or humans obviously necessitate the undertaking of safety studies. Accordingly, the first GLP-compliant toxicology and safety study of a phage lysin revealed no sign of toxicity or adverse effect in rats in a trial carried out by Jun and co-workers (Jun *et al.*, 2014). Although some side effects were recorded when lysin administration was continued for more than one week in dogs, these were resolved within 1hr and were suggested to be due to immune response to the lysin (Jun *et al.*, 2014). Studies such as this will advance the use of lysins as therapeutic candidates in the control of pathogenic bacteria in animals and humans.

1.6.3 Biofilm elimination by lysins

An important feature of many pathogenic bacteria is their ability to form biofilms, resulting in their tolerance to many antimicrobial agents (Otto, 2008; Sanchez-Vizuete *et al.*, 2015) and lysins possess the potential to eliminate these structures. The most frequently recognised causative agents of biofilm-associated infections are the staphylococci (Otto, 2008) and lysins with the ability to disrupt their associated biofilms have been reported. Sass and Bierbaum provided experimental evidence that the phi11 lysin was capable of eliminating *S. aureus* biofilms (Sass and Bierbaum, 2007). The lysin was also suggested to destabilise biofilm structure by rapid lysis of sessile cells embedded within extracellular matrix. Another phage-encoded lysin reported to eliminate staphylococcal biofilm was LysH5. This enzyme was capable of reducing bacterial population in biofilms formed by either *S. aureus* or *S. epidermidis* including persister cells (a bacterial subpopulation that show multidrug resistance) (Gutiérrez *et al.*, 2014). It was also reported that subinhibitory concentrations of this enzyme completely inhibited staphylococcal biofilm for some of the strains tested in this study. Other lysins reported to eliminate staphylococcal biofilms include SAL-2 (Son *et al.*, 2010), $CHAP_k$ (Fenton *et al.*, 2013), SAL-1 (Jun *et al.*, 2013), PlyGRCS (Linden *et al.*, 2015) and Ply187 (Singh *et al.*, 2014).

1.6.4 Diagnostic applications

Pioneering work in the laboratory of Loessner has shown that lysins also have a potential application in the detection and quantification of bacterial pathogens in food materials (Schmelcher *et al.*, 2012). Essentially, the lysin's CWBD with its affinity for specific cell wall structures in the host bacterium has been exploited in a few bacterial genera, namely *Listeria*, *Bacillus* and *Clostridium* (Kretzer *et al.*, 2007). Indeed to date, a variety of bacterial detection technologies involving the CWBD have been reported. One involved the use of fluorescent protein attached to lysin's CWBD (Schmelcher *et al.*, 2010; Gerova *et al.*, 2011).

Another approach incorporated the development of CWBD-based surface plasmon resonance (SPR) technology (Kong *et al.*, 2015). Here, CWBD was genetically engineered by attaching glutathione S-transferase to its N-terminal. This allowed immobilisation of the engineered CWBD unto glutathione chips. The use of paramagnetic beads coated with endolysin-derived CWBD proteins in the development of immobilisation and magnetic separation technology has also been reported (Kretzer *et al.*, 2007).

The detection technologies mentioned above have allowed for several practical applications in the use of CWBD for detection of bacterial pathogens. For example, CWBD immobilisation onto a glutathione chip allowed for specific and quantitative detection of *Bacillus cereus* and the SPR response intensity was significantly higher than that of antibody-based chip used in comparison (Kong *et al.*, 2015). Also, milk contaminated with *L. monocytogenes* was detected using paramagnetic beads coated with CWBD-derived proteins. The average recovery rates recorded for both plating and real time PCR based detection was 97.8% and 70.1%, respectively (Walcher *et al.*, 2010).

1.6.5 Other applications of lysins

Lysins also have a potential application as narrow spectrum disinfectants and this has been investigated by Hoopes and co-workers (Hoopes *et al.*, 2009). The streptococcal lysin PlyC was reported as the first protein-based narrow-spectrum disinfectant against *Streptococcus equi*. The enzyme was also reported to be 1,000 times more active than the commonly used disinfectant virkon S as 1µg of the enzyme sterilised 10^8 CFU/ml of *S. equi* culture in 30mins (Hoopes *et al.*, 2009).

Lysins have also been reported as antimicrobial candidates for the control of lactic acid bacterial contaminations in fuel ethanol fermentation (Roach *et al.*, 2013). Here, the streptococcal lysin λSa2 was reported to exhibit lytic activity against majority of LAB tested.

This enzyme was also capable of reducing *L. fermentum* in a mock fermentation of corn fiber hydrolysate (Roach *et al.*, 2013).

1.7 Conclusion

Lysins have increased potential as effective antibacterial agents against infectious pathogens. Their specific nature makes these enzymes and/or their phages good candidates to complement increasingly redundant antibiotic therapy, but in an approach that is far more specific than antibiotics. The application of protein engineering has the potential to significantly improve lysin activity for various biotechnological applications.

Table 1: Typical applications of recombinant phage lysins

Lysin name	Application	Reference
LysZ5	Controlled *Listeria monocytogenes* in soy milk	Zhang *et al.*, 2012
LysH5	Acted in Synergy with nisin to control *Staphylococcus aureus* in milk	García *et al.*, 2010
Cpl-7S	Reduced population of *Streptococcus pneumoniae* in infected zebrafish model providing a 99% survival rate	Díez-Martínez *et al.*, 2013
Cpl-1	Protected mice model infected with *Streptococcus pneumonia* in aerosolised form	Doehn *et al.*, 2013
$CHAP_k$	Completely eliminated *S. aureus* in nares of mice models as displayed in *In vivo* imaging system (IVIS)	Fenton *et al.*, 2010
SAL-1	Preformulated as SAL-200 with lysin as active pharmaceutical ingredient	Jun *et al.*, 2013
PBC1	Utilised CWBD in the development of surface plasmon resonance (SPR) technology	Kong *et al.*, 2015
Ply500 and Ply118	Utilised CWBD in development of magnetic separation technology for immobilisation and separation of bacterial cells	Kretzer *et al.*, 2007
λSa2	Controlled *Lactobacillus fermentum* contaminate in a mock fermentation of corn fibre hydrolysate	Roach *et al.*, 2013
Ply500	Covalent attachment to silica nano particles allowed for decontamination of *Listeria innocua* on iceberg lettuce	Solanki *et al.*, 2013

1.8 References

1. Abaev, I., Foster-Frey, J., Korobova, O., Shishkova, N., Kiseleva, N., Kopylov, P., Pryamchuk, S., Schmelcher, M., Becker, S. C. and Donovan, D. M. (2013) 'Staphylococcal Phage 2638A endolysin is lytic for *Staphylococcus aureus* and harbors an inter-lytic-domain secondary translational start site', *Applied Microbiology and Biotechnology*, 97(8), pp. 3449–3456. doi: 10.1007/s00253-012-4252-4.
2. Auclair, S. M., Bhanu, M. K. and Kendall, D. A. (2012) 'Signal peptidase I: Cleaving the way to mature proteins', *Protein Science*, 21(1), pp. 13–25. doi: 10.1002/pro.757.
3. Bateman, A. and Rawlings, N. D. (2003) 'The CHAP domain: A large family of amidases including GSP amidase and peptidoglycan hydrolases', *Trends in Biochemical Sciences*, pp. 234–237. doi: 10.1016/S0968-0004(03)00061-6.
4. Becker, S. C., Foster-Frey, J. and Donovan, D. M. (2008) 'The phage K lytic enzyme LysK and lysostaphin act synergistically to kill MRSA', *FEMS Microbiology Letters*, 287(2), pp. 185–191. doi: 10.1111/j.1574-6968.2008.01308.x.
5. Berry, J., Rajaure, M., Pang, T. and Young, R. (2012) 'The spanin complex is essential for lambda lysis', *Journal of Bacteriology*, 194(20), pp. 5667–5674. doi: 10.1128/JB.01245-12.
6. Berry, J., Savva, C., Holzenburg, A. and Young, R. (2010) 'The lambda spanin components Rz and Rz1 undergo tertiary and quaternary rearrangements upon complex formation', *Protein Science*, 19(10), pp. 1967–1977. doi: 10.1002/pro.485.
7. Bläsi, U., Fraisl, P., Chang, C. Y., Zhang, N. and Young, R. Y. (1999) 'The C-terminal sequence of the λ holin constitutes a cytoplasmic regulatory domain', *Journal of Bacteriology*, 181(9), pp. 2922–2929.
8. Bläsi, U., Nam, K., Hartz, D., Gold, L. and Young, R. (1989) 'Dual translational initiation sites control function of the lambda S gene.', *The EMBO Journal*, 8(11), pp. 3501–3510.
9. Briers, Y., Walmagh, M., Grymonprez, B., Biebl, M., Pirnay, J. P., Defraine, V., Michiels, J.,

Cenens, W., Aertsen, A., Miller, S. and Lavigne, R. (2014) 'Art-175 is a highly efficient antibacterial against multidrug-resistant strains and persisters of *Pseudomonas aeruginosa*', *Antimicrobial Agents and Chemotherapy*, 58(7), pp. 3774–3784. doi: 10.1128/AAC.02668-14.

10. Cullen, L. and McClean, S. (2015) 'Bacterial adaptation during chronic respiratory infections.', *Pathogens (Basel, Switzerland)*, 4(1), pp. 66–89. doi: 10.3390/pathogens4010066.
11. Daniel, A., Euler, C., Collin, M., Chahales, P., Gorelick, K. J. and Fischetti, V. a. (2010) 'Synergism between a novel chimeric lysin and oxacillin protects against infection by methicillin-resistant *Staphylococcus aureus*', *Antimicrobial Agents and Chemotherapy*, 54(4), pp. 1603–1612. doi: 10.1128/AAC.01625-09.
12. Díez-Martínez, R., De Paz, H. D., García-Fernández, E., Bustamante, N., Euler, C. W., Fischetti, V. a., Menendez, M. and García, P. (2015) 'A novel chimeric phage lysin with high *in vitro* and *in vivo* bactericidal activity against *Streptococcus pneumoniae*.', *The Journal of Antimicrobial Chemotherapy*, 70(6), pp. 1763–73. doi: 10.1093/jac/dkv038.
13. Díez-Martínez, R., De Paz, H., Bustamante, N., García, E., Menéndez, M. and García, P. (2013) 'Improving the lethal effect of Cpl-7, a pneumococcal phage lysozyme with broad bactericidal activity, by inverting the net charge of its cell wall-binding module', *Antimicrobial Agents and Chemotherapy*, 57(11), pp. 5355–5365. doi: 10.1128/AAC.01372-13.
14. Doehn, J. M., Fischer, K., Reppe, K., Gutbier, B., Tschernig, T., Hocke, A. C., Fischetti, V. A., Löffler, J., Suttorp, N., Hippenstiel, S. and Witzenrath, M. (2013) 'Delivery of the endolysin Cpl-1 by inhalation rescues mice with fatal pneumococcal pneumonia', *Journal of Antimicrobial Chemotherapy*, 68(9), pp. 2111–2117. doi: 10.1093/jac/dkt131.
15. Donovan, D. M., Lardeo, M. and Foster-Frey, J. (2006) 'Lysis of staphylococcal mastitis

pathogens by bacteriophage phi11 endolysin', *FEMS Microbiology Letters*, 265(1), pp. 133–139. doi: 10.1111/j.1574-6968.2006.00483.x.

16. Dunne, M., Mertens, H. D. T., Garefalaki, V., Jeffries, C. M., Thompson, A., Lemke, E. A., Svergun, D. I., Mayer, M. J., Narbad, A. and Meijers, R. (2014) 'The CD27L and CTP1L endolysins targeting clostridia contain a built-in trigger and release factor', *PLoS Pathogens*. Edited by G. Zhang, 10(7), p. e1004228. doi: 10.1371/journal.ppat.1004228.
17. Eugster, M. R. and Loessner, M. J. (2012) 'Wall teichoic acids restrict access of bacteriophage endolysin Ply118, Ply511, and Plyp40 cell wall binding domains to the *Listeria monocytogenes* peptidoglycan', *Journal of Bacteriology*, 194(23), pp. 6498–6506. doi: 10.1128/JB.00808-12.
18. Fenton, M., Casey, P. G., Hill, C., Gahan, C. G., Ross, R. P., McAuliffe, O., O'Mahony, J., Maher, F. and Coffey, A. (2010) 'The truncated phage lysin CHAP(k) eliminates *Staphylococcus aureus* in the nares of mice.', *Bioengineered Bugs*, 1(6), pp. 404–407. doi: 10.4161/bbug.1.6.13422.
19. Fenton, M., Keary, R., McAuliffe, O., Ross, R. P., O'Mahony, J. and Coffey, A. (2013) 'Bacteriophage-derived peptidase CHAP(K) eliminates and prevents staphylococcal biofilms.', *International Journal of Microbiology*, 2013, p. 625341. doi: 10.1155/2013/625341.
20. Fernandes, S., Proença, D., Cantante, C., Silva, F. A., Leandro, C., Lourenço, S., Milheiriço, C., de Lencastre, H., Cavaco-Silva, P., Pimentel, M. and São-José, C. (2012) 'Novel chimerical endolysins with broad antimicrobial activity against methicillin-resistant *Staphylococcus aureus*.', *Microbial Drug Resistance (Larchmont, N.Y.)*, 18(3), pp. 333–343. doi: 10.1089/mdr.2012.0025.
21. Fischetti, V. A. (2008) 'Bacteriophage lysins as effective antibacterials', *Current Opinion in Microbiology*, 11(5), pp. 393–400. doi: 10.1016/j.mib.2008.09.012.

22. Gaeng, S., Scherer, S., Neve, H. and Loessner, M. J. (2000) ‘Gene cloning and expression and secretion of *Listeria monocytogenes* bacteriophage-lytic enzymes in *Lactococcus lactis*.’, *Applied and Environmental Microbiology*, 66(7), pp. 2951–2958. doi: 10.1128/AEM.66.7.2951-2958.2000.

23. García, P., Martínez, B., Rodríguez, L. and Rodríguez, A. (2010) ‘Synergy between the phage endolysin LysH5 and nisin to kill *Staphylococcus aureus* in pasteurized milk’, *International Journal of Food Microbiology*. Elsevier B.V., 141(3), pp. 151–155. doi: 10.1016/j.ijfoodmicro.2010.04.029.

24. Gerova, M., Halgasova, N., Ugorcakova, J. and Bukovska, G. (2011) ‘Endolysin of bacteriophage BFK20: Evidence of a catalytic and a cell wall binding domain’, *FEMS Microbiology Letters*, 321(2), pp. 83–91. doi: 10.1111/j.1574-6968.2011.02312.x.

25. Gervasi, T., Horn, N., Wegmann, U., Dugo, G., Narbad, A. and Mayer, M. J. (2014) ‘Expression and delivery of an endolysin to combat *Clostridium perfringens*’, *Applied Microbiology and Biotechnology*, 98(6), pp. 2495–2505. doi: 10.1007/s00253-013-5128-y.

26. Gründling, A., Smith, D. L., Bläsi, U. and Young, R. (2000) ‘Dimerization between the holin and holin inhibitor of phage lambda.’, *Journal of Bacteriology*, 182(21), pp. 6075–6081. doi: 10.1128/JB.182.21.6075-6081.2000.

27. Guo, T., Zhang, C., Liu, W., Wang, S. and Kong, J. (2015) ‘Functional analysis of the N-terminal region of endolysin Lyb5 encoded by *Lactobacillus fermentum* bacteriophage φPYB5.’, *International Journal of Food Microbiology*. Elsevier B.V., 203, pp. 1–7. doi: 10.1016/j.ijfoodmicro.2015.02.033.

28. Gutiérrez, D., Ruas-Madiedo, P., Martínez, B., Rodríguez, A. and García, P. (2014) ‘Effective removal of staphylococcal biofilms by the endolysin LysH5.’, *PloS One*, 9(9), p. e107307. doi: 10.1371/journal.pone.0107307.

29. Hermoso, J. A., Monterroso, B., Albert, A., Galán, B., Ahrazem, O., García, P., Martínez-

Ripoll, M., García, J. L. and Menéndez, M. (2003) 'Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1', *Structure*, 11(10), pp. 1239–1249. doi: 10.1016/j.str.2003.09.005.

30. Hoopes, J. T., Stark, C. J., Kim, H. A., Sussman, D. J., Donovan, D. M. and Nelson, D. C. (2009) 'Use of a bacteriophage lysin, PlyC, as an enzyme disinfectant against *Streptococcus equi*', *Applied and Environmental Microbiology*, 75(5), pp. 1388–1394. doi: 10.1128/AEM.02195-08.

31. Horgan, M., O'Flynn, G., Garry, J., Cooney, J., Coffey, A., Fitzgerald, G. F., Paul Ross, R. and McAuliffe, O. (2009) 'Phage lysin LysK can be truncated to its CHAP domain and retain lytic activity against live antibiotic-resistant staphylococci', *Applied and Environmental Microbiology*, 75(3), pp. 872–874. doi: 10.1128/AEM.01831-08.

32. Huang, G., Shen, X., Gong, Y., Dong, Z., Zhao, X., Shen, W., Wang, J., Hu, F. and Peng, Y. (2014) 'Antibacterial properties of *Acinetobacter baumannii* phage Abp1 endolysin (PlyAB1).', *BMC Infectious Diseases*, 14(1), p. 681. doi: 10.1186/s12879-014-0681-2.

33. Jado, I., López, R., García, E., Fenoll, A., Casal, J., García, P., Pallares, R., de la Campa, A. G., Bouza, E., Baquero, F., Soriano, F., Prieto, J., Pallarés, R., Liñares, J., Garau, J., Martínez Lacasa, J., Latorre, C., Pérez-Trallero, E., García de Lomas, J. and Fleites, A. (2003) 'Phage lytic enzymes as therapy for antibiotic-resistant *Streptococcus pneumoniae* infection in a murine sepsis model', *Journal of Antimicrobial Chemotherapy*, 52(6), pp. 967–973. doi: 10.1093/jac/dkg485.

34. Jun, S. Y., Jung, G. M., Son, J. S., Yoon, S. J., Choi, Y. J. and Kang, S. H. (2011) 'Comparison of the antibacterial properties of phage endolysins SAL-1 and LysK', *Antimicrobial Agents and Chemotherapy*, 55(4), pp. 1764–1767. doi: 10.1128/AAC.01097-10.

35. Jun, S. Y., Jung, G. M., Yoon, S. J., Choi, Y. J., Koh, W. S., Moon, K. S. and Kang, S. H.

(2014) ‘Preclinical safety evaluation of intravenously administered SAL200 containing the recombinant phage endolysin SAL-1 as a pharmaceutical ingredient’, *Antimicrobial Agents and Chemotherapy*, 58(4), pp. 2084–2088. doi: 10.1128/AAC.02232-13.

36. Jun, S. Y., Jung, G. M., Yoon, S. J., Oh, M. D., Choi, Y. J., Lee, W. J., Kong, J. C., Seol, J. G. and Kang, S. H. (2013) ‘Antibacterial properties of a pre-formulated recombinant phage endolysin, SAL-1’, *International Journal of Antimicrobial Agents*. Elsevier B.V., 41(2), pp. 156–161. doi: 10.1016/j.ijantimicag.2012.10.011.

37. Kakikawa, M., Yokoi, K. J., Kimoto, H., Nakano, M., Kawasaki, K. I., Taketo, A. and Kodaira, K. I. (2002) ‘Molecular analysis of the lysis protein Lys encoded by *Lactobacillus plantarum* phage φg1e’, *Gene*, 299(1–2), pp. 227–234. doi: 10.1016/S0378-1119(02)01076-4.

38. Kikkawa, H. S., Ueda, T., Suzuki, S. I. and Yasuda, J. (2008) ‘Characterization of the catalytic activity of the gamma-phage lysin, PlyG, specific for *Bacillus anthracis*.’, *FEMS Microbiology Letters*, 286(2), pp. 236–240. doi: 10.1111/j.1574-6968.2008.01280.x.

39. Kong, M., Sim, J., Kang, T., Nguyen, H. H., Park, H. K., Chung, B. H. and Ryu, S. (2015) ‘A novel and highly specific phage endolysin cell wall binding domain for detection of *Bacillus cereus*.’, *European Biophysics Journal : EBJ*. Springer Berlin Heidelberg, 44(6), pp. 437–446. doi: 10.1007/s00249-015-1044-7.

40. Korndörfer, I. P., Danzer, J., Schmelcher, M., Zimmer, M., Skerra, A. and Loessner, M. J. (2006) ‘The crystal structure of the bacteriophage PSA endolysin reveals a unique fold responsible for specific recognition of *Listeria* cell walls’, *Journal of Molecular Biology*, 364(4), pp. 678–689. doi: 10.1016/j.jmb.2006.08.069.

41. Kretzer, J. W., Lehmann, R., Schmelcher, M., Banz, M., Kim, K. P., Korn, C. and Loessner, M. J. (2007) ‘Use of high-affinity cell wall-binding domains of bacteriophage endolysins for immobilization and separation of bacterial cells’, *Applied and Environmental Microbiology*,

73(6), pp. 1992–2000. doi: 10.1128/AEM.02402-06.

42. Lai, M. J., Lin, N. T., Hu, A., Soo, P. C., Chen, L. K., Chen, L. H. and Chang, K. C. (2011) 'Antibacterial activity of *Acinetobacter baumannii* phage ϕAB2 endolysin (LysAB2) against both Gram-positive and Gram-negative bacteria.', *Applied Microbiology and Biotechnology*, 90(2), pp. 529–539. doi: 10.1007/s00253-011-3104-y.

43. Lim, J. A., Shin, H., Kang, D. H. and Ryu, S. (2012) 'Characterization of endolysin from a *Salmonella typhimurium*-infecting bacteriophage SPN1S', *Research in Microbiology*. Elsevier Masson SAS, 163(3), pp. 233–241. doi: 10.1016/j.resmic.2012.01.002.

44. Lim, J., Shin, H., Heu, S. and Ryu, S. (2014) 'Exogenous lytic activity of SPN9CC endolysin against Gram-negative bacteria', *Journal of Microbiology and Biotechnology*, 24(6), pp. 803–811.

45. Linden, S. B., Zhang, H., Heselpoth, R. D., Shen, Y., Schmelcher, M., Eichenseher, F. and Nelson, D. C. (2015) 'Biochemical and biophysical characterization of PlyGRCS, a bacteriophage endolysin active against methicillin-resistant *Staphylococcus aureus*', *Applied Microbiology and Biotechnology*, 99(2), pp. 741–752. doi: 10.1007/s00253-014-5930-1.

46. Loeffler, J. M. and Fischetti, V. A. (2003) 'Synergistic lethal effect of a combination of phage lytic enzymes with different activities on penicillin-sensitive and -resistant *Streptococcus pneumoniae* strains', *Antimicrobial Agents and Chemotherapy*, 47(1), pp. 375–377. doi: 10.1128/AAC.47.1.375-377.2003.

47. Loeffler, J. M., Nelson, D. and Fischetti, V. A (2001) 'Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase.', *Science (New York, N.Y.)*, 294(5549), pp. 2170–2172. doi: 10.1126/science.1066869.

48. Lood, R., Raz, A., Molina, H., Euler, C. W. and Fischetti, V. A. (2014) 'A highly active and negatively charged *Streptococcus pyogenes* lysin with a rare d-Alanyl-l-Alanine endopeptidase activity protects mice against streptococcal bacteremia', *Antimicrobial Agents*

and Chemotherapy, 58(6), pp. 3073–3084. doi: 10.1128/AAC.00115-14.

49. Low, L. Y., Yang, C., Perego, M., Osterman, A. and Liddington, R. (2011) 'Role of net charge on catalytic domain and influence of cell wall binding domain on bactericidal activity, specificity, and host range of phage lysins', *Journal of Biological Chemistry*, 286(39), pp. 34391–34403. doi: 10.1074/jbc.M111.244160.

50. Martins, A. (2013) 'Mechanisms of resistance in bacteria: An evolutionary approach', *The Open Microbiology Journal*, 7(1), pp. 53–58. doi: 10.2174/1874285801307010053.

51. Mayer, M. J., Gasson, M. J. and Narbad, A. (2012) 'Genomic sequence of bacteriophage ATCC 8074-B1 and activity of its endolysin and engineered variants against *Clostridium sporogenes'*, *Applied and Environmental Microbiology*, 78(10), pp. 3685–3692. doi: 10.1128/AEM.07884-11.

52. Mayer, M. J., Narbad, A. and Gasson, M. J. (2008) 'Molecular characterization of a *Clostridium difficile* bacteriophage and its cloned biologically active endolysin', *Journal of Bacteriology*, 190(20), pp. 6734–6740. doi: 10.1128/JB.00686-08.

53. Nelson, D., Loomis, L. and Fischetti, V. A (2001) 'Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme.', *Proceedings of the National Academy of Sciences of the United States of America*, 98(7), pp. 4107–4112. doi: 10.1073/pnas.061038398.

54. O'Flaherty, S., Coffey, A., Meaney, W., Fitzgerald, G. F. and Ross, R. P. (2005) 'The recombinant phage lysin LysK has a broad spectrum of lytic activity against clinically relevant staphylococci, including methicillin-resistant *Staphylococcus aureus*', *Journal of Bacteriology*, 187(20), pp. 7161–7164. doi: 10.1128/JB.187.20.7161-7164.2005.

55. Obeso, J. M., Martínez, B., Rodríguez, A. and García, P. (2008) 'Lytic activity of the recombinant staphylococcal bacteriophage ΦH5 endolysin active against *Staphylococcus aureus* in milk', *International Journal of Food Microbiology*. Elsevier B.V., 128(2), pp. 212–

218. doi: 10.1016/j.ijfoodmicro.2008.08.010.

56. Oliveira, H., Thiagarajan, V., Walmagh, M., Sillankorva, S., Lavigne, R., Neves-Petersen, M. T., Kluskens, L. D. and Azeredo, J. (2014) 'A thermostable *Salmonella* phage endolysin, Lys68, with broad bactericidal properties against Gram-negative pathogens in presence of weak acids.', *PloS One*. Edited by K. Zeth, 9(10), p. e108376. doi: 10.1371/journal.pone.0108376.

57. Otto, M. (2008) 'Staphylococcal biofilms', in *Current Topics in Microbiology and Immunology*, pp. 207–228. doi: 10.1007/978-3-540-75418-3_10.

58. Oyston, P. C. F., Fox, M. A., Richards, S. J. and Clark, G. C. (2009) 'Novel peptide therapeutics for treatment of infections', *Journal of Medical Microbiology*, 58(8), pp. 977–987. doi: 10.1099/jmm.0.011122-0.

59. Pastagia, M., Euler, C., Chahales, P., Fuentes-Duculan, J., Krueger, J. G. and Fischetti, V. A. (2011) 'A novel chimeric lysin shows superiority to mupirocin for skin decolonization of methicillin-resistant and -sensitive *Staphylococcus aureus* strains', *Antimicrobial Agents and Chemotherapy*, 55(2), pp. 738–744. doi: 10.1128/AAC.00890-10.

60. Payne, K. M. and Hatfull, G. F. (2012) 'Mycobacteriophage endolysins: Diverse and modular enzymes with multiple catalytic activities', *PLoS One*. Edited by J. Nigou, 7(3), p. e34052. doi: 10.1371/journal.pone.0034052.

61. Plotka, M., Kaczorowska, A. K., Stefanska, A., Morzywolek, A., Fridjonsson, O. H., Dunin-Horkawicz, S., Kozlowski, L., Hreggvidsson, G. O., Kristjansson, J. K., Dabrowski, S., Bujnicki, J. M. and Kaczorowski, T. (2014) 'Novel highly thermostable endolysin from *Thermus scotoductus* MAT2119 bacteriophage Ph2119 with amino acid sequence similarity to eukaryotic peptidoglycan recognition proteins', *Applied and Environmental Microbiology*, 80(3), pp. 886–895. doi: 10.1128/AEM.03074-13.

62. Pohane, A. A., Patidar, N. D. and Jain, V. (2015) 'Modulation of domain-domain interaction

and protein function by a charged linker: a case study of mycobacteriophage D29 endolysin.', *FEBS Letters*. Federation of European Biochemical Societies, 589(6), pp. 695–701. doi: 10.1016/j.febslet.2015.01.036.

63. Proença, D., Fernandes, S., Leandro, C., Silva, F. A., Santos, S., Lopes, F., Mato, R., Cavaco-Silva, P., Pimentel, M. and São-José, C. (2012) 'Phage endolysins with broad antimicrobial activity against *Enterococcus faecalis* clinical strains.', *Microbial Drug Resistance (Larchmont, N.Y.)*, 18(3), pp. 322–332. doi: 10.1089/mdr.2012.0024.

64. Rashel, M., Uchiyama, J., Ujihara, T., Uehara, Y., Kuramoto, S., Sugihara, S., Yagyu, K. I., Muraoka, A., Sugai, M., Hiramatsu, K., Honke, K. and Matsuzaki, S. (2007) 'Efficient elimination of multidrug-resistant *Staphylococcus aureus* by cloned lysin derived from bacteriophage phi MR11.', *The Journal of Infectious Diseases*, 196(8), pp. 1237–1247. doi: 10.1086/521305.

65. Regulski, K., Courtin, P., Kulakauskas, S. and Chapot-Chartier, M. P. (2013) 'A novel type of peptidoglycan-binding domain highly specific for amidated D-Asp cross-bridge, identified in *Lactobacillus casei* bacteriophage endolysins', *Journal of Biological Chemistry*, 288(28), pp. 20416–20426. doi: 10.1074/jbc.M112.446344.

66. Rigden, D. J., Jedrzejas, M. J. and Galperin, M. Y. (2003) 'Amidase domains from bacterial and phage autolysins define a family of gamma-D,L-glutamate-specific amidohydrolases.', *Trends in Biochemical Sciences*, 28(5), pp. 230–234. doi: 10.1016/S0968-0004(03)00062-8.

67. Roach, D. R., Khatibi, P. a, Bischoff, K. M., Hughes, S. R. and Donovan, D. M. (2013) 'Bacteriophage-encoded lytic enzymes control growth of contaminating *Lactobacillus* found in fuel ethanol fermentations.', *Biotechnology for Biofuels*, 6(1), p. 20. doi: 10.1186/1754-6834-6-20.

68. Rodríguez-Cerrato, V., García, P., del Prado, G., García, E., Gracia, M., Huelves, L., Ponte, C., López, R. and Soriano, F. (2007) '*In vitro* interactions of LytA, the major pneumococcal

autolysin, with two bacteriophage lytic enzymes (Cpl-1 and Pal), cefotaxime and moxifloxacin against antibiotic-susceptible and -resistant *Streptococcus pneumoniae* strains', *Journal of Antimicrobial Chemotherapy*, 60(5), pp. 1159–1162. doi: 10.1093/jac/dkm342.

69. Rodríguez-Rubio, L., Gutiérrez, D., Martínez, B., Rodríguez, A. and García, P. (2012) 'Lytic activity of LysH5 endolysin secreted by *Lactococcus lactis* using the secretion signal sequence of bacteriocin Lcn972', *Applied and Environmental Microbiology*, 78(9), pp. 3469–3472. doi: 10.1128/AEM.00018-12.

70. Rodríguez-Rubio, L., Martínez, B., Rodríguez, A., Donovan, D. M., Götz, F. and García, P. (2013) 'The Phage lytic proteins from the *Staphylococcus aureus* bacteriophage vB_SauS-phiIPLA88 display multiple active catalytic domains and do not trigger staphylococcal resistance', *PLoS One*, 8(5), pp. 1–7. doi: 10.1371/journal.pone.0064671.

71. Sanchez-Vizuete, P., Orgaz, B., Aymerich, S., Le Coq, D. and Briandet, R. (2015) 'Pathogens protection against the action of disinfectants in multispecies biofilms.', *Frontiers in Microbiology*, 6(705), p. 1-12. doi: 10.3389/fmicb.2015.00705.

72. Sanz-Gaitero, M., Keary, R., Garcia-Doval, C., Coffey, A. and van Raaij, M. J. (2013) 'Crystallization of the CHAP domain of the endolysin from *Staphylococcus aureus* bacteriophage K', *Acta Crystallographica Section F Structural Biology and Crystallization Communications*, 69(Pt 12), pp. 1393–1396. doi: 10.1107/S1744309113030133.

73. Sao-Jose, C., Parreira, R., Vieira, G. and Santos, M. A. (2000) 'The N-terminal region of the *Oenococcus oeni* bacteriophage fOg44 lysin behaves as a bona fide signal peptide in *Escherichia coli* and as a cis-inhibitory element, preventing lytic activity on oenococcal cells', *Journal of Bacteriology*, 182(20), pp. 5823–5831. doi: 10.1128/JB.182.20.5823-5831.2000.

74. Sass, P. and Bierbaum, G. (2007) 'Lytic activity of recombinant bacteriophage phi11 and phi12 endolysins on whole cells and biofilms of *Staphylococcus aureus*.', *Applied and*

Environmental Microbiology, 73(1), pp. 347–52. doi: 10.1128/AEM.01616-06.

75. Schmelcher, M., Donovan, D. M. and Loessner, M. J. (2012) 'Bacteriophage endolysins as novel antimicrobials', *Future Microbiology*, 7(10), pp. 1147–1171. doi: 10.2217/fmb.12.97.

76. Schmelcher, M., Powell, A. M., Becker, S. C., Camp, M. J. and Donovan, D. M. (2012) 'Chimeric phage lysins act synergistically with lysostaphin to kill mastitis-causing *Staphylococcus aureus* in murine mammary glands', *Applied and Environmental Microbiology*, 78(7), pp. 2297–2305. doi: 10.1128/AEM.07050-11.

77. Schmelcher, M., Shabarova, T., Eugster, M. R., Eichenseher, F., Tchang, V. S., Banz, M. and Loessner, M. J. (2010) 'Rapid multiplex detection and differentiation of *Listeria* cells by use of fluorescent phage endolysin cell wall binding domains', *Applied and Environmental Microbiology*, 76(17), pp. 5745–5756. doi: 10.1128/AEM.00801-10.

78. Schmelcher, M., Tchang, V. S. and Loessner, M. J. (2011) 'Domain shuffling and module engineering of *Listeria* phage endolysins for enhanced lytic activity and binding affinity', *Microbial Biotechnology*, 4(5), pp. 651–662. doi: 10.1111/j.1751-7915.2011.00263.x.

79. Schuch, R., Nelson, D. and Fischetti, V. A (2002) 'A bacteriolytic agent that detects and kills *Bacillus anthracis*.', *Nature*, 418(6900), pp. 884–889. doi: 10.1038/nature01026.

80. Singh, P. K., Donovan, D. M. and Kumar, A. (2014) 'Intravitreal injection of the chimeric phage endolysin Ply187 protects mice from *Staphylococcus aureus* endophthalmitis.', *Antimicrobial Agents and Chemotherapy*, 58(8), pp. 4621–4629. doi: 10.1128/AAC.00126-14.

81. Smith, D. L., Struck, D. K., Scholtz, J. M. and Young, R. (1998) 'Purification and biochemical characterization of the lambda holin', *Journal of Bacteriology*, 180(9), pp. 2531–2540.

82. Solanki, K., Grover, N., Downs, P., Paskaleva, E. E., Mehta, K. K., Lee, L., Schadler, L. S., Kane, R. S. and Dordick, J. S. (2013) 'Enzyme-based listericidal nanocomposites.', *Scientific*

Reports, 3, p. 1584. doi: 10.1038/srep01584.

83. Son, B., Yun, J., Lim, J. A., Shin, H., Heu, S. and Ryu, S. (2012) 'Characterization of LysB4, an endolysin from the *Bacillus cereus*-infecting bacteriophage B4', *BMC Microbiology*. BioMed Central Ltd, 12(1), p. 33. doi: 10.1186/1471-2180-12-33.

84. Son, J. S., Lee, S. J., Jun, S. Y., Yoon, S. J., Kang, S. H., Paik, H. R., Kang, J. O. and Choi, Y. J. (2010) 'Antibacterial and biofilm removal activity of a podoviridae *Staphylococcus aureus* bacteriophage SAP-2 and a derived recombinant cell-wall-degrading enzyme', *Applied Microbiology and Biotechnology*, 86(5), pp. 1439–1449. doi: 10.1007/s00253-009-2386-9.

85. Summer, E. J., Berry, J., Tran, T. A. T., Niu, L., Struck, D. K. and Young, R. (2007) 'Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts', *Journal of Molecular Biology*, 373(5), pp. 1098–1112. doi: 10.1016/j.jmb.2007.08.045.

86. Swift, S., Seal, B., Garrish, J., Oakley, B., Hiett, K., Yeh, H. Y., Woolsey, R., Schegg, K., Line, J. and Donovan, D. (2015) 'A thermophilic phage endolysin fusion to a *Clostridium perfringens*-specific cell wall binding domain creates an anti-*Clostridium* antimicrobial with improved thermostability', *Viruses*, 7(6), pp. 3019–3034. doi: 10.3390/v7062758.

87. Tišáková, L., Vidová, B., Farkašovská, J. and Godány, A. (2014) 'Bacteriophage endolysin Lyt μ1/6: Characterization of the C-terminal binding domain', *FEMS Microbiology Letters*, 350(2), pp. 199–208. doi: 10.1111/1574-6968.12338.

88. Turner, M. S., Waldherr, F., Loessner, M. J. and Giffard, P. M. (2007) 'Antimicrobial activity of lysostaphin and a *Listeria monocytogenes* bacteriophage endolysin produced and secreted by lactic acid bacteria', *Systematic and Applied Microbiology*, 30(1), pp. 58–67. doi: 10.1016/j.syapm.2006.01.013.

89. Vasala, A., Välkkilä, M., Caldentey, J. and Alatossava, T. (1995) 'Genetic and biochemical characterization of the *Lactobacillus delbrueckii* subsp. *lactis* bacteriophage LL-H lysin.',

Applied and Environmental Microbiology, 61(11), pp. 4004–4011.

90. Walcher, G., Stessl, B., Wagner, M., Eichenseher, F., Loessner, M. J. and Hein, I. (2010) 'Evaluation of paramagnetic beads coated with recombinant *Listeria* phage endolysin-derived cell-wall-binding domain proteins for separation of *Listeria monocytogenes* from raw milk in combination with culture-based and real-time polymerase chain reaction-based quantification', *Foodborne Pathogens and Disease*, 7(9), pp. 1019–1024. doi: 10.1089/fpd.2009.0475.
91. Walmagh, M., Briers, Y., dos Santos, S. B., Azeredo, J. and Lavigne, R. (2012) 'Characterization of modular bacteriophage endolysins from *Myoviridae* phages OBP, 201φ2-1 and PVP-SE1', *PLoS One*, 7(5), p. e36991. doi: 10.1371/journal.pone.0036991.
92. White, R., Chiba, S., Pang, T., Dewey, J. S., Savva, C. G., Holzenburg, A., Pogliano, K. and Young, R. (2011) 'Holin triggering in real time.', *Proceedings of the National Academy of Sciences of the United States of America*, 108(2), pp. 798–803. doi: 10.1073/pnas.1011921108.
93. Wittebole, X., De Roock, S. and Opal, S. M. (2014) 'A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens.', *Virulence*, 5(1), pp. 226–235. doi: 10.4161/viru.25991.
94. Yoong, P., Schuch, R., Nelson, D. and Fischetti, V. A. (2004) 'Identification of a broadly active phage lytic enzyme with lethal activity against antibiotic-resistant *Enterococcus faecalis* and *Enterococcus faecium*', *Journal of Bacteriology*, 186(14), pp. 4808–4812. doi: 10.1128/JB.186.14.4808-4812.2004.
95. Young, R. (1992) 'Bacteriophage lysis: mechanism and regulation.', *Microbiological Reviews*, 56(3), pp. 430–481.
96. Young, R. (2002) 'Bacteriophage holins: deadly diversity.', *Journal of Molecular Microbiology and Biotechnology*, 4(1), pp. 21–36.

97. Young, R., Wang, I. N. and Roof, W. D. (2000) ‘Phages will out: Strategies of host cell lysis’, *Trends in Microbiology*, 8(3), pp. 120–128. doi: 10.1016/S0966-842X(00)01705-4.

98. Zhang, H., Bao, H., Billington, C., Hudson, J. A. and Wang, R. (2012) ‘Isolation and lytic activity of the *Listeria* bacteriophage endolysin LysZ5 against *Listeria monocytogenes* in soya milk’, *Food Microbiology*. Elsevier Ltd, 31(1), pp. 133–136. doi: 10.1016/j.fm.2012.01.005.

99. Zhang, W., Mi, Z., Yin, X., Fan, H., An, X., Zhang, Z., Chen, J. and Tong, Y. (2013) ‘Characterization of *Enterococcus faecalis* phage IME-EF1 and its endolysin’, *PLoS One*. Edited by M. J. Horsburgh, 8(11), p. e80435. doi: 10.1371/journal.pone.0080435.

100. Zheng, Y., Struck, D. K. and Young, R. (2009) ‘Purification and functional characterization of phiX174 lysis protein E.’, *Biochemistry*, 48(22), pp. 4999–5006. doi: 10.1021/bi900469g.

Chapter II

Comparison of *Staphylococcus* phage K with close phage relatives commonly employed in phage therapeutics

A manuscript based on this chapter has been published in the journal *Antibiotics MDPI* and the TEM work was done at the Max Rubner-Institute in Germany by Dr. Horst Neve.

2.1 Abstract

The increase in antibiotic resistance in pathogenic bacteria is a public health danger requiring alternative treatment options, and this has led to renewed interest in phage therapy. In this respect, we describe the distinct host ranges of *Staphylococcus* phage K, and two other K-like phages against 23 isolates, including 21 methicillin-resistant *S. aureus* (MRSA) representative sequence types representing the Irish National MRSA Reference Laboratory collection. The two K-like phages were isolated from the *Fersisi* therapeutic phage mix from the Tbilisi Eliava Institute, and were designated B1 (vB_SauM_B1) and JA1 (vB_SauM_JA1). The sequence relatedness of B1 and JA1 to phage K was observed to be 95% and 94%, respectively. In terms of host range on the 23 *Staphylococcus* isolates, B1 and JA1 infected 73.9% and 78.2%, respectively whereas K infected only 43.5%. Eleven open reading frames (ORFs) present in both phages B1 and JA1 but absent in phage K were identified by comparative genomic analysis. These ORFs were also found to be present in the genomes of phages (Team 1, vB_SauM-fRuSau02, Sb_1 and ISP) that are components of several commercial phage mixtures with reported wide host ranges. This is the first comparative study of therapeutic staphylococcal phages within the recently described genus *Kayvirus*.

2.2 Introduction

Staphylococcus aureus (*S. aureus*) is an opportunistic and important pathogen in clinical and health-care settings, causing a wide variety of diseases commonly involving the skin, soft tissue, bone, and joints (Tong *et al.*, 2015). It is also a well-known causative agent of prosthetic joint infections (PJI), cardiac device infections, and intravascular catheter infections (Tong *et al.*, 2015). *S. aureus* pathogenicity is due, in part, to its ability to acquire and express a wide array of virulence factors, as well as antimicrobial resistance determinants (Shore *et al.*, 2008), an example of which involves the acquisition of the staphylococcal cassette chromosome (SCCmec) leading to the development of methicillin resistance in *S. aureus* (Hiramatsu *et al*., 2001). Methicillin-resistant *S. aureus* (MRSA) was first reported in 1961 (Jevons, 1961), and has since been observed to cause serious infections in hospitals worldwide. Reports of MRSA clones resistant to the majority of antibiotics are a growing concern (Klein et al., 2007). As such, new treatment options are needed.

Bacteriophages (phages) are biological entities composed of either DNA or RNA enclosed within a protein coat (O'Flaherty et al., 2009). They are highly specific, with most phages capable of infecting only a single bacterial species (O'Flaherty *et al.*, 2009; Schmelcher and Loessner, 2014), and studies on these viruses have been performed since the late 19th century (Wittebole *et al*., 2014). The phage infection process usually begins with the recognition of the receptor on the bacterial cell surface by its receptor binding protein (Bertozzi *et al*., 2016). In natural environments bacterial hosts have evolved many mechanisms to protect themselves from phage attack to include; adsorption blocking, DNA injection blocking, restriction-modification system (R/M), abortive infection, and the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas systems (Hyman and Abedon, 2010; Labrie *et al*., 2010). In turn, phages have evolved several strategies for overcoming these

systems to ensure their survival in the phage-host co-evolutionary race (Hall *et al.*, 2011; Hall *et al*., 2013; Samson *et al*., 2013).

The use of phages as therapeutics to eliminate pathogenic bacteria dates back to experiments conducted by Felix d'Herelle in 1919 at a French hospital to treat dysentery (Sulakvelidze *et al.*, 2001). Since then, a wide range of phage therapy trials have been undertaken, many with very promising results (Sulakvelidze *et al*., 2001; Abedon *et al*., 2011). Pyophage and Intesti-phage are among the commercial phage mixtures currently produced at the Eliava Institute. Metagenomic studies on these phage mixtures have been reported (Zschach *et al*., 2015; Villarroel *et al*., 2017) and the staphylococcal phages Sb-1 and ISP are key components of Pyophage (Kvachadze *et al*., 2011; Vandersteegen *et al*., 2011). Other phages isolated from these commercial phages mixes have also been reported (Markoishvili *et al*., 2002; Jikia *et al*., 2005; El Haddad *et al*., 2014; Leskinen *et al*., 2017). Phages like vB_SauM-fRuSau02 was isolated from a phage mix produced by Microgen (Moscow, Russia) (Leskinen *et al*., 2017) and Team 1 was isolated from PhageBioDerm, a wound healing preparation consisting of a biodegradable polymer impregnated with an antibiotic and lytic phages (Markoishvili *et al*., 2002; Jikia *et al*., 2005; El Haddad *et al*., 2014). These phages all possess a wide host range against a number of clinically relevant *S. aureus* isolates, demonstrating the efficacy of such commercial phage mixtures in treating a range of bacterial infections (Markoishvili *et al*., 2002; Jikia *et al*., 2005; Kvachadze *et al*., 2011; Vandersteegen *et al*., 2011; El Haddad *et al*., 2014; Leskinen *et al*., 2017).

In this paper, we employed another phage mixture from the Eliava Institute, namely the Fersisi phage mix. Fersisi is a relatively new combination developed approximately 15–20 years ago on the basis of Pyophage, although with fewer phage components. Two phages from this mix were designated B1 (vB_SauM_B1) and JA1 (vB_SauM_JA1). Phage K, on the other hand, is a well-known phage being the type phage of the recently designated genus

Kayvirus of the subfamily *Spounavirinae* (Adriaenssens *et al*., 2017). The exact origin of phage K is unknown, but descriptions of the phage are made as far back as 1949 (Rountree, 1949; O'Flaherty *et al*., 2005). An initial host range study involving this phage reported it to be ineffective against many MRSA strains (O'Flaherty *et al*., 2005). Thus, phages B1 and JA1 were compared (on the basis of their host range) to phage K to explore possible host range differences and it was observed that both phages had broader host ranges. A comparative study was performed on their genomes and the genomes of similar phages from other commercial phage mixtures (Team 1, vB_SauM-fRuSau02, Sb_1 and ISP) with reported wide host ranges, to provide molecular insight into the differences in host range encountered in this study.

2.3 Materials and Methods

2.3.1 Bacterial strains, phage and growth requirement

Phages B1 and JA1 were isolated from a commercial phage cocktail purchased from the George Eliava Institute of Bacteriophage, Microbiology and Virology, Tbilisi, Georgia. The MRSA strains utilized in this study were all acquired from the Irish National MRSA Reference Laboratory, Dublin, Ireland (Shore *et al.*, 2008) with the exception of DPC5246 and CIT281189, which are routine propagation strains utilized in our laboratory (O'Flaherty *et al.*, 2004, 2005). These strains were routinely cultured in Brain Heart Infusion broth (BHI; Sigma-Aldrich, St. Louis, MO, USA) at 37°C with shaking or on BHI plates containing 1.5% (w/v) bacteriological agar (Sigma-Aldrich, USA). All strains were stocked in BHI containing 40% (v/v) glycerol and stored at −80°C.

2.3.2 CsCl gradient purification

Isopycnic centrifugation through CsCl gradients was performed as previously described (Sambrook and Russell, 2001), with a number of modifications. A high titer phage lysate (>1 × 10^9 plaque forming units [PFU/ml]), was precipitated using polyethylene glycol (15% (w/v) PEG8000, 1M NaCl) at 4°C overnight and centrifuged, after which the pellet was resuspended in TMN buffer (10mM Tris-HCl pH 7.4, 10mM $MgSO_4 \cdot 7H_2O$, 0.5M NaCl). The resulting phage preparation was placed onto a CsCl step gradient composed of 1.3, 1.5, and 1.7g/ml layers and spun in a 100 Ti rotor (Beckman Coulter, Brea, CA, USA) at 200,480 g for 3hrs at 4°C. The resulting phage preparations were dialyzed in Tris-HCl buffer (10mM, pH 7.5) at 4°C.

2.3.3 Phage host range and adsorption study

Host range assay was performed for phages B1, JA1, and K using the plaque assay plating technique (Tables 2 and 3). This was done in triplicate for three independent experiments. The efficiency of plaquing (EOP) was determined by dividing the phage titer on each test strain by the phage titer of the reference strain (*S. aureus* DPC5246, in the case of phages B1 and K, and *S. aureus* CIT281189 for phage JA1) (Gutiérrez *et al.*, 2015). An adsorption assay was performed according to the protocol previously described elsewhere with some modification (Li *et al.*, 2016). Briefly, MRSA strains were grown to an optical density (OD) of 0.2 at 600nm (estimated cell count at 10^8 colony forming unit [CFU/ml]) and 100µl of cells were mixed with 100µl of respective phage titered at approximately 1×10^7 PFU/ml for a multiplicity of infection (MOI) of 0.1. The resulting mixtures were incubated at room temperature for 5min to allow for phage adsorption. The bound phages were separated from the free phages by centrifugation at 14,000 rpm for 5min. Adsorption of the phage on each strain was determined by subtracting the number of unbound phage (per ml) from the total input PFU/ml. Adsorption efficiency was expressed as a percentage relative to the propagating strain DPC5246.

2.3.4 Transmission electron microscopy

Electron microscopic analysis was performed following negative staining of the CsCl gradient prepared phages on freshly prepared carbon films with 2% (w/v) uranyl acetate. Electron micrographs were taken using a Tecnai 10 transmission electron microscope (FEI Thermo Fisher, Eindhoven, the Netherlands) at an acceleration voltage of 80kV with a MegaView G2 CDD camera (EMSIS, Muenster, Germany).

2.3.5 Phage DNA isolation

Phage DNA extraction was performed on CsCl purified high titer phages. These were initially treated with $MgCl_2$ followed by pre-treatment with DNase and RNase for 60min at 37°C. Following that subsequent treatment with SDS, EDTA and proteinase K with further incubation for 60min at 55°C were performed. DNA extractions were then performed on the pre-treated samples with phenol/chloroform/isoamyl alcohol (25:24:1 v/v/v) and chloroform/isoamyl alcohol (24:1 v/v). DNA precipitation was achieved using sodium acetate and 95% (v/v) ethanol. DNA quality and quantity were estimated using a Nanodrop (ND-1000) and visualized following agarose gel electrophoresis

2.3.6 Phage DNA sequencing

DNA sequencing was performed with a high throughput Illumina HiSeq system sequencing (GATC Biotech, Konstanz, Germany). Library preparation was performed by DNA fragmentation together with adapter ligation. The libraries were then measured and quantified on a Fragment Analyzer and then sequenced to generate 2 × 300bp paired-end reads. *De novo* assembly was performed using CLC Bio Genomics Workbench v8.0 (Aarhus, Denmark).

2.3.7 Bioinformatic analysis

Open reading frames (ORFs) for the sequenced phages were predicted with Glimmer (Delcher, 1999) and GenemarkS (Besemer *et al.*, 2001). Putative functions were assigned to these ORFs using BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins), HHpred (https://toolkit.tuebingen.mpg.de/#/tools/hhpred; [Söding *et al*., 2005]) and InterProscan (http://www.ebi.ac.uk/interpro/search/sequence-search; [Mitchell *et al*., 2015]). Transfer RNA was predicted using tRNAscan-SE (http://lowelab.ucsc.edu/ tRNAscan-SE/; (Lowe and Eddy, 1997)) and ARAGORN (http://130.235.46.10/ARAGORN/; (Laslett and Canback, 2004)). Potential promoters were

predicted using MEME (Multiple Em for Motif Elicitation) (http://meme-suite.org/tools/meme; (Bailey *et al.*, 2009)), followed by manual curation. Potential Rho-independent terminators were identified using ARNold (http://rna.igmors.u-psud.fr/toolbox/arnold; (Naville *et al*., 2011)) with Mfold QuikFold (http://unafold.rna.albany.edu/?q=DINAMelt/Quickfold; [Zucker *et al*., 2003]) using RNA energy rules 3.0 to verify predictions. Artemis Comparison Tool (ACT) was used for the identification of feature variations between the genomes of phages, with homology being assessed with BLASTN (Carver *et al.*, 2005) Genome comparison maps between phages were visualized using the Easyfig visualization tool (Sullivan *et al*., 2011). K-like *Staphylococcus* phages used in comparative studies were K (KF766114), Team 1 (KC012913), vB_SauM-fRuSau02 (MF398190), Sb-1 (HQ163896) and ISP (FR852584).

2.3.8 Nucleotide sequence accession number

The genome sequence for phages B1 and JA1 were deposited into GenBank under the accession numbers MG656408 and MF405094, respectively.

2.4 Results and Discussion

2.4.1 Origin of phages B1 and JA1

Phages B1 and JA1 were isolated from the Fersisi commercial phage mixtures; batch 010112 (B1) and F-062015 (JA1). This product is used in the treatment of staphylococcal and streptococcal infections. For the isolation of B1, phage enrichment was carried out using staphylococcal host cultured from the sonicate fluid of a hospital patient suffering from PJI. DPC5246 was subsequently used as propagating host for B1, as a prophage was encountered in the PJI strain. Phage enrichment in the isolation of JA1 was done using the Cork Institute of Technology (CIT) collection strain *S. aureus* CIT281189. Both the PJI strain and CIT281189 were insensitive to phage K.

2.4.1 Morphology and host range of phages K, B1 and JA1

Phages B1 and JA1 exhibited typical characteristics of phages belonging to the *Myoviridae* family, similar to the reported morphology of phage K (O'Flaherty *et al*., 2005). All three phages possessed an A1 morphology (Ackermann, 1996), displaying an icosahedral head as well as a long contractile tail. They also contained a structure previously described as knob-like appendages by O'Flaherty *et al* (2005), extending from their base plates (likely "clumped/aggregated" base plate appendices) and clearly visible in Figure 1. Estimations were made on the dimensions of these phages (Table 1). Capsid heights were estimated as 92.9 ± 4.0nm (B1), 87.0 ± 2.1nm (JA1) and 92.9 ± 3.8nm (K). Tail dimension were also estimated as 233.0 ± 4.4 × 23.4 ± 1.2nm (B1), 231.5 ± 4.7 × 22.7 ± 0.9nm (JA1), and 227.5 ± 5.5 × 23.8 ± 1.0nm (K), and base plates/knobs complexes were estimated as 30.1 ± 1.8 × 47.2 ± 3.7nm (B1), 32.5 ± 7.9 × 45.8 ± 1.4nm (JA1), and 36.6 ± 5.1 × 41.7 ± 2.6nm (K).

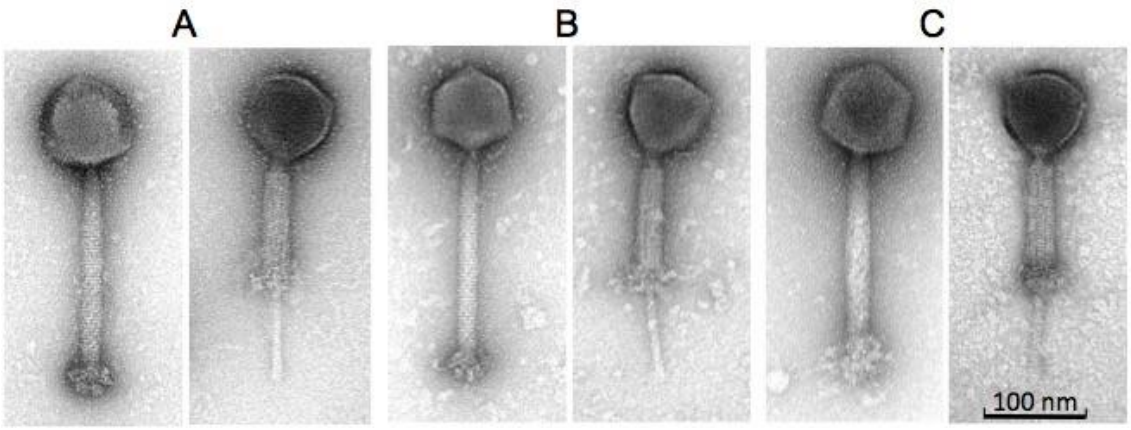


Figure 1. Transmission electron micrographs of phages B1 (A), JA1 (B), and K (C) showing their icosahedral capsid and their long contractile tail (both extended and contracted).

Table 1. Dimensions of staphylococcal phages B1, JA1, and K derived from micrographs obtained from transmission electron microscopy.

Phages	**Head (nm)**	**Tail length (nm) (incl. "knob")**	**Tail width (nm)**	**Baseplate "knob" length (nm)**	**Baseplate "knob" width (nm)**
B1	92.9 ± 4.0 (*n* = 11)	233.0 ± 4.4 (*n* = 12)	23.4 ± 1.2 (*n* = 12)	30.1 ± 1.8 (*n* = 12)	47.2 ± 3.7 (*n* = 10)
JA1	87.0 ± 2.1 (*n* = 9)	231.5 ± 4.7 (*n* = 9)	22.7 ± 0.9 (*n* = 9)	32.5 ± 7.9 (*n* = 9)	45.8 ± 1.4 (*n* = 9)
K	92.9 ± 3.8 (*n* = 16)	227.5 ± 5.5 (*n* = 16)	23.8 ± 1.0 (*n* = 16)	36.6 ± 5.1 (*n* = 16)	41.7 ± 2.6 (*n* = 16)

Owing to the similar morphology of all three phages, a host range study was conducted to explore possible differences in host spectra across a number of hospital isolates. Twenty-one of these isolates represented the entire collection of MRSA sequence-types identified in Ireland by the National MRSA Reference Laboratory (Dublin, Ireland), and includes the commonly encountered ST22-MRSA-IV, which has been predominant in Irish hospitals since the late 1990s (Rossney et al., 2006). The other two S. aureus strains used in this study were

included as additional phage propagation strains. Host range was assessed by plaque assay technique on lawns of various MRSA strains listed in Table 2. The efficiency of plaquing (EOP) was used to represent the degree to which each of the phages studied infected all 23 staphylococcal strains. Phage JA1 had the broadest host range, forming plaques on 18 out of the 23 staphylococcal strains examined. B1 also had a broad host range and was capable of forming plaques on 17 isolates (with some in common with the 18 lysed by phage JA1). Phage K had the narrowest host range, forming plaques on only 10 of the isolates (including its propagating strain DPC5246). All 23 staphylococcal strains were effectively lysed by at least one of the three phages, with the exception of E1139 (IV) ST45 and E1185 (IV) ST12, whose EOP were significantly low at 3.88×10^{-6} and 1.16×10^{-6} respectively; as well as 3488 (VV) ST8, which was resistant to all three phages. Plaque size ranged from 0.5mm to 1.5mm, with a halo occurring in some instances (Table 3 and Supplementary Materials, Figure S1). The wide host range encountered in this study is common among K-like phages and has been reported for other staphylococcal K-like phages, such as JD007, which infected 95% of S. aureus isolates obtained from several hospitals in Shanghai, China (Cui et al., 2017).

Table 2. Host ranges of staphylococcal phages B1, JA1, and K against methicillin-resistant *Staphylococcus aureus* (MRSA) strains from the Irish National Reference Laboratory (St. James's Hospital Dublin, Ireland) including the efficiency of plaquing (EOP) of these strains**.**

***S. aureus* strain**	**Phage K**	**Phage B1**	**Phage JA1**
DPC5246*	1.00 ± 0.0	1.00 ± 0.0	8.98×10^{-1} ± 0.8
CIT281189*	No infection	No infection	1.00 ± 0.0
0.0066 (IIIV) ST239	No infection	No infection	2.59 ± 2.5
0.1206 (IV) ST250	No infection	3.89×10^{-1} ± 0.3	1.35 ± 1.2
0.1239 (III) ST239	No infection	1.46×10^{-1} ± 0.1	4.17×10^{-2} ± 0.0
0.1345 (II) ST5	No infection	No infection	2.08×10^{-1} ± 0.1
0073 (III) ST239	No infection	3.21×10^{-1} ± 0.2	No infection
0104 (III) ST239	No infection	3.95×10^{-1} ± 0.2	1.82 ± 1.6
0220 (II) ST5	3.03×10^{-1} ± 0.1	2.17×10^{-1} ± 0.2	2.38×10^{-1} ± 0.2
0242 (IV) ST30	4.43×10^{-1} ± 0.1	5.23×10^{-1} ± 0.5	4.90×10^{-1} ± 0.3
0308 (IA) ST247	1.40 ± 0.2	1.36 ± 1.3	1.71 ± 1.6
3045 (IIV) ST8	No infection	4.93×10^{-2} ± 0.0	1.69 ± 0.7
3144 (IIV) ST8	No infection	1.21 ± 1.0	2.17 ± 1.2
3488 (VV) ST8	No infection	No infection	No infection
3581 (IA) ST247	No infection	No infection	9.26×10^{-1} ± 0.7
3594 (II) ST36	4.38×10^{-1} ± 0.1	8.67×10^{-1} ± 0.4	1.06 ± 0.7
3596 (IIV) ST8	2.49×10^{-4} ± 0.0	1.29 ± 0.9	3.59 ± 2.7
E1038 (IIV) ST8	1.27×10^{-4} ± 0.0	2.02×10^{-1} ± 0.2	1.89 ± 1.4
E1139 (IV) ST45	No infection	3.88×10^{-6} ± 0.0	No infection
E1174 (IV) ST22	7.03×10^{-1} ± 0.7	3.11×10^{-1} ± 0.2	No infection
E1185 (IV) ST12	1.16×10^{-6} ± 0.0	No infection	No infection
E1202 (II) ST496	No infection	4.79×10^{-1} ± 0.2	9.49×10^{-1} ± 0.8
M03/0073 (III) ST239	1.76 ± 0.5	1.51 ± 0.8	2.30 ± 0.7

* *S. aureus* strains for phage propagation; data is represented as means ± standard deviations based on triplicate measurements.

Table 3. Zone sizes and morphologies of B1, JA1, and K plaques formed on MRSA strains collected from the Irish National MRSA Reference Laboratory (St. James's Hospital Dublin, Ireland).

***S. aureus* strain**	**Phage K**	**Phage B1**	**Phage JA1**
DPC5246	2mm	1mm with halo to 2mm	1mm with halo to 2mm
CIT281189	No plaques	No plaques	1.5mm
0.0066 (IIV) ST239	No plaques	No plaques	1mm
0.1206 (IV) ST250	No plaques	2mm	0.5mm with halo to 1mm
0.1239 (III) ST239	No plaques	0.5mm, faint plaques	1mm
0.1345 (II) ST5	No plaques	No plaques	1mm
0073 (III) ST239	No plaques	0.5mm	No plaques
0104 (III) ST239	No plaques	0.5mm	1mm
0220 (II) ST5	0.5mm	1mm	1mm
0242 (IV) ST30	1mm	1.5mm	1.5mm
0308 (IA) ST247	1mm	1mm	0.5mm, faint plaques
3045 (IIV) ST8	No plaques	1mm	1mm
3144 (IIV) ST8	No plaques	1.5mm, faint plaques	1mm
3488 (VV) ST8	No plaques	0.5mm, faint plaques	0.5mm with halo to 1mm
3581 (IA) ST247	No plaques	No plaques	1mm
3594 (II) ST36	1.5mm	1mm	1.5mm
3596 (IIV) ST8	0.5mm	0.5mm with halo to 1.5mm	0.5mm with halo to 1.5mm
E1038 (IIV) ST8	0.5mm, faint plaques	0.5mm, faint plaques	1.5mm
E1139 (IV) ST45	No plaques	0.5mm, faint plaques	No plaques
E1174 (IV) ST22	0.5mm, faint plaques	0.5mm	No plaques
E1185 (IV) ST12	0.5mm, faint plaques	No plaques	No plaques
E1202 (II) ST496	No plaques	1mm	0.5mm
M03/0073 (III) ST239	2mm	0.5mm with halo to 1.5 mm	0.5mm with halo to 1.5mm

2.4.3 Phage adsorption on phage resistant isolates

While some level of phage insensitivity was encountered against all three phages, phage K was the frequently insensitive virion to the *S. aureus* strains tested, and thus, was chosen to evaluate whether or not adsorption inhibition played a role in its insensitivity. Phage K was able to adsorb to all phage-insensitive strains to approximately the same extent as the propagating strain DPC5246. This rules out the possibility of adsorption inhibition playing a role in the narrow host range encountered with phage K in comparison to both phages B1 and JA1 (Supplementary Materials, Figure S2). Additionally, adsorption studies with phages B1 and JA1 indicated that adsorption did not play a role in the differences observed (Supplementary Materials, Figure S3, S4).

2.4.4 Genome comparison between phages B1, JA1 and K

The genome of phage K is 139,831bp in size with long terminal repeats (LTRs) of 8,486bp (Gill, 2014). Genomes of similar sizes were obtained for phages B1 and JA1, these being 140,808bp and 139,484bp, respectively. Examination of sequence reads allowed the identification of LTRs for these phages, due to the identification of a region within their genomes with roughly double the average number of reads, these regions being 8,076bp and 7,651bp in size for phages B1 and JA1, respectively. This approach to the determination of terminal repeats has been utilized for a number of phages (Fouts *et al*., 2013; Li *et al*., 2014; Buttimer *et al.*, 2017). The sequences of all three phages, when analyzed, contained the 12bp inverted repeat sequences 5'-TAAGTACCTGGG-3' and 5'-CCCAGGTACTTA-3', which separates the LTRs from the non-redundant part of the phage DNA, and are characteristic of K-like phages (Łobocka *et al.*, 2012; El Haddad *et al*., 2014). Thus, the entire packaged genome sizes are 148,884bp (B1), 147,135bp (JA1), and 148,317bp (K). Phage K possessed 212 ORFs in its genome (O'Flaherty *et al*., 2004; Gill, 2014), whereas phages B1 and JA1

possessed 219 (Supplementary Materials, Table S1) and 215 ORFs (Supplementary Materials, Table S2) respectively.

Nucleotide pairwise sequence alignment based on BLASTN revealed phages B1 and JA1 (including their LTRs) to be 99% identical to each other, thus can be considered different isolates of the same phage species (Adriaenssens and Brister, 2017). On the other hand, phages B1 and JA1 (including their LTRs) showed 95% and 94% identity (respectively) to phage K, placing these phages on the boundary of speciation.

The examination of 100bp sequences upstream of each ORFs on the non-redundant genome of these phages, using MEME (Bailey *et al.*, 2009), identified 44 and 43 RpoD-like promoters for phages B1 and JA1, respectively. It was observed that these promoters where heavily concentrated in regions with ORFs encoding short hypothetical proteins and those with functions associated with nucleotide metabolism and DNA replication, rather than those associated with virion structure (Supplementary Materials, Table S3, S4). A similar finding was also reported with K-like phage vB_SauM-fRuSau02 (Leskinen *et al.*, 2017). Additionally, 30 Rho-independent terminators were identified on the non-redundant genomes for both B1 and JA1 (Supplementary Materials, Table S5, S6).

Four ORFs present in phage B1 were observed to be absent in JA1 (Table 4). These ORFs encoded two putative terminal repeat-encoded proteins (PhageB1_009, 016) and two other proteins of unknown function (phageB1_202, 203). Although both B1 and JA1 had similar content of ORFs with minor difference between their genomes, both phages varied in their host range on the *S. aureus* strains they infected. This variation is likely attributed to the difference encountered in their genome. Additionally, multiple ORFs present in phage K but absent in both B1 and JA1 were encountered (Figure 2, Table 5). Furthermore, ORFs present

in both phages B1 and JA1 but absent in K were also encountered (Figure 2, Table 6). These ORFs are discussed below.

Table 4. List of missing ORFs predicted to be present in phage B1 but absent in phage JA1.

ORFs	**Amino acid numbers**	**Protein size (kDa)**	**Predicted function**
PhageB1_009	112	13.5	Terminal repeat encoded protein
PhageB1_016	107	12.4	Terminal repeat encoded protein
PhageB1_202	32	3.5	Unknown
PhageB1_203	104	11.6	Unknown

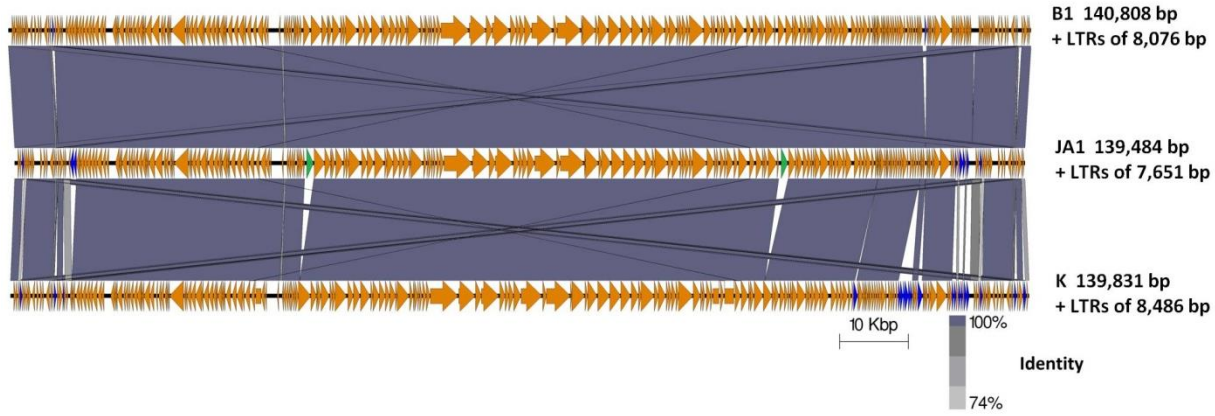


Figure 2. Genome comparison of phages B1, JA1, and K (including their long terminal repeats) using currently available annotations employing BLASTN and visualized with Easyfig. Regions of sequence similarity are connected by the shaded area, using a grey scale; genome maps consisting of orange arrows indicating the location of ORFs along the phage genomes, with unshared ORFs highlighted in blue with those indicating unshared homing endonuclease highlighted in green.

Table 5. List of missing ORFs and their predicted putative functions absent in both phages B1 and JA1 but present in phage K.

ORFs	**Amino acid number**	**Predicted Protein size (kDa)**	**Predicted function**
PhageK_004	108	12.7	Unknown
PhageK_016*	107	12.4	Unknown
PhageK_019	57	4.7	Unknown
PhageK_020	89	10.2	Unknown
PhageK_168	185	21.7	Predicted to contain a transmembrane region based on InterProScan
PhageK_187	101	11.7	Unknown
PhageK_188	123	13.8	Predicted to contain a transmembrane region based on InterProScan
PhageK_189	78	9.2	Unknown
PhageK_190	175	20.6	Predicted as a putative metallophoshatase
PhageK_191	106	12.9	Unknown
PhageK_192	76	8.9	Predicted to contain a transmembrane region based on InterProScan
PhageK_196	226	25.8	Unknown
PhageK_205	83	9.7	Unknown
PhageK_206	98	11.2	Unknown
PhageK_208	99	11.6	Unknown
PhageK_209	75	8.9	Unknown
PhageK_211	117	13.9	Predicted to possess a transmembrane region based on InterProScan
PhageK_212	128	15.6	Unknown

* ORF that phage JA1 does not share with phage K

Table 6. List of missing ORFs and their predicted function absent in phage K but present in phages B1 and JA1.

ORFs	**Amino acid number**	**Predicted Protein size (kDa)**	**Predicted function**
PhageJA1_003 (PhageB1_003)	96	11.3	Unknown
PhageJA1_020 (PhageB1_022)	161	19.1	Unknown
PhageJA1_021 (PhageB1_023)	135	16.5	Unknown
PhageJA1_084 (PhageB1_087)	323	39.6	Predicted as a putative endonuclease interrupting the terminase large subunit [PhageJA1_083 (PhageB1_086) and PhageJA1_085 (PhageB1_088)]
PhageJA1_152 (PhageB1_155)	322	38.3	Predicted as a putative endonuclease containing a LAGLIDADG-like domain and an Intein splicing domain and interrupts the DNA repair protein [PhageJA1_151 (PhageB1_154) and PhageJA1_153 (PhageB1_156)]
PhageJA1_206 (PhageB1_212)	73	8.9	Unknown
PhageJA1_208 (PhageB1_214)	169	20.3	HHpred indicates homology to cell wall hydrolases
PhageJA1_209 (PhageB1_215)	109	12.6	Unknown
PhageJA1_211 (PhageB1_217)	104	12.0	Unknown
PhageJA1_212 (PhageB1_218)	55	6.5	Unknown
PhageJA1_213 (PhageB1_219)	33	3.7	Predicted to possess a transmembrane region based on InterProScan

2.4.5 Characteristic features of phage K ORFs absent in both JA1 and B1

Seventeen ORFs present in phage K were absent in both phages B1 and JA1, with one additional ORF found not to be shared between JA1 and K. These ORFs are listed in Table 5. No function could be assigned to these with the exception of phageK_190, which based on NCBI conserved domain search possessed a metallophosphatase-like domain (cd07390; E value; $3.94x10^{-30}$) and is a member of the metallophosphatase (MPP) superfamily. Families within this superfamily of enzymes are functionally diverse, involved in the cleavage of phosphoester bonds, and include Mre11/SbcD-like exonucleases, Dbr1-like RNA lariat debranching enzymes, YfcE-like phosphodiesterases, purple acid phosphatases (PAPs), YbbF-like UDP-2,3-diacylglucosamine hydrolases, and acid sphingomyelinases (ASMases) (Matange *et al*., 2015).

2.4.6 Characteristic features of phages B1 and JA1 ORFs absent in phage K

Eleven ORFs present in both phages B1 and JA1 were absent in phage K (Table 6). No putative function could be assigned to the majority of these ORFs based on BLASTP, InterProScan or HHpred analysis, with the exception of phageJA1_084 (phageB1_087) and phageJA1_152 (phageB1_155), which encoded homing endonucleases interrupting both the terminase large subunit and the DNA repair protein, respectively. These homing endonucleases are site-specific DNA endonucleases capable of initiating DNA breaks leading to repair and recombination event that results in the integration of this endonuclease ORF into a gene that was previously lacking it (Gogarten and Hilario, 2006). The presence of these mobile genetic elements is common among known staphylococcal phages of the subfamily *Spounavirinae*, and these endonucleases ORFs are known to insert themselves into essential phage genes (Vandersteegen *et al*., 2013; Leskinen *et al*., 2017). Additionally, HHpred analysis indicated ORFs PhageJA1_208 and PhageB1_214 to possess remote homology to cell-degrading proteins. The majority of these ORFs were found to be located next to the

genome termini of JA1 and B1, with genes located in this region having been previously reported in similar phages to be expressed early in phage development (Łobocka *et al*., 2012). Such proteins are usually involved in subversion of the host's machinery to aid phage takeover (Wei and Stewart, 1993; Stewart *et al*., 1998).

2.4.7 Comparison of phages K, B1 and JA1 with other similar therapeutic phages (Team1, vB_SauM-fRuSau02, Sb-1 and ISP)

Four additional staphylococcal phages that originate in commercial phage therapeutic mixtures are Team1, vB_SauM-fRuSau02, Sb-1 and ISP, as discussed earlier (Markoishvili *et al*., 2002; Jikia *et al.*, 2005; Kvachadze *et al*., 2011; Vandersteegen *et al*., 2011; El Haddad *et al*., 2014; Leskinen *et al*., 2017). These phages were also reported to possess wide host ranges towards a number of clinically relevant *S. aureus* strains. Although similar, these phages have several feature differences from each other and from phages B1 and JA1. Comparison of nucleotide identities (BLASTN) with phage K shows that they belong to the genus *Kayvirus* (Supplementary Materials, Table S7) possessing genomes of similar sizes, apart from Sb-1, being smaller than would be expected, suggesting the genome submission may have been incomplete (Figure 3). Additionally, the arrangement of ORFs is quite similar. Furthermore, tRNA genes of these phages were also examined. All seven phages were found to possess the same four tRNA genes for methionine, tryptophan, phenylalanine, and aspartic acid (Supplementary Materials, Table S8). The eleven ORFs which were present in B1 and JA1 but absent in K (Table 6, Supplementary Materials, Figure S5) were similarly present in Team 1, vB_SauM-fRuSau02, Sb-1 and ISP. And likewise, the ORFs present in K, but absent in both B1 and JA1, were also missing in these phages. However, vB_SauM-fRuSau02 possesses a much shorter putative tail protein (RS_159) of 73 amino acids compared to the phage K counterpart (PhageK_151) of 170 amino acids. Non-hypothetical proteins that differed between these phages were a membrane protein (Phage B1_180, PhageJA1_177, and

Phage_170) and an ATPase-like protein (Protein id: CCA65911.1 for phage ISP). Other ORFs that differed among these phages were mostly hypothetical proteins.

S. aureus employ several defense strategies against viral attack (Hyman and Abedon, 2010; Seed, 2015) and these, such as restriction modification systems (Roberts *et al*., 2013) and CRISPR-Cas systems (Cao *et al*., 2016), may vary from strain to strain. These defenses along with several variations encountered at the genetic level across phages B1, JA1, and K may explain the differences in host ranges observed in this study.

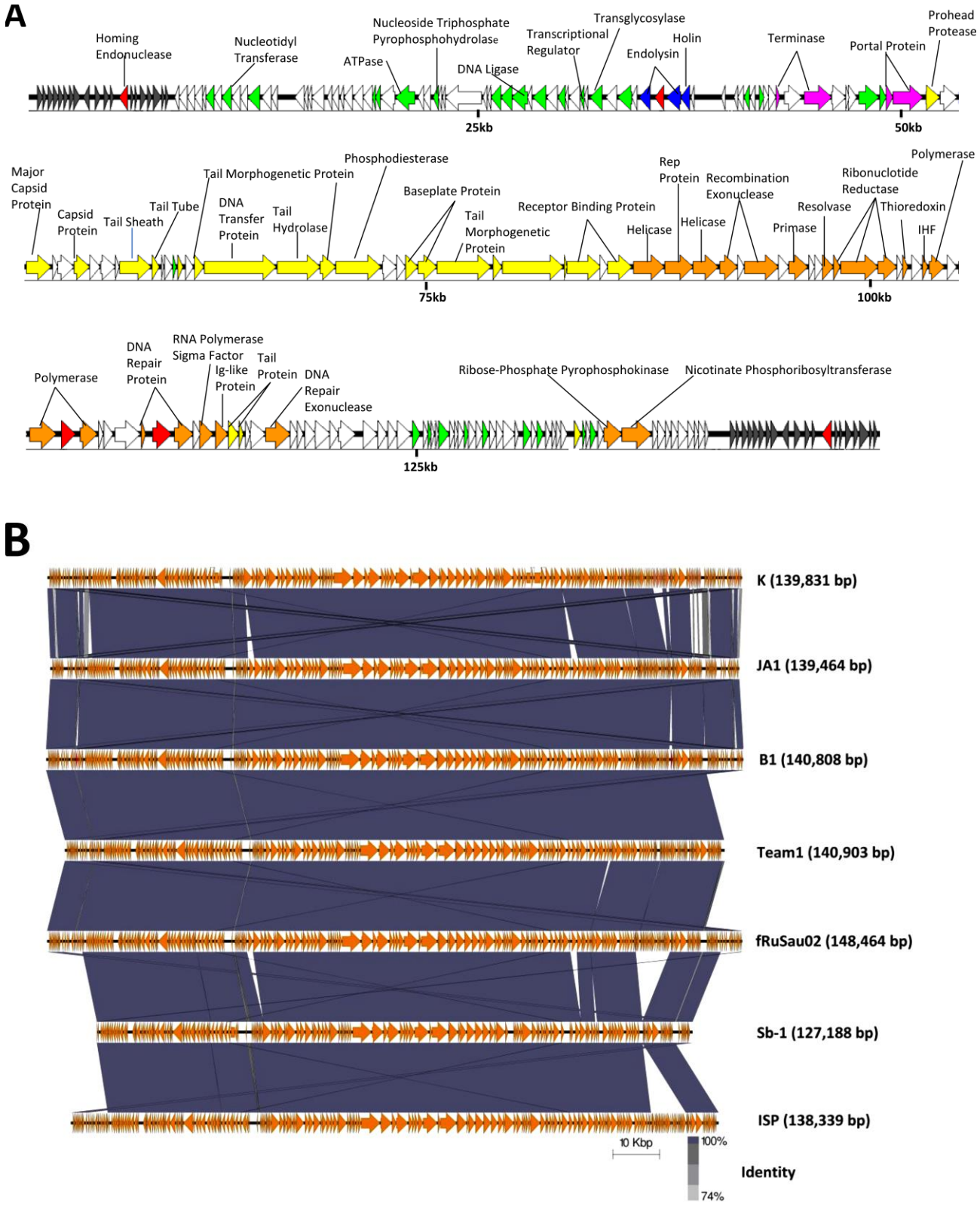


Figure 3. Genome organization of phage JA1 with its predicted genes indicated by arrows (A). Genome comparison of phage K with the six staphylococcal phages employed in commercial phages mixture consisting of B1, JA1, Team 1 (Markoishvili *et al.*, 2002; Jikia *et al.*, 2005; El Haddad *et al.*, 2014), vB_SauM-fRuSau02 (Leskinen *et al.*, 2017), Sb-1 (Kvachadze *et al.*, 2011) and ISP (Vandersteegen *et al.*, 2011) using currently available annotations employing BLASTN and visualized with Easyfig (B).

2.5 Conclusions

Host range of three highly similar phages was performed in this study, and it was identified that phages B1 and JA1 from the Fersisi commercial phage mix had a much broader host range in comparison to phage K on a representative Irish bank of clinical MRSA sequence type isolates. Comparisons of their genomes lead to the identification of several ORFs absent in phage K, but present in both phages B1 and JA1. These ORFs were also identified in several other staphylococcal phages sourced from commercial phage mixtures (B1, JA1, Team 1 (Markoishvili *et al*., 2002; Jikia *et al*., 2005; El Haddad *et al*., 2014), vB_SauM-fRuSau02 (Leskinen *et al*., 2017), Sb-1 (Kvachadze *et al*., 2011) and ISP (Vandersteegen *et al*., 2011), also with a reported wide host range. The exact role of these ORFs is currently unknown. However, these ORFs along with several variations encountered at the genetic level between these phages may, in part, explain their different host range. Unfortunately, information is lacking on the influences of various phage resistance systems, which may be active in *Staphylococcus aureus*. Phage research also needs to focus more on elucidation of the functions of hypothetical proteins to allow greater understanding of how phages overcome such systems.

2.6 References

1. Abedon, S. T., Kuhl, S. J., Blasdel, B. G. and Kutter, E. M. (2011) 'Phage treatment of human infections', *Bacteriophage*, 1(2), pp. 66–85. doi: 10.4161/bact.1.2.15845.

2. Ackermann, H. W. (1996) 'Frequency of morphological phage descriptions in 1995.', *Archives of Virology*, 141(2), pp. 209–218. doi: 10.1007/s007050170120.

3. Adriaenssens, E. and Brister, J. R. (2017) 'How to name and classify your phage: An informal guide', *Viruses*, 9(4), p. 70. doi: 10.3390/v9040070.

4. Adriaenssens, E. M., Krupovic, M., Knezevic, P., Ackermann, H. W., Barylski, J., Brister, J. R., Clokie, M. R. C., Duffy, S., Dutilh, B. E., Edwards, R. A., Enault, F., Jang, H. Bin, Klumpp, J., Kropinski, A. M., Lavigne, R., Poranen, M. M., Prangishvili, D., Rumnieks, J., Sullivan, M. B., Wittmann, J., Oksanen, H. M., Gillis, A. and Kuhn, J. H. (2017) 'Taxonomy of prokaryotic viruses: 2016 update from the ICTV bacterial and archaeal viruses subcommittee', *Archives of Virology*, 162(4), pp. 1153–1157. doi: 10.1007/s00705-016-3173-4.

5. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. and Noble, W. S. (2009) 'MEME SUITE: tools for motif discovery and searching.', *Nucleic Acids Research*, 37(Web Server issue), pp. W202-208. doi: 10.1093/nar/gkp335.

6. Bertozzi Silva, J., Storms, Z. and Sauvageau, D. (2016) 'Host receptors for bacteriophage adsorption', *FEMS Microbiology Letters*, 363(4), pp. 1–11. doi: 10.1093/femsle/fnw002.

7. Besemer, J., Lomsadze, A. and Borodovsky, M. (2001) 'GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions.', *Nucleic Acids Research*. Oxford University Press, 29(12), pp. 2607–2618.

8. Buttimer, C., Hendrix, H., Oliveira, H., Casey, A., Neve, H., McAuliffe, O., Paul Ross, R., Hill, C., Noben, J. P., O'Mahony, J., Lavigne, R. and Coffey, A. (2017) 'Things are getting hairy: Enterobacteria bacteriophage vB_PcaM_CBB', *Frontiers in Microbiology*, 8(44), pp. 1–16. doi: 10.3389/fmicb.2017.00044.

9. Cao, L., Gao, C. H., Zhu, J., Zhao, L., Wu, Q., Li, M. and Sun, B. (2016) 'Identification and functional study of type III-A CRISPR-Cas systems in clinical isolates of *Staphylococcus aureus*', *International Journal of Medical Microbiology*. Elsevier GmbH., 306(8), pp. 686–696. doi: 10.1016/j.ijmm.2016.08.005.

10. Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M.-A., Barrell, B. G. and Parkhill, J. (2005) 'ACT: the Artemis comparison tool', *Bioinformatics*, 21(16), pp. 3422–3423. doi: 10.1093/bioinformatics/bti553.

11. Cui, Z., Feng, T., Gu, F., Li, Q., Dong, K., Zhang, Y., Zhu, Y., Han, L., Qin, J. and Guo, X. (2017) 'Characterization and complete genome of the virulent *Myoviridae* phage JD007 active against a variety of *Staphylococcus aureus* isolates from different hospitals in Shanghai, China', *Virology Journal*, 14(1), p. 26. doi: 10.1186/s12985-017-0701-0.

12. Delcher, A. (1999) 'Improved microbial gene identification with GLIMMER', *Nucleic Acids Research*, 27(23), pp. 4636–4641. doi: 10.1093/nar/27.23.4636.

13. Fouts, D. E., Klumpp, J., Bishop-Lilly, K. A., Rajavel, M., Willner, K. M., Butani, A., Henry, M., Biswas, B., Li, M., Albert, M. J., Loessner, M. J., Calendar, R. and Sozhamannan, S. (2013) 'Whole genome sequencing and comparative genomic analyses of two *Vibrio cholerae* O139 Bengal-specific Podoviruses to other N4-like phages reveal extensive genetic diversity', *Virology Journal*, 10, p. 165. doi: 10.1186/1743-422X-10-165.

14. Gill, J. J. (2014) 'Revised genome sequence of *Staphylococcus aureus* bacteriophage K',

2(1), pp. 12–13. doi: 10.1128/genomeA.01173-13.

15. Gogarten, J. P. and Hilario, E. (2006) ‘Inteins, introns, and homing endonucleases: recent revelations about the life cycle of parasitic genetic elements.’, *BMC Evolutionary Biology*. BioMed Central, 6, p. 94. doi: 10.1186/1471-2148-6-94.

16. Gutiérrez, D., Vandenheuvel, D., Martínez, B., Rodríguez, A., Lavigne, R. and García, P. (2015) ‘Two phages, phiIPLA-RODI and phiIPLA-C1C, lyse mono- and dual-species staphylococcal biofilms’, *Applied and Environmental Microbiology*, 81(10), pp. 3336–3348. doi: 10.1128/AEM.03560-14.

17. El Haddad, L., Abdallah, N. Ben, Plante, P. L., Dumaresq, J., Katsarava, R., Labrie, S., Corbeil, J., St-Gelais, D. and Moineau, S. (2014) ‘Improving the safety of *Staphylococcus aureus* polyvalent phages by their production on a *Staphylococcus xylosus* strain’, *PLoS One*, 9(7), pp. 1–10. doi: 10.1371/journal.pone.0102600.

18. Hall, A. R., Scanlan, P. D. and Buckling, A. (2011) ‘Bacteria-phage coevolution and the emergence of generalist pathogens’, *The American Naturalist*, 177(1), pp. 44–53. doi: 10.1086/657441.

19. Hall, J. P. J., Harrison, E. and Brockhurst, M. A. (2013) ‘Viral host-adaptation: Insights from evolution experiments with phages’, *Current Opinion in Virology*. Elsevier B.V., 3(5), pp. 572–577. doi: 10.1016/j.coviro.2013.07.001.

20. Hiramatsu, K., Cui, L., Kuroda, M. and Ito, T. (2001) ‘The emergence and evolution of methicillin-resistant *Staphylococcus aureus*.’, *Trends in Microbiology*, 9(10), pp. 486–493.

21. Hyman, P. and Abedon, S. T. (2010) ‘Bacteriophage host range and bacterial resistance.’, *Advances in Applied Microbiology*. 1st edn. Elsevier Inc., 70(10), pp. 217–248. doi: 10.1016/S0065-2164(10)70007-1.

22. Jevons, M. P. (1961) '"Celbenin" - resistant Staphylococci', *Bmj*, 1(5219), pp. 124–125. doi: 10.1136/bmj.1.5219.124-a.

23. Jikia, D., Chkhaidze, N., Imedashvili, E., Mgaloblishvili, I., Tsitlanadze, G., Katsarava, R., Morris, J. G. and Sulakvelidze, A. (2005) 'The use of a novel biodegradable preparation capable of the sustained release of bacteriophages and ciprofloxacin, in the complex treatment of multidrug-resistant *Staphylococcus aureus*-infected local radiation injuries caused by exposure to Sr90', *Clinical and Experimental Dermatology*, 30(1), pp. 23–26. doi: 10.1111/j.1365-2230.2004.01600.x.

24. Klein, E., Smith, D. L. and Laxminarayan, R. (2007) 'Hospitalizations and deaths caused by methicillin-resistant *Staphylococcus aureus* , United States, 1999–2005', *Emerging Infectious Diseases*, 13(12), pp. 1840–1846. doi: 10.3201/eid1312.070629.

25. Kvachadze, L., Balarjishvili, N., Meskhi, T., Tevdoradze, E., Skhirtladze, N., Pataridze, T., Adamia, R., Topuria, T., Kutter, E., Rohde, C. and Kutateladze, M. (2011) 'Evaluation of lytic activity of staphylococcal bacteriophage Sb-1 against freshly isolated clinical pathogens', *Microbial Biotechnology*, 4(5), pp. 643–650. doi: 10.1111/j.1751-7915.2011.00259.x.

26. Labrie, S. J., Samson, J. E. and Moineau, S. (2010) 'Bacteriophage resistance mechanisms', *Nature Reviews Microbiology*. Nature Publishing Group, 8(5), pp. 317–327. doi: 10.1038/nrmicro2315.

27. Laslett, D. and Canback, B. (2004) 'ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences', *Nucleic Acids Research*. Oxford University Press, 32(1), pp. 11–16. doi: 10.1093/nar/gkh152.

28. Leskinen, K., Tuomala, H., Wicklund, A., Horsma-Heikkinen, J., Kuusela, P., Skurnik, M.

and Kiljunen, S. (2017) ‘Characterization of vB_SauM-fRuSau02, a Twort-like bacteriophage isolated from a therapeutic phage cocktail’, *Viruses*, 9(9), p. 258. doi: 10.3390/v9090258.

29. Li, S., Fan, H., An, X., Fan, H., Jiang, H., Chen, Y. and Tong, Y. (2014) ‘Scrutinizing virus genome termini by high-throughput sequencing’, *PLoS One*, 9(1). p. e85806 doi: 10.1371/journal.pone.0085806.

30. Li, X., Koç, C., Kühner, P., Stierhof, Y. D., Krismer, B., Enright, M. C., Penadés, J. R., Wolz, C., Stehle, T., Cambillau, C., Peschel, A. and Xia, G. (2016) ‘An essential role for the baseplate protein Gp45 in phage adsorption to *Staphylococcus aureus*.’, *Scientific Reports*. Nature Publishing Group, 6, p. 26455. doi: 10.1038/srep26455.

31. Łobocka, M., Hejnowicz, M. S., Dąbrowski, K., Gozdek, A., Kosakowski, J., Witkowska, M., Ulatowska, M. I., Weber-Dąbrowska, B., Kwiatek, M., Parasion, S., Gawor, J., Kosowska, H. and Głowacka, A. (2012) ‘Genomics of staphylococcal Twort-like phages--potential therapeutics of the post-antibiotic era.’, *Advances in Virus Research*, 83, pp. 143–216. doi: 10.1016/B978-0-12-394438-2.00005-0.

32. Lowe, T. M. and Eddy, S. R. (1997) ‘tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence.’, *Nucleic Acids Research*, 25(5), pp. 955–64.

33. Markoishvili, K., Tsitlanadze, G., Katsarava, R., Morris, J. G. and Sulakvelidze, A. (2002) ‘A novel sustained-release matrix based on biodegradable poly(ester amide)s and impregnated with bacteriophages and an antibiotic shows promise in management of infected venous stasis ulcers and other poorly healing wounds’, *International Journal of Dermatology*, 41(7), pp. 453–458. doi: 10.1046/j.1365-4362.2002.01451.x.

34. Matange, N., Podobnik, M. and Visweswariah, S. S. S. (2015) ‘Metallophosphoesterases:

structural fidelity with functional promiscuity.', *The Biochemical Journal*, 467(2), pp. 201–216. doi: 10.1042/BJ20150028.

35. Mitchell, A., Chang, H. Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C., McMenamin, C., Nuka, G., Pesseat, S., Sangrador-Vegas, A., Scheremetjew, M., Rato, C., Yong, S. Y., Bateman, A., Punta, M., Attwood, T. K., Sigrist, C. J. A., Redaschi, N., Rivoire, C., Xenarios, I., Kahn, D., Guyot, D., Bork, P., Letunic, I., Gough, J., Oates, M., Haft, D., Huang, H., Natale, D. A., Wu, C. H., Orengo, C., Sillitoe, I., Mi, H., Thomas, P. D. and Finn, R. D. (2015) 'The InterPro protein families database: the classification resource after 15 years.', *Nucleic Acids Research*, 43(Database issue), pp. D213-221. doi: 10.1093/nar/gku1243.

36. Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A. and Gautheret, D. (2011) 'ARNold: A web tool for the prediction of rho-independent transcription terminators', *RNA Biology*, 8(1). pp. 11-13 doi: 10.4161.rna.8.1.13346.

37. O'Flaherty, S., Coffey, A., Edwards, R., Meaney, W., Fitzgerald, G. F. and Ross, R. P. (2004) 'Genome of staphylococcal phage K: A new lineage of *Myoviridae* infecting Gram-positive bacteria with a low G+C content', *Journal of Bacteriology*, 186(9), pp. 2862–2871. doi: 10.1128/JB.186.9.2862-2871.2004.

38. O'Flaherty, S., Ross, R. P. and Coffey, A. (2009) 'Bacteriophage and their lysins for elimination of infectious bacteria', *FEMS Microbiology Reviews*, 33(4), pp. 801–819. doi: 10.1111/j.1574-6976.2009.00176.x.

39. O'Flaherty, S., Ross, R. P., Meaney, W., Fitzgerald, G. F., Elbreki, M. F. and Coffey, A. (2005) 'Potential of the polyvalent anti-*Staphylococcus* bacteriophage K for control of antibiotic-resistant staphylococci from hospitals', *Applied and Environmental Microbiology*, 71(4), pp. 1836–1842. doi: 10.1128/AEM.71.4.1836-1842.2005.

40. Roberts, G. A., Houston, P. J., White, J. H., Chen, K., Stephanou, A. S., Cooper, L. P., Dryden, D. T. F. and Lindsay, J. A. (2013) ‘Impact of target site distribution for Type i restriction enzymes on the evolution of methicillin-resistant *Staphylococcus aureus* (MRSA) populations’, *Nucleic Acids Research*, 41(15), pp. 7472–7484. doi: 10.1093/nar/gkt535.

41. Rossney, A. S., Lawrence, M. J., Morgan, P. M., Fitzgibbon, M. M., Shore, A., Coleman, D. C., Keane, C. T. and O’Connell, B. (2006) ‘Epidemiological typing of MRSA isolates from blood cultures taken in Irish hospitals participating in the European Antimicrobial Resistance Surveillance System (1999-2003)’, *European Journal of Clinical Microbiology and Infectious Diseases*, 25(2), pp. 79–89. doi: 10.1007/s10096-006-0091-5.

42. Rountree, P. M. (1949) ‘The serological differentiation of staphylococcal bacteriophages.’, *Journal of General Microbiology*, 3(2), pp. 164–173. doi: 10.1099/00221287-3-2-164.

43. Sambrook, J. and Russell, D. W. (2001) ‘Purification of bacteriophage lamda particles by isopycnic centrifugation through CsCl gradients’, in *Molecular Cloning: A Laboratory Manual, Volume 1*. Cold Spring Harbor Laboratory Press, p. 2.47.

44. Samson, J. E., Magadán, A. H., Sabri, M. and Moineau, S. (2013) ‘Revenge of the phages: Defeating bacterial defences’, *Nature Reviews Microbiology*. Nature Publishing Group, 11(10), pp. 675–687. doi: 10.1038/nrmicro3096.

45. Schmelcher, M. and Loessner, M. J. (2014) ‘Application of bacteriophages for detection of foodborne pathogens’, *Bacteriophage*, 4(2), p. e28137. doi: 10.4161/bact.28137.

46. Seed, K. D. (2015) ‘Battling phages: How bacteria defend against viral attack.’, *PLoS Pathogens*, 11(6), p. e1004847. doi: 10.1371/journal.ppat.1004847.

47. Shore, A. C., Rossney, A. S., O’Connell, B., Herra, C. M., Sullivan, D. J., Humphreys, H. and Coleman, D. C. (2008) ‘Detection of staphylococcal cassette chromosome mec-

associated DNA segments in multiresistant methicillin-susceptible *Staphylococcus aureus* (MSSA) and identification of *Staphylococcus epidermidis* ccrAB4 in both methicillin-resistant *S. aureus* and MSSA', *Antimicrobial Agents and Chemotherapy*, 52(12), pp. 4407–4419. doi: 10.1128/AAC.00447-08.

48. Söding, J., Biegert, A. and Lupas, A. N. (2005) 'The HHpred interactive server for protein homology detection and structure prediction.', *Nucleic Acids Research*, 33(Web Server issue), pp. W244-8. doi: 10.1093/nar/gki408.

49. Stewart, C. R., Gaslightwala, I., Hinata, K., Krolikowski, K. A., Needleman, D. S., Peng, A. S. Y., Peterman, M. A., Tobias, A. and Wei, P. (1998) 'Genes and regulatory sites of the "host-takeover module" in the terminal redundancy of *Bacillus subtilis* bacteriophage SPO1', *Virology*, 246(2), pp. 329–340. doi: 10.1006/viro.1998.9197.

50. Sulakvelidze, A., Alavidze, Z. and Morris, J. G. (2001) 'Bacteriophage therapy.', *Antimicrobial Agents and Chemotherapy*. American Society for Microbiology, 45(3), pp. 649–659. doi: 10.1128/AAC.45.3.649-659.2001.

51. Sullivan, M. J., Petty, N. K. and Beatson, S. A. (2011) 'Easyfig: a genome comparison visualizer', *Bioinformatics*, 27(7), pp. 1009–1010. doi: 10.1093/bioinformatics/btr039.

52. Tong, S. Y. C., Davis, J. S., Eichenberger, E., Holland, T. L. and Fowler, V. G. (2015) '*Staphylococcus aureus* infections: Epidemiology, pathophysiology, clinical manifestations, and management', *Clinical Microbiology Reviews*, 28(3), pp. 603–661. doi: 10.1128/CMR.00134-14.

53. Vandersteegen, K., Kropinski, A. M., Nash, J. H. E., Noben, J. P., Hermans, K. and Lavigne, R. (2013) 'Romulus and Remus, two phage isolates representing a distinct clade within the Twortlikevirus genus, display suitable properties for phage therapy applications.', *Journal of*

Virology, 87(6), pp. 3237–3247. doi: 10.1128/JVI.02763-12.

54. Vandersteegen, K., Mattheus, W., Ceyssens, P. J., Bilocq, F., de Vos, D., Pirnay, J. P., Noben, J. P., Merabishvili, M., Lipinska, U., Hermans, K. and Lavigne, R. (2011) 'Microbiological and molecular assessment of bacteriophage ISP for the control of *Staphylococcus aureus*', *PLoS ONE*, 6(9). P. e24418. doi: 10.1371/journal.pone.0024418.

55. Villarroel, J., Larsen, M. V., Kilstrup, M. and Nielsen, M. (2017) 'Metagenomic analysis of therapeutic PYO phage cocktails from 1997 to 2014', *Viruses*, 9(11). p. 328. doi: 10.3390/v9110328.

56. Wei, P. and Stewart, C. R. (1993) 'A cytotoxic early gene of *Bacillus subtilis* bacteriophage SPOl', *Journal of Bacteriology*, 175(24), pp. 7887–7900.

57. Wittebole, X., De Roock, S. and Opal, S. M. (2014) 'A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens.', *Virulence*, 5(1), pp. 226–235. doi: 10.4161/viru.25991.

58. Zschach, H., Joensen, K. G., Lindhard, B., Lund, O., Goderdzishvili, M., Chkonia, I., Jgenti, G., Kvatadze, N., Alavidze, Z., Kutter, E. M., Hasman, H. and Larsen, M. V. (2015) 'What can we learn from a metagenomic analysis of a georgian bacteriophage cocktail?', *Viruses*, 7(12), pp. 6570–6589. doi: 10.3390/v7122958.

59. Zuker, M. (2003) 'Mfold web server for nucleic acid folding and hybridization prediction.', *Nucleic Acids Research*, 31(13), pp. 3406–15.

Supplementary Materials for Chapter II

Table S1: Annotation of the Staphylococcal phage vB_SauM_B1 genome.

ORF	Predicted function	Start	stop	F/R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageB1_001	putative terminal repeat encoded protein	486	785	F	99	11.6	ATG	TAA	ORF150 [Staphylococcus virus G1]	100%	5.00E-65	YP_241022.1
PhageB1_002	putative terminal repeat encoded protein	801	986	F	61	6.8	ATG	TAG	TreB [Staphylococcus phage A5W]	100%	9.00E-31	AHL83358.1
PhageB1_003	putative terminal repeat encoded protein	1093	1383	F	96	11.3	ATG	TAA	ORF156 [Staphylococcus virus G1]	100%	2.00E-61	YP_241024.1
PhageB1_004	putative terminal repeat encoded protein	1383	1670	F	95	10.9	ATG	TAA	ORF158 [Staphylococcus virus G1]	100%	2.00E-61	YP_241025.1
PhageB1_005	putative terminal repeat encoded protein	1670	1963	F	97	11.5	ATG	TAA	ORF154 [Staphylococcus virus G1]	100%	3.00E-64	YP_241026.1
PhageB1_006	putative terminal repeat encoded protein	1967	2224	F	85	10.2	ATG	TAG	ORF175 [Staphylococcus virus G1]	100%	5.00E-55	YP_241027.1
PhageB1_007	putative terminal repeat encoded protein	2302	2541	F	79	9.2	ATG	TAG	ORF183 [Staphylococcus virus G1]	100%	2.00E-50	YP_241028.1
PhageB1_008	putative terminal repeat encoded protein	2552	2899	F	115	13.7	ATG	TGA	ORF125 [Staphylococcus virus G1]	100%	2.00E-77	YP_241029.1
PhageB1_009	putative terminal repeat encoded protein	3446	3108	R	112	13.5	ATG	TAA	ORF128 [Staphylococcus phage G1]	100%	2.00E-71	YP_241030.1
PhageB1_010	putative terminal repeat encoded protein	3757	4065	F	102	11.8	ATG	TAA	ORF145 [Staphylococcus virus G1]	100%	1.00E-69	YP_241031.1
PhageB1_011	putative terminal repeat encoded	4271	4555	F	94	11.0	ATG	TAA	ORF159 [Staphylococcus virus	100%	8.00E-64	YP_241032.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
	protein								G1]			
PhageB1_012	putative terminal repeat encoded protein	4630	4821	F	63	7.7	ATG	TAA	ORF221 [Staphylococcus virus G1]	100%	1.00E-37	YP_241033.1
PhageB1_013	putative HNH homing endonuclease	5626	5138	R	162	19.6	ATG	TAA	ORF085 [Staphylococcus virus G1]	100%	5.00E-115	YP_241035.1
PhageB1_014	putative terminal repeat encoded protein	5794	5952	F	52	6.1	ATG	TAA	hypothetical protein [Staphylococcus phage Team1]	100%	3.00E-27	YP_009098148.1
PhageB1_015	putative terminal repeat encoded protein	6022	6153	F	43	5.2	ATG	TAA	ORF297 [Staphylococcus virus G1]	100%	1.00E-21	YP_241036.1
PhageB1_016	putative terminal repeat encoded protein	6321	6644	F	107	12.4	ATG	TAA	ORF135 [Staphylococcus phage G1]	99%	3.00E-71	YP_241037.1
PhageB1_017	Hypothetical protein	6744	6980	F	78	9.1	ATG	TAG	hypothetical protein CPT_phageK_gp017 [Staphylococcus virus K]	100%	1.00E-47	YP_009041239.1
PhageB1_018	putative terminal repeat encoded protein	7060	7530	F	156	17.8	ATG	TAG	ORF092 [Staphylococcus virus G1]	100%	3.00E-105	YP_241038.1
PhageB1_019	putative terminal repeat encoded protein	7560	7685	F	41	4.6	ATG	TAA	ORF166 [Staphylococcus virus G1]	100%	2.00E-20	YP_241041.1
PhageB1_020	putative terminal repeat encoded protein	7770	7949	F	59	7.2	ATG	TAA	hypothetical protein [Staphylococcus phage ISP]	100%	4.00E-33	CCA65883.1
PhageB1_021	Hypothetical protein	8519	8283	R	78	9.6	ATG	TAA	hypothetical protein SA5_0153/152 [Staphylococcus phage SA5]	100%	4.00E-50	AFV80807.1
PhageB1_022	Hypothetical protein	9006	8521	R	161	19.1	ATG	TAA	ORF088 [Staphylococcus virus	100%	3.00E-111	YP_241045.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
									G1]			
PhageB1_ 023	Hypothetical protein	9426	9019	R	135	16.5	ATG	TAA	ORF109 [Staphylococcus virus G1]	100%	3.00E-94	YP_241046.1
PhageB1_ 024	Hypothetical protein	9857	9426	R	143	17.3	ATG	TAA	ORF103 [Staphylococcus virus G1]	100%	5.00E-96	YP_241047.1
PhageB1_ 025	Hypothetical protein	10051	9860	R	63	7.9	ATG	TAA	ORF224 [Staphylococcus virus G1]	100%	1.00E-36	YP_241048.1
PhageB1_ 026	putative membrane protein	10533	10048	R	161	18.3	ATG	TGA	conserved hypothetical protein [Staphylococcus phage Sb-1]	100%	1.00E-110	YP_008873528.1
PhageB1_ 027	Hypothetical protein	10957	10526	R	143	16.7	ATG	TAA	ORF104 [Staphylococcus virus G1]	100%	2.00E-98	YP_241050.1
PhageB1_ 028	putative nucleotidyl transferase	11513	10971	R	180	21.5	ATG	TAA	ORF073 [Staphylococcus virus G1]	100%	4.00E-126	YP_241051.1
PhageB1_ 029	Hypothetical protein	12013	11525	R	162	19.5	ATG	TAG	ORF086 [Staphylococcus virus G1]	100%	7.00E-118	YP_241052.1
PhageB1_ 030	Hypothetical protein	12424	12026	R	132	16.1	ATG	TAA	ORF111 [Staphylococcus virus G1]	100%	1.00E-89	YP_241053.1
PhageB1_ 031	putative serine threonine protein phosphatase I	13128	12421	R	235	27.7	ATG	TGA	serine threonine protein phosphatase I [Staphylococcus phage JD007]	100%	2.00E-172	YP_007112835.1
PhageB1_ 032	Hypothetical protein	13779	13228	R	183	21.0	ATG	TAA	ORF070 [Staphylococcus phage G1]	99%	2.00E-128	, YP_241055.1
PhageB1_ 033	putative tail protein	14115	13798	R	105	11.8	GTG	TAA	ORF138 [Staphylococcus virus	99%	4.00E-67	YP_241056.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
									G1]			
PhageB1_ 034	Hypothetical protein	15649	15101	R	182	21.9	ATG	TGA	ORF071 [Staphylococcus virus G1]	99%	8.00E-125	YP_241057.1
PhageB1_ 035	Hypothetical protein	15871	15653	R	72	8.4	ATG	TAA	ORF201 [Staphylococcus virus G1]	100%	4.00E-44	YP_241058.1
PhageB1_ 036	Hypothetical protein	16066	15872	R	64	7.6	ATG	TAA	ORF218 [Staphylococcus virus G1]	100%	2.00E-38	YP_241059.1
PhageB1_ 037	Hypothetical protein	16793	16056	R	245	28.7	ATG	TAA	ORF050 [Staphylococcus virus G1]	100%	3.00E-175	YP_241060.1
PhageB1_ 038	Hypothetical protein	16960	16856	R	34	4.1	ATG	TAA	ORF437 [Staphylococcus virus G1]	100%	1.00E-14	YP_241061.1
PhageB1_ 039	Hypothetical protein	17211	16972	R	79	9.4	ATG	TAA	gpORF020 [Staphylococcus phage A5W]	100%	3.00E-51	ACB89011.1
PhageB1_ 040	Hypothetical protein	17602	17213	R	129	15.2	ATG	TAA	ORF114 [Staphylococcus virus G1]	100%	7.00E-90	YP_241063.1
PhageB1_ 041	Hypothetical protein	17874	17701	R	57	6.8	ATG	TAA	ORF245 [Staphylococcus virus G1]	100%	7.00E-35	YP_241064.1
PhageB1_ 042	Hypothetical protein	18397	17915	R	160	18.8	ATG	TGA	ORF090 [Staphylococcus virus G1]	100%	7.00E-111	YP_241065.1
PhageB1_ 043	Hypothetical protein	18989	18447	R	180	20.4	ATG	TAA	ORF072 [Staphylococcus virus G1]	100%	4.00E-124	YP_241066.1
PhageB1_ 044	Hypothetical protein	19522	18989	R	177	20.7	ATG	TAA	ORF077 [Staphylococcus virus G1]	100%	1.00E-124	YP_241067.1
PhageB1_	putative	19689	19525	R	54	6.3	ATG	TAA	ORF256	100%	7.00E-	YP_241068.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
045	membrane protein								[Staphylococcus virus G1]		30	
PhageB1_046	putative membrane protein	19967	19692	R	91	10.9	ATG	TAA	ORF163 [Staphylococcus virus G1]	100%	8.00E-54	YP_241069.1
PhageB1_047	Hypothetical protein	20812	19967	R	281	31.7	ATG	TAA	ORF038 [Staphylococcus virus G1]	100%	0.00E+00	YP_241070.1
PhageB1_048	putative AAA family ATPase	21942	20824	R	372	42.2	ATG	TAG	ORF024 [Staphylococcus virus G1]	100%	0.00E+00	YP_241071.1
PhageB1_049	Hypothetical protein	22422	22096	R	108	13	GTG	TAA	ORF134 [Staphylococcus virus G1]	99%	3.00E-73	YP_241072.1
PhageB1_050	Hypothetical protein	22831	22415	R	138	16	ATG	TAA	ORF106 [Staphylococcus virus G1]	100%	5.00E-96	YP_241073.1
PhageB1_051	putative nucleoside triphosphate pyrophosphohydrolase	23268	22966	R	100	11.3	ATG	TAA	ORF149 [Staphylococcus virus G1]	100%	7.00E-65	YP_241074.1
PhageB1_052	Hypothetical protein	23456	23268	R	62	7.3	ATG	TAA	ORF228 [Staphylococcus virus G1]	100%	2.00E-35	YP_241075.1
PhageB1_053	Hypothetical protein	23661	23500	R	53	6.4	ATG	TAA	ORF259 [Staphylococcus virus G1]	100%	6.00E-30	YP_241076.1
PhageB1_054	Hypothetical protein	25709	23661	R	682	79.8	ATG	TAA	ORF007 [Staphylococcus virus G1]	100%	0.00E+00	YP_241077.1
PhageB1_055	Hypothetical protein	26050	25787	R	87	10.1	ATG	TAA	ORF172 [Staphylococcus virus G1]	100%	7.00E-56	YP_241079.1
PhageB1_	Hypothetical	26240	26067	R	57	6.7	TTG	TAG	hypothetical protein	98%	7.00E-	YP_00711281

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
056	protein								[Staphylococcus phage JD007]		32	2.1
PhageB1_057	MbpB (putative membrane protein)	26825	26247	R	192	21.4	ATG	TAG	ORF068 [Staphylococcus virus G1]	100%	7.00E-131	YP_241080.1
PhageB1_058	putative nucleoside 2-deoxyribosyltransferase	27444	26818	R	208	23.8	ATG	TAA	ORF061 [Staphylococcus virus G1]	100%	2.00E-151	YP_241081.1
PhageB1_059	putative DNA ligase	28333	27437	R	298	35	ATG	TAA	ORF032 [Staphylococcus virus G1]	100%	0.00E+00	YP_241082.1
PhageB1_060	Hypothetical protein	28557	28333	R	74	8.2	ATG	TAA	hypothetical protein [Staphylococcus phage JD007]	100%	1.00E-40	YP_007112808.1
PhageB1_061	putative PhoH-related protein	29366	28626	R	246	28.6	ATG	TAA	ORF049 [Staphylococcus virus G1]	100%	0.00E+00	YP_241083.1
PhageB1_062	Hypothetical protein	30032	29418	R	204	23	ATG	TAG	ORF063 [Staphylococcus virus G1]	100%	4.00E-145	YP_241084.1
PhageB1_063	putative ribonuclease	30473	30048	R	141	15.8	ATG	TAA	ORF096 [Staphylococcus virus G1]	100%	2.00E-95	YP_241085.1
PhageB1_064	Hypothetical protein	30654	30463	R	63	7.5	ATG	TAG	ORF222 [Staphylococcus virus G1]	100%	5.00E-38	YP_241086.1
PhageB1_065	Hypothetical protein	31318	30677	R	213	24.6	ATG	TAA	ORF057 [Staphylococcus virus G1]	100%	1.00E-144	YP_241087.1
PhageB1_066	putatitive transcriptional regulator	31538	31308	R	76	8.8	ATG	TAA	ORF187 [Staphylococcus virus G1]	100%	2.00E-47	YP_241088.1
PhageB1_067	Hypothetical protein	31768	31541	R	75	9.2	ATG	TAA	ORF190 [Staphylococcus virus	100%	1.00E-45	YP_241089.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
									G1]			
PhageB1_ 068	putative transglycosylase	32570	31878	R	230	24.8	ATG	TAA	ORF054 [Staphylococcus virus G1]	100%	3.00E-168	YP_241090.1
PhageB1_ 069	Hypothetical protein	33392	32757	R	211	24.8	ATG	TAA	ORF058 Staphylococcus virus G1	100%	1.00E-152	YP_241091.1
PhageB1_ 070	Putative membrane protein	34250	33459	R	263	29.4	ATG	TAA	ORF044 [Staphylococcus virus G1]	100%	0.00E+00	YP_241091.2
PhageB1_ 071	Hypothetical protein	34558	34250	R	102	12.2	ATG	TAA	ORF146 [Staphylococcus virus G1]	100%	8.00E-66	YP_241093.1
PhageB1_ 072	putative endolysin	35300	34671	R	209	23.1	ATG	TAG	ORF060 [Staphylococcus virus G1]	100%	3.00E-155	YP_241094.1
PhageB1_ 073	Hypothetical protein	35371	35297	R	24	2.8	TTG	TGA	exodeoxyribonuclease VII large subunit [Enterococcus termitis]	37%	2.80E+00	WP_069661600.1
PhageB1_ 074	putative HNH endonuclease	36071	35571	R	166	19.3	ATG	TAA	ORF084 [Staphylococcus virus G1]	99%	2.00E-115	YP_241095.1
PhageB1_ 075	putative endolysin	37034	36231	R	267	29.8	ATG	TAA	ORF042 [Staphylococcus virus G1]	100%	0.00E+00	YP_241096.1
PhageB1_ 076	putative holin	37537	37034	R	167	18.1	ATG	TAA	ORF083 [Staphylococcus virus G1]	100%	5.00E-117	YP_241097.1
PhageB1_ 077	Hypothetical protein	37807	37622	R	61	7.1	ATG	TAA	ORF233 [Staphylococcus virus G1]	100%	2.00E-34	YP_241098.1
PhageB1_ 078	Hypothetical protein	39572	39354	R	72	8.7	ATG	TAA	ORF200 [Staphylococcus virus G1]	100%	7.00E-46	YP_241099.1

ORF	Predicted function	Start	stop	F/R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageB1_079	Hypothetical protein	40259	40050	R	69	8.0	ATG	TAA	ORF207 [Staphylococcus virus G1]	100%	3.00E-43	YP_241100.1
PhageB1_080	Hypothetical protein	40604	40272	R	110	12.5	TTG	TAG	ORF209 [Staphylococcus virus G1]	99%	2.00E-69	YP_007112791.1
PhageB1_081	putative membrane protein	40943	40617	R	108	13.1	TTG	TAG	hypothetical protein [Staphylococcus phage IME-SA2]	99%	7.00E-71	AKC02471.1
PhageB1_082	Hypothetical protein	41242	40976	R	88	10.1	ATG	TAA	ORF169 [Staphylococcus virus G1]	100%	6.00E-51	YP_241102.1
PhageB1_083	Putative membrane protein	41503	41769	F	88	10.3	ATG	TAA	ORF168 [Staphylococcus virus G1]	100%	7.00E-55	YP_241103.1
PhageB1_084	Hypothetical protein	41747	42025	F	92	10.6	ATG	TGA	ORF161 [Staphylococcus virus G1]	100%	2.00E-61	YP_241104.1
PhageB1_085	Hypothetical protein	42022	42432	F	136	15.6	TTG	TAA	ORF133 [Staphylococcus virus G1]	99%	5.00E-92	YP_241105.1
PhageB1_086	putative terminase large subunit	42447	42644	F	65	7.7	ATG	TAA	terminase large subunit [Staphylococcus phage Team1]	100%	4.00E-41	YP_009098219.1
PhageB1_087	Hypothetical protein	42938	43909	F	323	38.6	TTG	TAA	hypothetical protein [Staphylococcus phage Team1]	99%	0.00E+00	YP_009098220.1
PhageB1_088	putative terminase large subunit	44050	45597	F	515	59.7	ATG	TAG	Ter [Staphylococcus phage MSA6]	100%	0.00E+00	AFN38730.1
PhageB1_089	putative structural protein	45590	46411	F	273	30.7	ATG	TAG	hypothetical protein [Staphylococcus phage Team1]	100%	0.00E+00	YP_009098222.1
PhageB1_090	Hypothetical protein	46398	46571	F	57	6.7	GTG	TGA	ORF235 [Staphylococcus virus	100%	4.00E-30	YP_240894.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
									G1]			
PhageB1_ 091	Hypothetical protein	46568	47047	F	159	18.5	ATG	TAA	ORF091 [Staphylococcus virus G1]	100%	1.00E-110	YP_240895.1
PhageB1_ 092	putative membrane protein	47140	48270	F	376	41.2	ATG	TAA	membrane protein [Staphylococcus phage Team1]	99%	0.00E+00	YP_009098225.1
PhageB1_ 093	putative membrane protein	48346	48696	F	116	13.1	TTG	TAA	ORF120 [Staphylococcus virus G1]	99%	4.00E-74	YP_240898.1
PhageB1_ 094	putative portal protein	48714	49085	F	123	14.5	TTG	TAG	hypothetical protein [Staphylococcus phage phiIPLA-RODI]	100%	9.00E-84	YP_009195910.1
PhageB1_ 095	putative portal protein	49089	50780	F	563	64.1	TTG	TAG	ORF014 [Staphylococcus virus G1]	99%	0.00E+00	YP_240900.1
PhageB1_ 096	putitive prohead protease	50974	51747	F	257	28.6	TTG	TAG	ORF048 [Staphylococcus virus G1]	99%	0.00E+00	YP_240901.1
PhageB1_ 097	Hypothetical protein	51766	52722	F	318	35.9	ATG	TAA	ORF029 [Staphylococcus virus G1],	100%	0.00E+00	YP_240902.1
PhageB1_ 098	putative major capsid protein	52838	54229	F	463	51.2	ATG	TAA	ORF016 [Staphylococcus virus G1]	100%	0.00E+00	YP_240903.1
PhageB1_ 099	Hypothetical protein	54321	54617	F	98	11.3	ATG	TAA	ORF151 [Staphylococcus virus G1]	100%	1.00E-60	YP_240904.1
PhageB1_ 100	Hypothetical protein	54630	55538	F	302	34.2	ATG	TAA	ORF030 [Staphylococcus virus G1]	100%	0.00E+00	YP_240905.1
PhageB1_ 101	Putative capsid protein	55552	56430	F	292	33.7	ATG	TAA	ORF034 [Staphylococcus virus G1]	100%	0.00E+00	YP_240906.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageB1_ 102	Hypothetical protein	56430	57050	F	206	23.8	ATG	TAA	ORF062 [Staphylococcus virus G1]	100%	1.00E-149	YP_240907.1
PhageB1_ 103	Hypothetical protein	57069	57905	F	278	31.8	ATG	TAG	ORF039 [Staphylococcus virus G1]	100%	0.00E+00	YP_240908.1
PhageB1_ 104	Hypothetical protein	57907	58122	F	71	8.3	ATG	TAA	ORF202 [Staphylococcus virus G1]	100%	3.00E-46	YP_240909.1
PhageB1_ 105	putative tail sheath protein	58149	59912	F	587	64.5	ATG	TAG	putative tail sheath protein [Staphylococcus virus K]	99%	0.00E+00	YP_009041322.1
PhageB1_ 106	putative tail tube protein	59985	60413	F	142	15.9	ATG	TAA	ORF105 [Staphylococcus virus G1]	100%	8.00E-101	YP_240911.1
PhageB1_ 107	Hypothetical protein	60510	60650	F	46	5.4	ATG	TAA	ORF293 [Staphylococcus virus G1]	100%	1.00E-23	YP_240912.1
PhageB1_ 108	Hypothetical protein	60693	61151	F	152	18.1	ATG	TAA	ORF093 [Staphylococcus virus G1]	100%	1.00E-108	YP_240913.1
PhageB1_ 109	Putative membrane protein	61164	61358	F	64	7.2	ATG	TAG	ORF215 [Staphylococcus virus G1]	100%	5.00E-35	YP_240914.1
PhageB1_ 110	putative virion component	61440	61751	F	103	12.3	ATG	TAA	ORF141 [Staphylococcus virus G1]	100%	9.00E-67	YP_240915.1
PhageB1_ 111	Hypothetical protein	61883	62341	F	152	18.1	ATG	TAA	ORF095 [Staphylococcus virus G1]	100%	3.00E-106	YP_240916.1
PhageB1_ 112	putative tail morphogenetic protein	62385	62921	F	178	20.9	ATG	TAA	ORF074 [Staphylococcus virus G1]	100%	7.00E-127	YP_240917.1
PhageB1_	putative DNA	62977	67032	F	1351	143.8	ATG	TAG	DNA transfer protein	100%	0.00E+	AKQ07126.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
113	transfer protein								[Staphylococcus phage IME-SA118]		00	
PhageB1_ 114	putative secretory antigen SsaA-like protein	67111	69537	F	808	91.2	ATG	TAA	secretory antigen SsaA-like protein [Staphylococcus phage IME-SA119]	99%	0.00E+ 00	AKQ07397.1
PhageB1_ 115	putative peptidoglycan hydrolase, tail morphogenetic protein E	69551	70438	F	295	34.6	ATG	TAA	ORF033 [Staphylococcus virus G1]	100%	0.00E+ 00	YP_240922.1
PhageB1_ 116	putative phosphodiesterase	70438	72984	F	848	96.1	ATG	TAA	ORF004 [Staphylococcus virus G1]	100%	0.00E+ 00	YP_240923.1
PhageB1_ 117	Hypothetical protein	73091	73882	F	263	29.3	ATG	TAA	ORF043 [Staphylococcus virus G1]	100%	0.00E+ 00	YP_240924.1
PhageB1_ 118	Hypothetical protein	73882	74406	F	174	20	ATG	TAA	ORF078 [Staphylococcus virus G1]	100%	2.00E- 122	YP_240925.1
PhageB1_ 119	putative baseplate protein	74406	75110	F	234	26.6	ATG	TAG	ORF052 [Staphylococcus virus G1]	100%	2.00E- 172	YP_240926.1
PhageB1_ 120	putative baseplate J protein	75125	76171	F	348	39.2	ATG	TAA	ORF027 [Staphylococcus virus G1]	100%	0.00E+ 00	YP_240927.1
PhageB1_ 121	putative tail morphogenetic protein F	76192	79251	F	1019	116.4	GTG	TAA	conserved hypothetical protein [Staphylococcus phage Sb-1]	99%	0.00E+ 00	YP_00887361 8.1
PhageB1_ 122	putative structural protein	79362	79883	F	173	19.2	ATG	TAA	ORF079 [Staphylococcus virus G1]	100%	2.00E- 123	YP_240929.1
PhageB1_ 123	putative adsorption-	79904	83362	F	1152	129.1	ATG	TAA	ORF002 [Staphylococcus virus	100%	0.00E+ 00	YP_240930.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
	associated tail protein								G1]			
PhageB1_ 124	Hypothetical protein	83411	83569	F	52	6.2	ATG	TAG	ORF262 [Staphylococcus virus G1]	100%	2.00E-27	YP_240931.1
PhageB1_ 125	putative capsid and scaffold protein	83570	85492	F	640	72.6	ATG	TAA	capsid and scaffold protein [Staphylococcus phage IME-SA2]	100%	0.00E+00	AKC02517.1
PhageB1_ 126	Hypothetical protein	85440	85889	F	124	14.6	ATG	TAA	ORF117 [Staphylococcus virus G1]	100%	9.00E-85	YP_240933.1
PhageB1_ 127	putative structural protein	85896	87272	F	458	50.4	ATG	TAG	putative structural protein [Staphylococcus phage SA5]	99%	0.00E+00	AFV80704.1
PhageB1_ 128	putative DNA helicase	87364	89112	F	582	67.2	ATG	TAG	ORF012 [Staphylococcus virus G1]	100%	0.00E+00	YP_240935.1
PhageB1_ 129	putative Rep protein	89124	90737	F	537	63.2	ATG	TAA	ORF013 [Staphylococcus virus G1]	100%	0.00E+00	YP_240936.1
PhageB1_ 130	putative DNA helicase	90730	92172	F	480	54.6	ATG	TAA	ORF015 [Staphylococcus virus G1]	100%	0.00E+00	YP_240937.1
PhageB1_ 131	putative recombination exonuclease	92251	93288	F	315	40.1	ATG	TAA	ORF028 [Staphylococcus virus G1]	100%	0.00E+00	YP_240938.1
PhageB1_ 132	Hypothetical protein	93288	93665	F	125	14.9	ATG	TAA	ORF110 [Staphylococcus virus G1]	100%	2.00E-86	YP_240939.1
PhageB1_ 133	putative recombination related exonuclease	93665	95584	F	639	73.4	ATG	TAA	ORF009 [Staphylococcus virus G1]	100%	0.00E+00	YP_240940.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageB1_ 134	Hypothetical protein	95584	96180	F	198	23.2	ATG	TAG	hypothetical protein [Staphylococcus phage IME-SA1]	100%	4.00E-143	AKC02281.1
PhageB1_ 135	putative DNA primase	96195	97262	F	355	40.9	ATG	TAG	ORF026 [Staphylococcus virus G1]	100%	0.00E+00	YP_240942.1
PhageB1_ 136	Hypothetical protein	97329	97667	F	112	13.0	ATG	TAA	ORF127 [Staphylococcus virus G1]	100%	1.00E-72	YP_240943.1
PhageB1_ 137	Hypothetical protein	97667	98119	F	150	17.1	ATG	TAA	ORF098 [Staphylococcus virus G1]	100%	9.00E-100	YP_240944.1
PhageB1_ 138	putative resolvase	98106	98714	F	202	23.6	ATG	TAA	ORF064 [Staphylococcus virus G1]	100%	7.00E-149	YP_240945.1
PhageB1_ 139	putative ribonuclotide reductase stimulatory protein	98731	99123	F	143	16.2	ATG	TAA	ribonucleotide reductase stimulatory protein [Staphylococcus phage Team1]	100%	3.00E-98	YP_009098273.1
PhageB1_ 140	putative ribonucleotide reductase of class Ib (aerobic) alpha subunit	99138	101252	F	704	80.2	ATG	TAG	ribonucleotide reductase of class Ib (aerobic) alpha subunit [Staphylococcus phage IME-SA2]	99%	0.00E+00	AKC02533.1
PhageB1_ 141	putative ribonucleotide reductase minor subunit	101266	102315	F	349	40.4	ATG	TAA	putative ribonucleotide reductase minor subunit [Staphylococcus phage GH15]	100%	0.00E+00	YP_007002259.1
PhageB1_ 142	Hypothetical protein	102333	102662	F	109	12.4	ATG	TAG	ORF130 [Staphylococcus virus G1]	100%	2.00E-73	YP_240949.1
PhageB1_	putative	102646	102966	F	106	12.1	ATG	TAA	thioredoxin-like	100%	7.00E-	YP_00711294

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
143	thioredoxin-like protein								protein [Staphylococcus phage JD007]		70	9.1
PhageB1_144	Hypothetical protein	103173	103769	F	198	23.5	ATG	TAA	ORF066 [Staphylococcus virus G1]	99%	1.00E-140	YP_240951.1
PhageB1_145	putative integration host factor	103779	104084	F	101	11.9	ATG	TAA	ORF147 [Staphylococcus virus G1]	100%	5.00E-67	YP_240952.1
PhageB1_146	putative DNA polymerase	104160	105032	F	290	33.2	ATG	TGA	putative DNA polymerase A [Staphylococcus virus K]	100%	0.00E+00	YP_009041363.1
PhageB1_147	Hypothetical protein	105198	105710	F	170	20.3	GTG	TAA	ORF081 [Staphylococcus virus G1]	99%	7.00E-119	YP_240954.1
PhageB1_148	putative DNA polymerase-associated exonuclease	105846	107189	F	447	52.8	ATG	TAA	PolA [Staphylococcus phage MSA6]	99%	0.00E+00	AFN38789.1
PhageB1_149	putative HNH endonuclease	107457	108164	F	235	27.5	ATG	TAA	putative HNH endonuclease [Staphylococcus virus K]	100%	2.00E-170	YP_009041365.1
PhageB1_150	putative DNA polymerase	108398	109258	F	286	32.9	ATG	TAA	DNA polymerase [Staphylococcus phage phiIPLA-RODI]	100%	0.00E+00	YP_009195964.1
PhageB1_151	Hypothetical protein	109327	109569	F	80	9	GTG	TAA	ORF181 [Staphylococcus virus G1]	99%	5.00E-50	YP_240959.1
PhageB1_152	Hypothetical protein	109586	110068	F	160	18.9	ATG	TAA	ORF089 [Staphylococcus virus G1]	100%	8.00E-116	YP_240960.1
PhageB1_	Hypothetical	110155	111426	F	423	46.9	ATG	TAA	ORF020	100%	0.00E+	YP_240961.1

ORF	Predicted function	Start	stop	F/R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
153	protein								[Staphylococcus virus G1]		00	
PhageB1_154	putative DNA repair protein	111486	111710	F	74	7.9	ATG	TAG	recombinase a [Staphylococcus phage Team1]	100%	8.00E-45	YP_009098288.1
PhageB1_155	putative endonuclease	112055	113023	F	322	38.3	ATG	TAA	endonuclease [Staphylococcus phage Team1]	100%	0.00E+00	YP_009098289.1
PhageB1_156	putative DNA repair protein	113171	114118	F	315	35.7	ATG	TAA	ORF021 [Staphylococcus virus G1]	100%	0.00E+00	YP_240962.1
PhageB1_157	Hypothetical protein	114122	114475	F	117	13.4	ATG	TAA	ORF121 [Staphylococcus virus G1]	100%	5.00E-80	YP_240963.1
PhageB1_158	putative RNA polymerase sigma factor	114462	115124	F	220	26.6	ATG	TAG	ORF056 [Staphylococcus virus G1]	100%	5.00E-157	YP_240964.1
PhageB1_159	putative Ig-like protein	115252	115884	F	210	23.2	ATG	TAA	hypothetical protein [Staphylococcus phage IME-SA1]	99%	2.00E-146	AKC02307.1
PhageB1_160	putative major tail ptotein	115907	116419	F	173	18.2	ATG	TAG	putative major tail protein [Staphylococcus phage SA5]	100%	6.00E-117	AFV80732.1
PhageB1_161	putative major tail protein	116434	116661	F	75	7.8	ATG	TAA	ORF189 [Staphylococcus virus G1]	100%	1.00E-45	YP_240967.1
PhageB1_162	Hypothetical protein	116757	117017	F	86	10.3	ATG	TAA	ORF174 [Staphylococcus virus G1]	100%	1.00E-55	YP_240968.1
PhageB1_163	Hypothetical protein	117021	117776	F	251	29.2	ATG	TAA	ORF046 [Staphylococcus virus G1]	100%	0.00E+00	YP_240969.1
PhageB1_164	putative DNA repair exonuclease	117769	119019	F	416	47.6	ATG	TAA	ORF022 [Staphylococcus virus	100%	0.00E+00	YP_240970.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
									G1]			
PhageB1_ 165	Hypothetical protein	119033	119401	F	122	14	ATG	TGA	ORF118 [Staphylococcus virus G1]	100%	2.00E-81	YP_240971.1
PhageB1_ 166	Hypothetical protein	119388	119699	F	103	12	ATG	TAG	ORF143 [Staphylococcus virus G1]	100%	7.00E-69	YP_240972.1
PhageB1_ 167	Hypothetical protein	119763	120299	F	178	20.8	ATG	TAA	ORF075 [Staphylococcus virus G1]	100%	4.00E-128	YP_240973.1
PhageB1_ 168	Hypothetical protein	120292	121059	F	255	30.1	ATG	TAG	ORF045 [Staphylococcus virus G1]	100%	0.00E+00	YP_240974.1
PhageB1_ 169	Hypothetical protein	121037	121483	F	148	17.3	ATG	TAA	ORF099 [Staphylococcus virus G1]	100%	2.00E-104	YP_240975.1
PhageB1_ 170	Hypothetical protein	121483	122346	F	287	32.4	ATG	TAG	ORF036 [Staphylococcus virus G1]	100%	0.00E+00	YP_240976.1
PhageB1_ 171	Hypothetical protein	122718	123449	F	243	28.4	ATG	TAG	ORF047 [Staphylococcus virus G1]	100%	5.00E-174	YP_240977.1
PhageB1_ 172	Hypothetical protein	123467	123925	F	152	17.8	ATG	TAG	ORF094 [Staphylococcus virus G1]	100%	5.00E-106	YP_240978.1
PhageB1_ 173	Hypothetical protein	123990	124433	F	147	17.5	ATG	TAA	ORF100 [Staphylococcus virus G1]	100%	2.00E-99	YP_240979.1
PhageB1_ 174	Hypothetical protein	124450	125154	F	234	27.4	ATG	TAA	ORF053 [Staphylococcus virus G1]	100%	2.00E-169	YP_240980.1
PhageB1_ 175	Putative membrane protein	125216	125614	F	132	15.4	ATG	TAA	ORF108 [Staphylococcus virus G1]	100%	4.00E-91	YP_240981.1
PhageB1_	Hypothetical	125761	126003	F	80	9.4	ATG	TAG	ORF182	100%	4.00E-	YP_240982.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
176	protein								[Staphylococcus virus G1]		49	
PhageB1_177	Putative membrane protein	126008	126172	F	54	6.3	ATG	TGA	ORF252 [Staphylococcus virus G1]	100%	5.00E-30	YP_240983.1
PhageB1_178	Hypothetical protein	126159	126338	F	59	7.1	TTG	TAA	hypothetical protein [Staphylococcus phage Team1]	98%	1.00E-33	YP_009098312.1
PhageB1_179	Hypothetical protein	126374	126550	F	58	7	ATG	TAA	ORF240 [Staphylococcus virus G1]	100%	3.00E-33	YP_240984.1
PhageB1_180	putative membrane protein	126540	127073	F	177	20.9	ATG	TAA	ORF076 [Staphylococcus virus G1]	100%	6.00E-124	YP_240985.1
PhageB1_181	Hypothetical protein	127088	127336	F	82	9.1	ATG	TAA	hypothetical protein [Staphylococcus phage S25-4]	100%	1.00E-45	YP_008854124.1
PhageB1_182	Hypothetical protein	127348	127524	F	58	7	ATG	TAA	ORF241 [Staphylococcus virus G1]	100%	1.00E-31	YP_240986.1
PhageB1_183	Hypothetical protein	127517	127813	F	98	11.3	ATG	TAA	ORF152 [Staphylococcus virus G1]	100%	3.00E-64	YP_240987.1
PhageB1_184	putative membrane protein	127861	128043	F	60	7.2	ATG	TAG	membrane protein [Staphylococcus phage Team1]	98%	6.00E-33	YP_009098318.1
PhageB1_185	Hypothetical protein	128056	128424	F	122	14.2	ATG	TAA	ORF119 [Staphylococcus virus G1]	100%	1.00E-82	YP_240989.1
PhageB1_186	Hypothetical protein	128437	128784	F	115	13	ATG	TAA	ORF124 [Staphylococcus virus G1]	100%	2.00E-77	YP_240990.1
PhageB1_187	putative membrane protein	128784	129062	F	92	10.2	ATG	TAA	putative membrane protein MbpI [Staphylococcus	100%	6.00E-57	AFN38827.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
									phage MSA6]			
PhageB1_ 188	Hypothetical protein	129132	129437	F	101	12.1	ATG	TAG	ORF140 [Staphylococcus virus G1]	100%	8.00E-68	YP_240992.1
PhageB1_ 189	Hypothetical protein	129452	129802	F	116	13.7	ATG	TAA	ORF122 [Staphylococcus virus G1]	100%	4.00E-77	YP_240993.1
PhageB1_ 190	Hypothetical protein	129802	130404	F	200	23.4	ATG	TAA	ORF065 [Staphylococcus virus G1]	100%	1.00E-145	YP_240994.1
PhageB1_ 191	Hypothetical protein	130418	130597	F	59	7.3	ATG	TAA	ORF237 [Staphylococcus virus G1]	100%	4.00E-35	YP_240995.1
PhageB1_ 192	Hypothetical protein	130601	130669	F	22	2.6	ATG	TAA	hypothetical protein 812_188 [Staphylococcus phage 812]	100%	3.00E-15	YP_009224598.1
PhageB1_ 193	Hypothetical protein	130733	130807	F	24	2.8	ATG	TAG	hypothetical protein 812_189 [Staphylococcus phage 812]	100%	8.00E-20	YP_009224599.1
PhageB1_ 194	putative membrane protein	130824	131225	F	133	15	ATG	TAA	ORF107 [Staphylococcus virus G1]	100%	2.00E-87	YP_240996.1
PhageB1_ 195	Hypothetical protein	131227	131487	F	86	10.1	ATG	TGA	ORF173 [Staphylococcus virus G1]	100%	6.00E-54	YP_240997.1
PhageB1_ 196	putative membrane protein	131539	131826	F	95	10.5	ATG	TAG	ORF157 [Staphylococcus virus G1]	100%	6.00E-60	YP_240999.1
PhageB1_ 197	Hypothetical protein	131837	131953	F	38	4.6	ATG	TAG	ORF362 [Staphylococcus virus G1]	100%	3.00E-15	YP_241000.1
PhageB1_ 198	Hypothetical protein	131943	132206	F	87	9.9	ATG	TAA	ORF170 [Staphylococcus virus	100%	1.00E-53	YP_241001.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
									G1]			
PhageB1_ 199	Hypothetical protein	132283	132462	F	59	6.4	ATG	TAA	ORF236 [Staphylococcus virus G1]	100%	6.00E-30	YP_241002.1
PhageB1_ 200	Hypothetical protein	132477	132740	F	87	10.3	ATG	TAA	ORF171 [Staphylococcus virus G1]	100%	2.00E-56	YP_241003.1
PhageB1_ 201	Hypothetical protein	132743	133060	F	105	12	ATG	TAA	ORF137 [Staphylococcus virus G1]	100%	8.00E-69	YP_241004.1
PhageB1_ 202	Hypothetical protein	133061	133159	F	32	3.5	GTG	TGA	hypothetical protein [Staphylococcus phage Team1]	97%	1.00E-12	YP_009098334.1
PhageB1_ 203	Hypothetical protein	133427	133741	F	104	11.6	ATG	TAA	hypothetical protein [Staphylococcus phage Team1]	100%	2.00E-68	YP_009098335.1
PhageB1_ 204	putative membrane protein	133830	133988	F	52	5.7	ATG	TAA	ORF263 [Staphylococcus virus G1]	100%	9.00E-24	YP_241007.1
PhageB1_ 205	Hypothetical protein	134023	134223	F	66	7.6	ATG	TAA	ORF211 [Staphylococcus virus G1]	100%	7.00E-42	YP_241008.1
PhageB1_ 206	putative membrane protein	134224	134514	F	96	11.1	ATG	TAA	ORF155 [Staphylococcus virus G1]	100%	6.00E-60	YP_241009.1
PhageB1_ 207	Hypothetical protein	134606	134914	F	102	12.0	ATG	TGA	ORF144 [Staphylococcus virus G1]	100%	3.00E-64	YP_241010.1
PhageB1_ 208	putative robose-phosphate pyrophosphokinase	134911	135819	F	302	35.2	ATG	TAA	ORF031 [Staphylococcus virus G1]	100%	0.00E+00	YP_241011.1
PhageB1_ 209	putative nicotinate phosphoribosyltransferase	135837	137306	F	489	56.1	ATG	TAA	putative nicotinate phosphoribosyltransferase [Staphylococcus	100%	0.00E+00	YP_009041424.1

ORF	Predicted function	Start	stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop Codon	Best match (Blastp results)	% Identity	E-value	Accession number
									virus K]			
PhageB1_ 210	Hypothetical protein	137385	137630	F	81	10	ATG	TAA	ORF178 [Staphylococcus virus G1]	100%	3.00E-52	YP_241013.1
PhageB1_ 211	Hypothetical protein	137650	138042	F	130	15.4	ATG	TAG	ORF113 [Staphylococcus virus G1]	100%	2.00E-87	YP_241014.1
PhageB1_ 212	Hypothetical protein	138044	138265	F	73	8.9	ATG	TAA	ORF194 [Staphylococcus virus G1]	100%	6.00E-45	YP_241015.1
PhageB1_ 213	Hypothetical protein	138331	138642	F	103	11.6	ATG	TAA	ORF142 [Staphylococcus virus G1]	100%	1.00E-66	YP_241016.1
PhageB1_ 214	Hypothetical protein	138645	139154	F	169	20.3	ATG	TAA	ORF082 [Staphylococcus virus G1]	100%	4.00E-119	YP_241017.1
PhageB1_ 215	Hypothetical protein	139156	139485	F	109	12.6	ATG	TAA	ORF131 [Staphylococcus virus G1]	100%	3.00E-74	YP_241018.1
PhageB1_ 216	Hypothetical protein	139491	139685	F	64	7.8	ATG	TAA	gpORF179 [Staphylococcus phage A5W]	100%	2.00E-37	ACB89172.1
PhageB1_ 217	Hypothetical protein	139709	140023	F	104	12	ATG	TAA	ORF139 [Staphylococcus virus G1]	100%	2.00E-67	YP_241019.1
PhageB1_ 218	Hypothetical protein	140038	140205	F	55	6.5	ATG	TAA	ORF225 [Staphylococcus virus G1]	100%	1.00E-30	YP_241020.1
PhageB1_ 219	Hypothetical protein	140242	140343	F	33	3.7	ATG	TAA	ORF445 [Staphylococcus virus G1]	100%	7.00E-14	YP_241021.1

Table S2: Annotation of the staphylococcal phage vB_SauM_JA1 genome.

ORF	Predicted function	Start	Stop	F/R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_001	putative terminal repeat encoded protein	486	785	F	99	11.6	ATG	TAA	ORF150 [Staphylococcus virus G1]	100%	5.00E-65	YP_241022.1
PhageJA1_002	putative terminal repeat encoded protein	801	986	F	61	6.8	ATG	TAG	TreB [Staphylococcus phage A5W]	100%	9.00E-31	AHL83358.1
PhageJA1_003	putative terminal repeat encoded protein	1093	1383	F	96	11.3	ATG	TAA	ORF156 [Staphylococcus virus G1]	100%	2.00E-61	YP_241024.1
PhageJA1_004	putative terminal repeat encoded protein	1383	1670	F	95	10.9	ATG	TAA	ORF158 [Staphylococcus virus G1]	100%	2.00E-61	YP_241025.1
PhageJA1_005	putative terminal repeat encoded protein	1670	1963	F	97	11.5	ATG	TAA	ORF154 [Staphylococcus virus G1]	100%	3.00E-64	YP_241026.1
PhageJA1_006	putative terminal repeat encoded protein	1967	2224	F	85	10.2	ATG	TAG	ORF175 [Staphylococcus virus G1]	100%	5.00E-55	YP_241027.1
PhageJA1_007	putative terminal repeat encoded protein	2302	2541	F	79	9.2	ATG	TAG	ORF183 [Staphylococcus virus G1]	100%	2.00E-50	YP_241028.1
PhageJA1_008	putative terminal repeat encoded protein	2552	2899	F	115	13.7	ATG	TGA	ORF125 [Staphylococcus virus G1]	100%	2.00E-77	YP_241029.1
PhageJA1_009	putative terminal repeat encoded protein	3755	4063	F	102	11.8	ATG	TAA	ORF145 [Staphylococcus virus G1]	100%	1.00E-69	YP_241031.1
PhageJA1_010	putative terminal repeat encoded protein	4269	4553	F	94	11	ATG	TAA	ORF159 [Staphylococcus virus G1]	100%	8.00E-64	YP_241032.1
PhageJA1_011	putative terminal repeat encoded protein	4628	4819	F	63	7.7	ATG	TAA	ORF221 [Staphylococcus virus G1]	100%	1.00E-37	YP_241033.1

ORF	Predicted function	Start	Stop	F/R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_012	putative HNH homing endonuclease	5624	5136	R	162	19.6	ATG	TAA	ORF085 [Staphylococcus virus G1]	100%	5.00E-115	YP_241035.1
PhageJA1_013	putative terminal repeat encoded protein	5792	5950	F	52	6.1	ATG	TAA	hypothetical protein [Staphylococcus phage Team1]	100%	3.00E-27	YP_009098148.1
PhageJA1_014	putative terminal repeat encoded protein	6020	6151	F	43	5.2	ATG	TAA	ORF297 [Staphylococcus virus G1]	100%	1.00E-21	YP_241036.1
PhageJA1_015	Hypothetical protein	6319	6555	F	78	9.1	ATG	TAG	hypothetical protein CPT_phageK_gp017 [Staphylococcus virus K]	100%	1.00E-47	YP_009041239.1
PhageJA1_016	putative terminal repeat encoded protein	6635	7105	F	156	17.8	ATG	TAG	ORF092 [Staphylococcus virus G1]	100%	3.00E-105	YP_241038.1
PhageJA1_017	putative terminal repeat encoded protein	7135	7260	F	41	4.6	ATG	TAA	ORF166 [Staphylococcus virus G1]	100%	2.00E-20	YP_241041.1
PhageJA1_018	putative terminal repeat encoded protein	7345	7524	F	59	7.2	ATG	TAA	hypothetical protein [Staphylococcus phage ISP]	100%	4.00E-33	CCA65883.1
PhageJA1_019	Hypothetical protein	8094	7858	R	78	9.6	ATG	TAA	hypothetical protein SA5_0153/152 [Staphylococcus phage SA5]	100%	4.00E-50	AFV80807.1
PhageJA1_020	Hypothetical protein	8581	8096	R	161	19.1	ATG	TAA	ORF088 [Staphylococcus virus G1]	100%	3.00E-111	YP_241045.1
PhageJA1_021	Hypothetical protein	9001	8594	R	135	16.5	ATG	TAA	ORF109 [Staphylococcus virus G1]	100%	3.00E-94	YP_241046.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_ 022	Hypothetical protein	9432	9001	R	143	17.3	ATG	TAA	ORF103 [Staphylococcus virus G1]	100%	5.00E-96	YP_241047.1
PhageJA1_ 023	Hypothetical protein	9626	9435	R	63	7.9	ATG	TAA	ORF224 [Staphylococcus virus G1]	100%	1.00E-36	YP_241048.1
PhageJA1_ 024	putative membrane protein	10108	9623	R	161	18.3	ATG	TGA	conserved hypothetical protein [Staphylococcus phage Sb-1]	100%	1.00E-110	YP_008873528 .1
PhageJA1_ 025	Hypothetical protein	10532	10101	R	143	16.7	ATG	TAA	ORF104 [Staphylococcus virus G1]	100%	2.00E-98	YP_241050.1
PhageJA1_ 026	putative nucleotidyl transferase	11088	10546	R	180	21.5	ATG	TAA	ORF073 [Staphylococcus virus G1]	100%	4.00E-126	YP_241051.1
PhageJA1_ 027	Hypothetical protein	11588	11100	R	162	19.5	ATG	TAG	ORF086 [Staphylococcus virus G1]	100%	7.00E-118	YP_241052.1
PhageJA1_ 028	Hypothetical protein	11999	11601	R	132	16.1	ATG	TAA	ORF111 [Staphylococcus virus G1]	100%	1.00E-89	YP_241053.1
PhageJA1_ 029	putative serine threonine protein phosphatase I	12703	11996	R	235	27.7	ATG	TGA	serine threonine protein phosphatase I [Staphylococcus phage JD007]	100%	2.00E-172	YP_007112835 .1
PhageJA1_ 030	Hypothetical protein	13261	12803	R	152	17.6	ATG	TAA	ORF070 [Staphylococcus virus G1]	100%	2.00E-105	YP_241055.1
PhageJA1_ 031	putative tail protein	13690	13373	R	105	11.8	GTG	TAA	ORF138 [Staphylococcus virus G1]	99%	4.00E-67	YP_241056.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_ 032	Hypothetical protein	15224	14676	R	182	21.9	ATG	TGA	ORF071 [Staphylococcus virus G1]	99%	8.00E-125	YP_241057.1
PhageJA1_ 033	Hypothetical protein	15446	15228	R	72	8.4	ATG	TAA	ORF201 [Staphylococcus virus G1]	100%	4.00E-44	YP_241058.1
PhageJA1_ 034	Hypothetical protein	15641	15447	R	64	7.6	ATG	TAA	ORF218 [Staphylococcus virus G1]	100%	2.00E-38	YP_241059.1
PhageJA1_ 035	Hypothetical protein	16368	15631	R	245	28.7	ATG	TAA	ORF050 [Staphylococcus virus G1]	100%	3.00E-175	YP_241060.1
PhageJA1_ 036	Hypothetical protein	16535	16431	R	34	4.1	ATG	TAA	ORF437 [Staphylococcus virus G1]	100%	1.00E-14	YP_241061.1
PhageJA1_ 037	Hypothetical protein	16786	16547	R	79	9.4	ATG	TAA	gpORF020 [Staphylococcus phage A5W]	100%	3.00E-51	ACB89011.1
PhageJA1_ 038	Hypothetical protein	17177	16788	R	129	15.2	ATG	TAA	ORF114 [Staphylococcus virus G1]	100%	7.00E-90	YP_241063.1
PhageJA1_ 039	Hypothetical protein	17449	17276	R	57	6.8	ATG	TAA	ORF245 [Staphylococcus virus G1]	100%	7.00E-35	YP_241064.1
PhageJA1_ 040	Hypothetical protein	17972	17490	R	160	18.8	ATG	TGA	ORF090 [Staphylococcus virus G1]	100%	7.00E-111	YP_241065.1
PhageJA1_ 041	Hypothetical protein	18564	18022	R	180	20.4	ATG	TAA	ORF072 [Staphylococcus virus G1]	100%	4.00E-124	YP_241066.1
PhageJA1_ 042	Hypothetical protein	19097	18564	R	177	20.7	ATG	TAA	ORF077 [Staphylococcus virus G1]	100%	1.00E-124	YP_241067.1

ORF	Predicted function	Start	Stop	F/R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_043	putative membrane protein	19264	19100	R	54	6.3	ATG	TAA	ORF256 [Staphylococcus virus G1]	100%	7.00E-30	YP_241068.1
PhageJA1_044	putative membrane protein	19542	19267	R	91	10.9	ATG	TAA	ORF163 [Staphylococcus virus G1]	100%	8.00E-54	YP_241069.1
PhageJA1_045	Hypothetical protein	20387	19542	R	281	31.7	ATG	TAA	ORF038 [Staphylococcus virus G1]	100%	0.00E+00	YP_241070.1
PhageJA1_046	putative AAA family ATPase	21517	20399	R	372	42.2	ATG	TAG	ORF024 [Staphylococcus virus G1]	100%	0.00E+00	YP_241071.1
PhageJA1_047	Hypothetical protein	21997	21671	R	108	13	GTG	TAA	ORF134 [Staphylococcus virus G1]	99%	3.00E-73	YP_241072.1
PhageJA1_048	Hypothetical protein	22406	21990	R	138	16	ATG	TAA	ORF106 [Staphylococcus virus G1]	100%	5.00E-96	YP_241073.1
PhageJA1_049	putative nucleoside triphosphate pyrophosphohydrolase	22841	22539	R	100	11.3	ATG	TAA	ORF149 [Staphylococcus virus G1]	100%	7.00E-65	YP_241074.1
PhageJA1_050	Hypothetical protein	23029	22841	R	62	7.3	ATG	TAA	ORF228 [Staphylococcus virus G1]	100%	2.00E-35	YP_241075.1
PhageJA1_051	Hypothetical protein	23234	23073	R	53	6.4	ATG	TAA	ORF259 [Staphylococcus virus G1]	100%	6.00E-30	YP_241076.1
PhageJA1_052	Hypothetical protein	25282	23234	R	682	79.8	ATG	TAA	ORF007 [Staphylococcus virus G1]	100%	0.00E+00	YP_241077.1
PhageJA1_053	Hypothetical protein	25623	25360	R	87	10.1	ATG	TAA	ORF172 [Staphylococcus virus G1]	100%	7.00E-56	YP_241079.1

ORF	Predicted function	Start	Stop	F/R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_054	Hypothetical protein	25813	25640	R	57	6.7	TTG	TAG	hypothetical protein [Staphylococcus phage JD007]	98%	7.00E-32	YP_007112812.1
PhageJA1_055	MbpB (putative membrane protein)	26398	25820	R	192	21.4	ATG	TAG	ORF068 [Staphylococcus virus G1]	100%	7.00E-131	YP_241080.1
PhageJA1_056	putative nucleoside 2-deoxyribosyltransferase	27017	26391	R	208	23.8	ATG	TAA	ORF061 [Staphylococcus virus G1]	100%	2.00E-151	YP_241081.1
PhageJA1_057	putative DNA ligase	27906	27010	R	298	35	ATG	TAA	ORF032 [Staphylococcus virus G1]	100%	0.00E+00	YP_241082.1
PhageJA1_058	Hypothetical protein	28130	27906	R	74	8.2	ATG	TAA	hypothetical protein [Staphylococcus phage JD007]	100%	1.00E-40	YP_007112808.1
PhageJA1_059	putative PhoH-related protein	28939	28199	R	246	28.6	ATG	TAA	ORF049 [Staphylococcus virus G1]	100%	0.00E+00	YP_241083.1
PhageJA1_060	Hypothetical protein	29605	28991	R	204	23	ATG	TAG	ORF063 [Staphylococcus virus G1]	100%	4.00E-145	YP_241084.1
PhageJA1_061	putative ribonuclease	30046	29621	R	141	15.8	ATG	TAA	ORF096 [Staphylococcus virus G1]	100%	2.00E-95	YP_241085.1
PhageJA1_062	Hypothetical protein	30227	30036	R	63	7.5	ATG	TAG	ORF222 [Staphylococcus virus G1]	100%	5.00E-38	YP_241086.1
PhageJA1_063	Hypothetical protein	30891	30250	R	213	24.6	ATG	TAA	ORF057 [Staphylococcus virus G1]	100%	1.00E-144	YP_241087.1

ORF	Predicted function	Start	Stop	F/R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_064	putatitive transcriptional regulator	31111	30881	R	76	8.8	ATG	TAA	ORF187 [Staphylococcus virus G1]	100%	2.00E-47	YP_241088.1
PhageJA1_065	Hypothetical protein	31341	31114	R	75	9.2	ATG	TAA	ORF190 [Staphylococcus virus G1]	100%	1.00E-45	YP_241089.1
PhageJA1_066	putative transglycosylase	32143	31451	R	230	24.8	ATG	TAA	ORF054 [Staphylococcus virus G1]	100%	3.00E-168	YP_241090.1
PhageJA1_067	Hypothetical protein	32965	32330	R	211	24.8	ATG	TAA	ORF058 Staphylococcus virus G1	100%	1.00E-152	YP_241091.1
PhageJA1_068	Putative membrane protein	33823	33032	R	263	22.5	ATG	TAA	ORF044 [Staphylococcus virus G1]	99%	0.00E+00	YP_241091.2
PhageJA1_069	Hypothetical protein	34131	33823	R	102	12.2	ATG	TAA	ORF146 [Staphylococcus virus G1]	100%	8.00E-66	YP_241093.1
PhageJA1_070	putative endolysin	34873	34244	R	209	23.1	ATG	TAG	ORF060 [Staphylococcus virus G1]	100%	3.00E-155	YP_241094.1
PhageJA1_070A	Hypothetical protein	34944	34870	R	24	2.8	TTG	TGA	exodeoxyribonuclease VII large subunit [Enterococcus termitis]	37%	2.80E+00	WP_069661600.1
PhageJA1_071	putative HNH endonuclease	35644	35144	R	166	19.2	ATG	TAA	ORF084 [Staphylococcus virus G1]	100%	2.00E-117	YP_241095.1
PhageJA1_072	putative endolysin	36607	35804	R	267	29.8	ATG	TAA	ORF042 [Staphylococcus virus G1]	100%	0.00E+00	YP_241096.1
PhageJA1_073	putative holin	37110	36607	R	167	18.1	ATG	TAA	ORF083 [Staphylococcus virus G1]	100%	5.00E-117	YP_241097.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_074	Hypothetical protein	37380	37195	R	61	7.1	ATG	TAA	ORF233 [Staphylococcus virus G1]	100%	2.00E-34	YP_241098.1
PhageJA1_075	Hypothetical protein	39145	38927	R	72	8.7	ATG	TAA	ORF200 [Staphylococcus virus G1]	99%	4.00E-45	YP_241099.1
PhageJA1_076	Hypothetical protein	39832	39623	R	69	8.0	ATG	TAA	ORF207 [Staphylococcus virus G1]	100%	3.00E-43	YP_241100.1
PhageJA1_077	Hypothetical protein	40177	39845	R	110	12.5	TTG	TAG	ORF209 [Staphylococcus virus G1]	99%	2.00E-69	YP_007112791.1
PhageJA1_078	putative membrane protein	40516	40190	R	108	13.1	TTG	TAG	hypothetical protein [Staphylococcus phage IME-SA2]	98%	5.00E-70	AKC02471.1
PhageJA1_079	Hypothetical protein	40815	40549	R	88	10.1	ATG	TAA	ORF169 [Staphylococcus virus G1]	100%	6.00E-51	YP_241102.1
PhageJA1_080	Putative membrane protein	41094	41360	F	88	10.3	ATG	TAA	ORF168 [Staphylococcus virus G1]	100%	7.00E-55	YP_241103.1
PhageJA1_081	Hypothetical protein	41338	41616	F	92	10.6	ATG	TGA	ORF161 [Staphylococcus virus G1]	100%	2.00E-61	YP_241104.1
PhageJA1_082	Hypothetical protein	41613	42023	F	136	15.6	TTG	TAA	ORF133 [Staphylococcus virus G1]	99%	4.00E-92	YP_241105.1
PhageJA1_083	putative terminase large subunit	42038	42235	F	65	7.7	ATG	TAA	terminase large subunit [Staphylococcus phage Team1]	100%	4.00E-41	YP_009098219.1

ORF	Predicted function	Start	Stop	F/R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_084	Hypothetical protein	42529	43500	F	323	38.6	TTG	TAA	hypothetical protein [Staphylococcus phage Team1]	99%	0.00E+00	YP_009098220.1
PhageJA1_085	putative terminase large subunit	43641	45188	F	515	59.7	ATG	TAG	Ter [Staphylococcus phage MSA6]	100%	0.00E+00	AFN38730.1
PhageJA1_086	putative structural protein	45181	46002	F	273	30.7	ATG	TAG	hypothetical protein [Staphylococcus phage Team1]	100%	0.00E+00	YP_009098222.1
PhageJA1_087	Hypothetical protein	45989	46162	F	57	6.7	GTG	TGA	ORF235 [Staphylococcus virus G1]	100%	4.00E-30	YP_240894.1
PhageJA1_088	Hypothetical protein	46159	46638	F	159	18.5	ATG	TAA	ORF091 [Staphylococcus virus G1]	100%	1.00E-110	YP_240895.1
PhageJA1_089	putative membrane protein	46755	47837	F	360	39.5	ATG	TAA	membrane protein [Staphylococcus phage Team1]	99%	0.00E+00	YP_009098225.1
PhageJA1_090	putative membrane protein	47914	48264	F	116	13.1	TTG	TAA	ORF120 [Staphylococcus virus G1]	99%	4.00E-74	YP_240898.1
PhageJA1_091	putative portal protein	48282	48653	F	123	14.5	TTG	TAG	hypothetical protein [Staphylococcus phage phiIPLA-RODI]	100%	9.00E-84	YP_009195910.1
PhageJA1_092	putative portal protein	48657	50348	F	563	64.1	TTG	TAG	ORF014 [Staphylococcus virus G1]	99%	0.00E+00	YP_240900.1
PhageJA1_093	putitive prohead protease	50542	51315	F	257	28.6	TTG	TAG	ORF048 [Staphylococcus virus G1]	99%	0.00E+00	YP_240901.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_094	Hypothetical protein	51334	52290	F	318	35.9	ATG	TAA	ORF029 [Staphylococcus virus G1],	100%	0.00E+00	YP_240902.1
PhageJA1_095	putative major capsid protein	52406	53797	F	463	51.2	ATG	TAA	ORF016 [Staphylococcus virus G1]	100%	0.00E+00	YP_240903.1
PhageJA1_096	Hypothetical protein	53889	54185	F	98	11.3	ATG	TAA	ORF151 [Staphylococcus virus G1]	100%	1.00E-60	YP_240904.1
PhageJA1_097	Hypothetical protein	54198	55106	F	302	34.2	ATG	TAA	ORF030 [Staphylococcus virus G1]	100%	0.00E+00	YP_240905.1
PhageJA1_098	Putative capsid protein	55120	55998	F	292	33.7	ATG	TAA	ORF034 [Staphylococcus virus G1]	100%	0.00E+00	YP_240906.1
PhageJA1_099	Hypothetical protein	55998	56618	F	206	23.8	ATG	TAA	ORF062 [Staphylococcus virus G1]	100%	1.00E-149	YP_240907.1
PhageJA1_100	Hypothetical protein	56637	57473	F	278	31.8	ATG	TAG	ORF039 [Staphylococcus virus G1]	100%	0.00E+00	YP_240908.1
PhageJA1_101	Hypothetical protein	57475	57690	F	71	8.3	ATG	TAA	ORF202 [Staphylococcus virus G1]	100%	3.00E-46	YP_240909.1
PhageJA1_102	putative tail sheath protein	57717	59480	F	587	64.5	ATG	TAG	putative tail sheath protein [Staphylococcus virus K]	99%	0.00E+00	YP_009041322.1
PhageJA1_103	putative tail tube protein	59553	59981	F	142	15.9	ATG	TAA	ORF105 [Staphylococcus virus G1]	100%	8.00E-101	YP_240911.1
PhageJA1_104	Hypothetical protein	60078	60218	F	46	5.4	ATG	TAA	ORF293 [Staphylococcus virus G1]	100%	1.00E-23	YP_240912.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_105	Hypothetical protein	60261	60719	F	152	18.1	ATG	TAA	ORF093 [Staphylococcus virus G1]	100%	1.00E-108	YP_240913.1
PhageJA1_106	Putative membrane protein	60732	60926	F	64	7.2	ATG	TAG	ORF215 [Staphylococcus virus G1]	100%	5.00E-35	YP_240914.1
PhageJA1_107	putative virion component	61008	61319	F	103	12.3	ATG	TAA	ORF141 [Staphylococcus virus G1]	100%	9.00E-67	YP_240915.1
PhageJA1_108	Hypothetical protein	61451	61909	F	152	18.1	ATG	TAA	ORF095 [Staphylococcus virus G1]	100%	3.00E-106	YP_240916.1
PhageJA1_109	putative tail morphogenetic protein	61953	62489	F	178	20.9	ATG	TAA	ORF074 [Staphylococcus virus G1]	100%	7.00E-127	YP_240917.1
PhageJA1_110	putative DNA transfer protein	62545	66600	F	1351	143.8	ATG	TAG	DNA transfer protein [Staphylococcus phage IME-SA118]	100%	0.00E+00	AKQ07126.1
PhageJA1_111	putative secretory antigen SsaA-like protein	66679	69105	F	808	91.3	ATG	TAA	secretory antigen SsaA-like protein [Staphylococcus phage IME-SA119]	100%	0.00E+00	AKQ07397.1
PhageJA1_112	putative peptidoglycan hydrolase, tail morphogenetic protein E	69119	70006	F	295	34.6	ATG	TAA	ORF033 [Staphylococcus virus G1]	100%	0.00E+00	YP_240922.1
PhageJA1_113	putative phosphodiesterase	70006	72552	F	848	96.1	ATG	TAA	ORF004 [Staphylococcus virus G1]	100%	0.00E+00	YP_240923.1
PhageJA1_114	Hypothetical protein	72659	73450	F	263	29.3	ATG	TAA	ORF043 [Staphylococcus virus G1]	100%	0.00E+00	YP_240924.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_115	Hypothetical protein	73450	73974	F	174	20	ATG	TAA	ORF078 [Staphylococcus virus G1]	100%	2.00E-122	YP_240925.1
PhageJA1_116	putative baseplate protein	73974	74678	F	234	26.6	ATG	TAG	ORF052 [Staphylococcus virus G1]	100%	2.00E-172	YP_240926.1
PhageJA1_117	putative baseplate J protein	74693	75739	F	348	39.2	ATG	TAA	ORF027 [Staphylococcus virus G1]	99%	0.00E+00	YP_240927.1
PhageJA1_118	putative tail morphogenetic protein F	75760	78819	F	1019	116.4	GTG	TAA	conserved hypothetical protein [Staphylococcus phage Sb-1]	99%	0.00E+00	YP_008873618.1
PhageJA1_119	putative structural protein	78930	79451	F	173	19.2	ATG	TAA	ORF079 [Staphylococcus virus G1]	100%	2.00E-123	YP_240929.1
PhageJA1_120	putative adsorption-associated tail protein	79472	82930	F	1152	129.1	ATG	TAA	ORF002 [Staphylococcus virus G1]	100%	0.00E+00	YP_240930.1
PhageJA1_121	Hypothetical protein	82979	83137	F	52	6.2	ATG	TAG	ORF262 [Staphylococcus virus G1]	100%	2.00E-27	YP_240931.1
PhageJA1_122	putative capsid and scaffold protein	83138	85060	F	640	72.6	ATG	TAA	capsid and scaffold protein [Staphylococcus phage IME-SA2]	99%	0.00E+00	AKC02517.1
PhageJA1_123	Hypothetical protein	85083	85457	F	124	14.6	ATG	TAA	ORF117 [Staphylococcus virus G1]	100%	9.00E-85	YP_240933.1
PhageJA1_124	putative structural protein	85464	86840	F	458	50.4	ATG	TAG	putative structural protein [Staphylococcus phage SA5]	99%	0.00E+00	AFV80704.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_125	putative DNA helicase	86932	88680	F	582	67.2	ATG	TAG	ORF012 [Staphylococcus virus G1]	100%	0.00E+00	YP_240935.1
PhageJA1_126	putative Rep protein	88692	90305	F	537	63.2	ATG	TAA	ORF013 [Staphylococcus virus G1]	100%	0.00E+00	YP_240936.1
PhageJA1_127	putative DNA helicase	90298	91740	F	480	54.6	ATG	TAA	ORF015 [Staphylococcus virus G1]	100%	0.00E+00	YP_240937.1
PhageJA1_128	putative recombination exonuclease	91819	92856	F	315	40.1	ATG	TAA	ORF028 [Staphylococcus virus G1]	100%	0.00E+00	YP_240938.1
PhageJA1_129	Hypothetical protein	92856	93233	F	125	14.9	ATG	TAA	ORF110 [Staphylococcus virus G1]	100%	2.00E-86	YP_240939.1
PhageJA1_130	putative recombination related exonuclease	93233	95152	F	639	73.4	ATG	TAA	ORF009 [Staphylococcus virus G1]	99%	0.00E+00	YP_240940.1
PhageJA1_131	Hypothetical protein	95152	95748	F	198	23.2	ATG	TAG	hypothetical protein [Staphylococcus phage IME-SA1]	100%	2.00E-143	AKC02281.1
PhageJA1_132	putative DNA primase	95763	96830	F	355	40.9	ATG	TAG	ORF026 [Staphylococcus virus G1]	100%	0.00E+00	YP_240942.1
PhageJA1_133	Hypothetical protein	96897	97235	F	112	13	ATG	TAA	ORF127 [Staphylococcus virus G1]	100%	1.00E-72	YP_240943.1
PhageJA1_134	Hypothetical protein	97235	97686	F	150	17	ATG	TAA	ORF098 [Staphylococcus virus G1]	100%	1.00E-101	YP_240944.1
PhageJA1_135	putative resolvase	97674	98282	F	202	23.6	ATG	TAA	ORF064 [Staphylococcus virus G1]	100%	7.00E-149	YP_240945.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_136	putative ribonuclotide reductase stimulatory protein	98260	98691	F	143	16.2	ATG	TAA	ribonucleotide reductase stimulatory protein [Staphylococcus phage Team1]	100%	3.00E-98	YP_009098273.1
PhageJA1_137	putative ribonucleotide reductase of class Ib (aerobic) alpha subunit	98706	100820	F	704	80.1	ATG	TAG	ribonucleotide reductase of class Ib (aerobic) alpha subunit [Staphylococcus phage IME-SA2]	100%	0.00E+00	AKC02533.1
PhageJA1_138	putative ribonucleotide reductase minor subunit	100834	101883	F	349	40.4	ATG	TAA	putative ribonucleotide reductase minor subunit [Staphylococcus phage GH15]	100%	0.00E+00	YP_007002259.1
PhageJA1_139	Hypothetical protein	101901	102230	F	109	12.4	ATG	TAG	ORF130 [Staphylococcus virus G1]	100%	2.00E-73	YP_240949.1
PhageJA1_140	putative thioredoxin-like protein	102214	102534	F	106	12.1	ATG	TAA	thioredoxin-like protein [Staphylococcus phage JD007]	100%	7.00E-70	YP_007112949.1
PhageJA1_141	Hypothetical protein	102741	103337	F	198	23.5	ATG	TAA	ORF066 [Staphylococcus virus G1]	100%	1.00E-141	YP_240951.1
PhageJA1_142	putative integration host factor	103347	103652	F	101	11.9	ATG	TAA	ORF147 [Staphylococcus virus G1]	100%	5.00E-67	YP_240952.1
PhageJA1_143	putative DNA polymerase	103728	104600	F	290	33.2	ATG	TGA	putative DNA polymerase A [Staphylococcus virus K]	100%	0.00E+00	YP_009041363.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_144	Hypothetical protein	104766	105278	F	170	20.3	GTG	TAA	ORF081 [Staphylococcus virus G1]	99%	7.00E-119	YP_240954.1
PhageJA1_145	putative DNA polymerase-associated exonuclease	105414	106757	F	447	52.8	ATG	TAA	PolA [Staphylococcus phage MSA6]	100%	0.00E+00	AFN38789.1
PhageJA1_146	putative HNH endonuclease	107025	107732	F	235	27.5	ATG	TAA	putative HNH endonuclease [Staphylococcus virus K]	100%	2.00E-170	YP_009041365.1
PhageJA1_147	putative DNA polymerase	107966	108826	F	286	32.9	ATG	TAA	DNA polymerase [Staphylococcus phage phiIPLA-RODI]	100%	0.00E+00	YP_009195964.1
PhageJA1_148	Hypothetical protein	108895	109137	F	80	9	GTG	TAA	ORF181 [Staphylococcus virus G1]	99%	5.00E-50	YP_240959.1
PhageJA1_149	Hypothetical protein	109154	109636	F	160	18.9	ATG	TAA	ORF089 [Staphylococcus virus G1]	100%	8.00E-116	YP_240960.1
PhageJA1_150	Hypothetical protein	109723	110994	F	423	46.9	ATG	TAA	ORF020 [Staphylococcus virus G1]	100%	0.00E+00	YP_240961.1
PhageJA1_151	putative DNA repair protein	111054	111278	F	74	7.9	ATG	TAG	recombinase a [Staphylococcus phage Team1]	100%	8.00E-45	YP_009098288.1
PhageJA1_152	putative endonuclease	111623	112591	F	322	38.3	ATG	TAA	endonuclease [Staphylococcus phage Team1]	100%	0.00E+00	YP_009098289.1
PhageJA1_153	putative DNA repair protein	112739	113686	F	315	35.7	ATG	TAA	ORF021 [Staphylococcus virus G1]	100%	0.00E+00	YP_240962.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_ 154	Hypothetical protein	113690	114043	F	117	13.4	ATG	TAA	ORF121 [Staphylococcus virus G1]	100%	5.00E-80	YP_240963.1
PhageJA1_ 155	putative RNA polymerase sigma factor	114030	114692	F	220	26.6	ATG	TAG	ORF056 [Staphylococcus virus G1]	100%	5.00E-157	YP_240964.1
PhageJA1_ 156	putative Ig-like protein	114820	115443	F	207	23	ATG	TAA	hypothetical protein [Staphylococcus phage IME-SA1]	100%	1.00E-148	AKC02307.1
PhageJA1_ 157	putative major tail ptotein	115457	115978	F	173	18.2	ATG	TAG	putative major tail protein [Staphylococcus phage SA5]	100%	6.00E-117	AFV80732.1
PhageJA1_ 158	putative major tail protein	115993	116220	F	75	7.8	ATG	TAA	ORF189 [Staphylococcus virus G1]	100%	1.00E-45	YP_240967.1
PhageJA1_ 159	Hypothetical protein	116316	116576	F	86	10.3	ATG	TAA	ORF174 [Staphylococcus virus G1]	100%	1.00E-55	YP_240968.1
PhageJA1_ 160	Hypothetical protein	116580	117335	F	251	29.2	ATG	TAA	ORF046 [Staphylococcus virus G1]	100%	0.00E+00	YP_240969.1
PhageJA1_ 161	putative DNA repair exonuclease	117328	118578	F	416	47.6	ATG	TAA	ORF022 [Staphylococcus virus G1]	100%	0.00E+00	YP_240970.1
PhageJA1_ 162	Hypothetical protein	118592	118960	F	122	14	ATG	TGA	ORF118 [Staphylococcus virus G1]	100%	2.00E-81	YP_240971.1
PhageJA1_ 163	Hypothetical protein	118947	119258	F	103	12	ATG	TAG	ORF143 [Staphylococcus virus G1]	100%	7.00E-69	YP_240972.1
PhageJA1_ 164	Hypothetical protein	119322	119858	F	178	20.8	ATG	TAA	ORF075 [Staphylococcus virus G1]	100%	4.00E-128	YP_240973.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_ 165	Hypothetical protein	119851	120618	F	255	30.1	ATG	TAG	ORF045 [Staphylococcus virus G1]	100%	0.00E+ 00	YP_240974.1
PhageJA1_ 166	Hypothetical protein	120596	121042	F	148	17.3	ATG	TAA	ORF099 [Staphylococcus virus G1]	100%	2.00E- 104	YP_240975.1
PhageJA1_ 167	Hypothetical protein	121042	121905	F	287	32.4	ATG	TAG	ORF036 [Staphylococcus virus G1]	100%	0.00E+ 00	YP_240976.1
PhageJA1_ 168	Hypothetical protein	122277	123008	F	243	28.4	ATG	TAG	ORF047 [Staphylococcus virus G1]	100%	5.00E- 174	YP_240977.1
PhageJA1_ 169	Hypothetical protein	123026	123484	F	152	17.8	ATG	TAG	ORF094 [Staphylococcus virus G1]	100%	4.00E- 106	YP_240978.1
PhageJA1_ 170	Hypothetical protein	123549	123992	F	147	17.5	ATG	TAA	ORF100 [Staphylococcus virus G1]	100%	2.00E- 99	YP_240979.1
PhageJA1_ 171	Hypothetical protein	124009	124713	F	234	27.4	ATG	TAA	ORF053 [Staphylococcus virus G1]	100%	2.00E- 169	YP_240980.1
PhageJA1_ 172	Putative membrane protein	124775	125173	F	132	15.4	ATG	TAA	ORF108 [Staphylococcus virus G1]	100%	4.00E- 91	YP_240981.1
PhageJA1_ 173	Hypothetical protein	125320	125562	F	80	9.4	ATG	TAG	ORF182 [Staphylococcus virus G1]	100%	4.00E- 49	YP_240982.1
PhageJA1_ 174	Putative membrane protein	125567	125731	F	54	6.3	ATG	TGA	ORF252 [Staphylococcus virus G1]	100%	5.00E- 30	YP_240983.1
PhageJA1_ 175	Hypothetical protein	125718	125897	F	59	7.1	TTG	TAA	hypothetical protein [Staphylococcus phage Team1]	98%	1.00E- 33	YP_009098312 .1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_176	Hypothetical protein	125933	126109	F	58	7	ATG	TAA	ORF240 [Staphylococcus virus G1]	100%	3.00E-33	YP_240984.1
PhageJA1_177	putative membrane protein	126099	126632	F	177	20.9	ATG	TAA	ORF076 [Staphylococcus virus G1]	100%	6.00E-124	YP_240985.1
PhageJA1_178	Hypothetical protein	126647	126895	F	82	9.1	ATG	TAA	hypothetical protein [Staphylococcus phage S25-4]	100%	1.00E-45	YP_008854124.1
PhageJA1_179	Hypothetical protein	126907	127083	F	58	7	ATG	TAA	ORF241 [Staphylococcus virus G1]	100%	1.00E-31	YP_240986.1
PhageJA1_180	Hypothetical protein	127076	127372	F	98	11.3	ATG	TAA	ORF152 [Staphylococcus virus G1]	100%	3.00E-64	YP_240987.1
PhageJA1_181	putative membrane protein	127425	127607	F	60	7.2	ATG	TAG	membrane protein [Staphylococcus phage Team1]	100%	1.00E-33	YP_009098318.1
PhageJA1_182	Hypothetical protein	127620	127988	F	122	14.2	ATG	TAA	ORF119 [Staphylococcus virus G1]	100%	1.00E-82	YP_240989.1
PhageJA1_183	Hypothetical protein	128001	128348	F	115	13	ATG	TAA	ORF124 [Staphylococcus virus G1]	100%	2.00E-77	YP_240990.1
PhageJA1_184	putative membrane protein	128348	128626	F	92	10.2	ATG	TAA	putative membrane protein MbpI [Staphylococcus phage MSA6]	100%	6.00E-57	AFN38827.1
PhageJA1_185	Hypothetical protein	128696	129001	F	101	12.1	ATG	TAG	ORF140 [Staphylococcus virus G1]	100%	8.00E-68	YP_240992.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_186	Hypothetical protein	129016	129366	F	116	13.7	ATG	TAA	ORF122 [Staphylococcus virus G1]	100%	4.00E-77	YP_240993.1
PhageJA1_187	Hypothetical protein	129366	129968	F	200	23.4	ATG	TAA	ORF065 [Staphylococcus virus G1]	100%	1.00E-145	YP_240994.1
PhageJA1_188	Hypothetical protein	129982	130161	F	59	7.3	ATG	TAA	ORF237 [Staphylococcus virus G1]	100%	4.00E-35	YP_240995.1
PhageJA1_188A	Hypothetical protein	130165	130233	F	22	2.6	ATG	TAA	hypothetical protein 812_188 [Staphylococcus phage 812]	100%	3.00E-15	YP_009224598.1
PhageJA1_189	Hypothetical protein	130297	130371	F	24	2.8	ATG	TAG	hypothetical protein 812_189 [Staphylococcus phage 812]	100%	8.00E-20	YP_009224599.1
PhageJA1_190	putative membrane protein	130388	130789	F	133	15	ATG	TAA	ORF107 [Staphylococcus virus G1]	100%	2.00E-87	YP_240996.1
PhageJA1_191	Hypothetical protein	130791	131051	F	86	10.1	ATG	TGA	ORF173 [Staphylococcus virus G1]	100%	6.00E-54	YP_240997.1
PhageJA1_192	putative membrane protein	131103	131390	F	95	10.5	ATG	TAG	ORF157 [Staphylococcus virus G1]	100%	6.00E-60	YP_240999.1
PhageJA1_193	Hypothetical protein	131401	131517	F	38	4.6	ATG	TAG	ORF362 [Staphylococcus virus G1]	100%	3.00E-15	YP_241000.1
PhageJA1_194	Hypothetical protein	131507	131770	F	87	9.9	ATG	TAA	ORF170 [Staphylococcus virus G1]	100%	1.00E-53	YP_241001.1
PhageJA1_195	Hypothetical protein	131847	132026	F	59	6.4	ATG	TAA	ORF236 [Staphylococcus virus G1]	100%	6.00E-30	YP_241002.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_196	Hypothetical protein	132041	132304	F	87	10.3	ATG	TAA	ORF171 [Staphylococcus virus G1]	100%	2.00E-56	YP_241003.1
PhageJA1_197	Hypothetical protein	132307	132624	F	105	12	ATG	TAA	ORF137 [Staphylococcus virus G1]	100%	8.00E-69	YP_241004.1
PhageJA1_198	putative membrane protein	132795	132953	F	52	5.7	ATG	TAA	ORF263 [Staphylococcus virus G1]	100%	9.00E-24	YP_241007.1
PhageJA1_199	Hypothetical protein	132988	133188	F	66	7.6	ATG	TAA	ORF211 [Staphylococcus virus G1]	100%	7.00E-42	YP_241008.1
PhageJA1_200	putative membrane protein	133189	133479	F	96	11.1	ATG	TAA	ORF155 [Staphylococcus virus G1]	100%	6.00E-60	YP_241009.1
PhageJA1_201	Hypothetical protein	133571	133870	F	99	11.7	ATG	TAA	ORF144 [Staphylococcus virus G1]	100%	3.00E-62	YP_241010.1
PhageJA1_202	putative robose-phosphate pyrophosphokinase	133876	134784	F	302	35.2	ATG	TAA	ORF031 [Staphylococcus virus G1]	100%	0.00E+00	YP_241011.1
PhageJA1_203	putative nicotinate phosphoribosyltransferase	134802	136271	F	489	56.1	ATG	TAA	putative nicotinate phosphoribosyltransferase [Staphylococcus virus K]	100%	0.00E+00	YP_009041424.1
PhageJA1_204	Hypothetical protein	136350	136595	F	81	10	ATG	TAA	ORF178 [Staphylococcus virus G1]	100%	3.00E-52	YP_241013.1
PhageJA1_205	Hypothetical protein	136615	137007	F	130	15.4	ATG	TAG	ORF113 [Staphylococcus virus G1]	100%	2.00E-87	YP_241014.1

ORF	Predicted function	Start	Stop	F/ R	Size (aa)	Mw (kDa)	Start Codon	Stop codon	Best match (Blastp results)	% Identity	E-value	Accession number
PhageJA1_206	Hypothetical protein	137009	137230	F	73	8.9	ATG	TAA	ORF194 [Staphylococcus virus G1]	100%	6.00E-45	YP_241015.1
PhageJA1_207	Hypothetical protein	137296	137607	F	103	11.6	ATG	TAA	ORF142 [Staphylococcus virus G1]	100%	1.00E-66	YP_241016.1
PhageJA1_208	Hypothetical protein	137610	138119	F	169	20.3	ATG	TAA	ORF082 [Staphylococcus virus G1]	100%	4.00E-119	YP_241017.1
PhageJA1_209	Hypothetical protein	138121	138450	F	109	12.6	ATG	TAA	ORF131 [Staphylococcus virus G1]	100%	3.00E-74	YP_241018.1
PhageJA1_210	Hypothetical protein	138456	138650	F	64	7.8	ATG	TAA	gpORF179 [Staphylococcus phage A5W]	100%	2.00E-37	ACB89172.1
PhageJA1_211	Hypothetical protein	138674	138988	F	104	12	ATG	TAA	ORF139 [Staphylococcus virus G1]	100%	2.00E-67	YP_241019.1
PhageJA1_212	Hypothetical protein	139003	139170	F	55	6.5	ATG	TAA	ORF225 [Staphylococcus virus G1]	100%	1.00E-30	YP_241020.1
PhageJA1_213	Hypothetical protein	139207	139308	F	33	3.7	ATG	TAA	ORF445 [Staphylococcus virus G1]	100%	7.00E-14	YP_241021.1

Table S3. Predicted Rho-like promoters of *Staphylococcus* phage B1 found using MEME.

no.	Promoter	start	stop	-35	spacer	-10
1	B1P_13	5,687	5,659	TTGACA	TTAAGACCGAATTATTA	TATAAT
2	B1P_14	5,733	5,761	TTGACT	TTAATATCATTATAGTT	TAATAT
3	B1P_15	5,956	5,984	TTGACA	ACCTAGAAACAACATGT	TAATAT
4	B1P_16	6,253	6,281	TTGACA	GTCACTTGAAACCATGA	TATTAT
5	B1P_17	6,658	6,686	TTGACT	TTCAAGCCCTACAATGT	TATTAT
6	B1P_18	6,989	7,017	TTGACA	TCCTAACATATAGATGG	TAATAT
7	B1P_31	13,184	13,156	TTGACT	TTTTTTACTAAGTATGG	TAAGAT
8	B1P_37	16,852	16,824	TTGACA	TTATTATCAATATATGT	TATTAT
9	B1P_40	17,662	17,634	TTGACA	AAATATAAAAAATAGTG	TATAGT
10	B1P_41	17,937	17,909	ATGACT	TAGAAAAAGACCTATGA	TATATT
11	B1P_48	22,004	21,976	TTGACA	AATACAAATACTTGTAA	TATAAT
12	B1P_54	25,768	25,740	TTGACA	AATATTATTTACTATGG	TATGAT
13	B1P_52	23,499	23,471	TTGACA	ATAGTATCATAATATGA	TATAAT
14	B1P_60	28,616	28,588	TTGACA	AATCCCCTTAGTTATGG	TATAAT
15	B1P_64	30,702	30,674	TTGAGT	TAGTTATTAATTTAAAA	TAAAAT
16	B1P_67	31,829	31,801	TTGACT	TCATAAGTTAACTATGC	TATAAT
17	B1P_69	33,442	33,414	TTGACA	TAGGTGGTTTTTTATGC	TATAGT
18	B1P_68	32,658	32,630	TTGCGT	TATTTAAAGATATATGT	TATGAT
19	B1P_71	34,618	34,590	TTGACA	AAATTAAATACATAGTG	TATAGT
20	B1P_74	36,120	36,092	TTGACA	ACATAATAACTTTCCTA	TATACT
21	B1P_78	39,639	39,611	TTGACT	TATTTATCAATATAGTA	TATAGT
22	B1P_107	60,449	60,477	TTGACA	CTTTAAAATTTATATGT	TATTAT
23	B1P_108	60,634	60,662	TTGACA	ATTATAATTAACTAAGG	TATATT
24	B1P_109	61,137	61,165	TTGACA	ATTCAATAAGGAGGTAT	TATAAT
25	B1P_110	61,378	61,406	TTGACA	AATTAAAACTAATAAAT	TATAAT
26	B1P_114	67,043	67,071	TTGACA	CAAGAGTAGTATCATAA	TATACT
27	B1P_128	87,304	87,332	TTGACT	TGAAAAGGATTCTGTGG	TATACT
28	B1P_131	92,191	92,219	TTGACA	TTTTATATGTTAGGTGG	TATAAT
29	B1P_136	97,265	97,293	TTGACC	TTAGAGAAGTTTTATGT	TATACT
30	B1P_146	104,101	104,129	TTGACA	AGGTTTAAAATATATGG	TATAGT
31	B1P_151	109,267	109,295	TTGACA	ATATAGTTAACTTATGT	TATACT
32	B1P_153	110,090	110,118	TCATAA	ATATAAAAAACTATGT	TATAAT
33	B1P_162	116,697	116,725	TTGACA	ATTTATAATATCTATGA	TACACT
34	B1P_167	119,714	119,742	TTGACT	CTTTTTACTATATATGG	TATATT
35	B1P_173	123,928	123,956	TTGACA	GCTCCTATAGTTTATGA	TATAGT
36	B1P_176	125,707	125,735	TTGACT	CTCTTTTTGTTTTATGG	TATATT
37	B1P_179	126,322	126,350	TTGACA	AGAACAAATAAGTGTAG	TATAGT
38	B1P_184	127,793	127,821	TTGACA	GATGAAGCATTTTAATA	TATACT
39	B1P_188	129,072	129,100	TTGACA	CCTTTGTACTTTTGTAT	TATACT
40	B1P_193	130,673	130,701	TTGACA	ATTGAGTATACATAGGT	TATACT
41	B1P_199	132,225	132,253	TTGACA	TTAGGTTTCTTTTATTA	TATACT
42	B1P_204	133,772	133,800	TTGACA	GCAGGTATTTTTTATAG	TATACT
43	B1P_210	137,326	137,354	TTGACA	AAGGGAGTTTTTTATTA	TATAGT
44	B1P_213	138,281	138,309	TTGACT	TAGGTAGGTATCTATTA	TATAAT

Table S4. Predicted Rho-like promoters of *Staphylococcus* phage JA1 found using MEME.

no.	Promoter	start	stop	-35	spacer	-10
1	JA1P_12	5,657	5,685	TTGACA	TTAAGACCGAATTATTA	TATAAT
2	JA1P_13	5,731	5,759	TTGACT	TTAATATCATTATAGTT	TAATAT
3	JA1P_14	5,954	5,982	TTGACA	ACCTAGAAACAACATGT	TAATAT
4	JA1P_15	6,251	6,279	TTGACA	GTCACTTGAAACCATGA	TATTAT
5	JA1P_16	6,564	6,592	TTGACA	TCCTAACATATAGATGG	TAATAT
6	JA1P_29	12,759	12,731	TTGACT	TTTTTTACTAAGTATGG	TAAGAT
7	JA1P_35	16427	16,399	TTGACA	TTATTATCAATATATGT	TATTAT
8	JA1P_38	17,237	17,209	TTGACA	AAATATAAAAAATAGTG	TATAGT
9	JA1P_39	17,512	17,484	ATGACT	TAGAAAAAGACCTATGA	TATATT
10	JA1P_46	21,579	21,551	TTGACA	AATACAAATACTTGTAA	TATAAT
11	JA1P_50	23,072	23,044	TTGACA	ATAGTATCATAATATGA	TATAAT
12	PJA1_52	25,341	25,313	TTGACA	AATATTATTTACTATGG	TATGAT
13	JA1P_58	28,189	28,161	TTGACA	AATCACCTTAGTTATGG	TATAAT
14	JA1P_62	30,275	30,247	TTGAGT	TAGTTATTAATTTAAAA	TAAAAT
15	JA1P_65	31,402	31,374	TTGACT	TCATAAGTTAACTATGC	TATAAT
16	JA1P_66	32,231	32,203	TTGCGT	TATTTAAAGATATATGT	TATGAT
17	JA1P_69	34,191	34,163	TTGACA	AAATTAAATACATAGTG	TATAGT
18	JA1P_67	33,015	32,987	TTGACA	TAGGTGGTTTTTTATGC	TATAGT
19	JA1P_71	35,693	35,665	TTGACA	ACATAATAACTTTCCTA	TATACT
20	JA1P_75	39,212	39,184	TTGACT	TATTTATCAATATAGTA	TATAGT
21	JA1P_105	60,202	60,230	TTGACA	ATTATAATTAACTAAGG	TATATT
22	JA1P_198	132,737	132,765	TTGACA	GCAGGTATTTTTTATAG	TATACT
23	JA1P_204	136,291	136,319	TTGACA	AAGGGAGTTTTTTATTA	TATAGT
24	JA1P_104	60,017	60,045	TTGACA	CTTTAAAATTTATATGT	TATTAT
25	JA1P_106	60,705	60,733	TTGACA	ATTCAATAAGGAGGTAT	TATAAT
26	JA1P_107	60,946	60,974	TTGACA	AATTAAAACTAATAAAT	TATAAT
27	JA1P_111	66,611	66,639	TTGACA	CAAGAGTAGTATCATAA	TATACT
28	JA1P_125	86,872	86,900	TTGACT	TGAAAAGGATTCTGTGG	TATACT
29	JA1P_127	91,759	91,787	TTGACA	TTTTATATGTTAGGTGG	TATAAT
30	JA1P_133	96,833	96,861	TTGACC	TTAGAGAAGTTTTATGT	TATACT
31	JA1P_143	103,669	103,697	TTGACA	AGGTTTAAAATATATGG	TATAGT
32	JA1P_148	108,835	108,863	TTGACA	ATATAGTTAACTTATGT	TATACT
33	JA1P_150	109,658	109,686	TTGACA	AATATAAAAAACTATGT	TATAAT
34	JA1P_159	116,256	116,284	TTGACA	ATTTATAATATCTATGA	TACACT
35	JA1P_164	119,273	119,301	TTGACT	CTTTTTACTATATATGG	TATATT
36	JA1P_170	123,487	123,515	TTGACA	GCTCCTATAGTTTATGA	TATAGT
37	JA1P_173	125,266	125,294	TTGACT	CTCTTTTTGTTTTATGG	TATATT
38	JA1P_176	125,881	125,909	TTGACA	AGAACAAATAAGTGTAG	TATAGT
39	JA1P_181	127,352	127,380	TTGACA	GATGAAGCATTTTAATA	TATACT
40	JA1P_185	128,636	128,665	TTGACA	CCTTTGTACTTTTGTAT	TATACT
41	JA1P_189	130,237	130,265	TTGACA	ATTGAGTATACATAAGT	TATACT
42	JA1P_195	141,789	131,817	TTGACA	TTAGGTTTCTTTTATTA	TATACT
43	JA1P_207	137,246	137,274	TTGACT	TAGGTAGGTATCTATTA	TATAAT

Table S5. High ΔG rho-independent terminators predicted in the genome Staphylococcus phage B1 identified using ARNold and QuikFold.

no.	Terminator	Coordinates	Sequence	ΔG kcal/mol
1	B1T_9	3056-3093	ACACTAGGAATAATATCCTAGTGTaTTTATTTTTGCGG	-12.8
2	B1T_8	3071-3104	CACTAGGATATTATTCCTAGTGTATTATATAATT	-11.8
3	B1T_12	4828-4865	TCCCTAGAAATCTAATCCTAGGGAaTTGTATAATTTTT	-9.8
4	B1T_13	4828-4865	TCCCTAGGATTAGATTTCTAGGGATTTTTATTTATT	-13.1
5	B1T_13a	5137-5103	AGAAAAGGGTTGACCTTTTCTtTTTTCTATAGTAT	-9
6	B1T_20	7960-7989	GAGGGAATAAAATCCCTCTTTTATTTTTAT	-9.6
7	B1T_21	8280-8247	GGAGGGATTTAATTTCCCTCTTTTTTTATTTTAG	-10.4
8	B1T_41	17691-17657	AGGCTACTTTAATTAGTAGCCTTTTTTTGTTGACA	-11.5
9	B1T_43	18458-18425	GCAGACTTTTAATAAGTCTGCTTTTCTCTTATAT	-11.6
10	B1T_51	22889-22852	CACCTTGCTTGTAGCCAAGCAGGGTGTTTTTTTTTTAT	-16.9
11	B1T_68	31869-31834	GACTAAGATTAATTTCTTAGTCtTTTTTTGTATATT	-10.3
12	B1T_70	33449-33417	CCACCTATTGACATAGGTGGTTTTTTATGCTAT	-10.5
13	B1T_72	34659-34625	AGACGGATTTTAAATCCGTCTaTTTTTTTTGCAAA	-10.8
14	B1T_92	48262-48291	GAGGAGTAATTACTCCTCTTTTTTGTTTGC	-10.6
15	B1T_95	50787-50820	AGCCTAGAATAAATCTAGGCTTTGTTTATTTTTT	-11
16	B1T_98	54261-54296	TAGGGTACAGTAAAATGTACCCTATTTATATTCTTT	-12.8
17	B1T_106	60421-60452	GACCAACTAAAAAGTTGGTCTTTTTTTATTGA	-11.3
18	B1T_112	62923-62958	GGGTGGTAGGTGATACTACCATCCTTATTTTTTTAA	-15.4
19	B1T_116	62923-62958	AGACCTATTAATTTAGGTCTTTTTTTAGTTGTA	-8.7
20	B1T_123	83367-83398	GAGGGGTTGATTGACCCCTCTTTATTTAATAA	-14.2
21	B1T_127	87272-87305	GACTAGGAGAAATTTCCTAGTCTTTTTTTTTCTT	-12.3
22	B1T_140	101361-10139	TTGGGAGCAAGGAATCTCCCAATTTTGGACTCCT	-9.1
23	B1T_145	104073-104106	GAAGAGAAATAATTCTCTTCtTTTTTTATTGACA	-9.1
24	B1T_153	111422-111462	GAGTGCCTTAGAGCACTCTTTTATTTGAGA	-9
25	B1T_161	116669-116700	GACCAACTAAAAAGTTGGTCTTTTTTTATTGA	-11.3
26	B1T_166	119699-119732	GAGTCAAGTCTTTACTTGACTCTTTTTACTATAT	-12
27	B1T_175	125692-125725	GAGTCAAGTTAATTCTTGACTCTCTTTTTGTTTT	-11.5
28	B1T_181	127410-127446	GAAGGTAGAGAATAAGCTACCTTCTTCTACTCCTATT	-11.2
29	B1T_203	133755-133797	TACCTGTTGACAGCCTGTTGACAGCAGGTATTTTTTAT AGTAT	-14.1
30	B1T_209	137317-137354	AACTCCCTATTGACAAAGGGAGTTtTTTATTATATAGT	-10.8

Table S6. High ΔG rho-independent terminators predicted in the genome Staphylococcus phage JA1 identified using ARNold and QuikFold.

Terminator	Coordinates	Sequence	ΔG kcal/mol
JA1T_8	3069-3102	CACTAGGATATTATTCCTAGTGTATTATATAATT	-11.8
JA1T_11	4840-4875	TCCCTAGGATTAGATTTCTAGGGATTTTTATTTATT	-13.1
JA1T_12	5135-5101	AGAAAAGGGTTGACCTTTTCTtTTTTCTATAGTAT	-9
JA1T_18	7535-7564	GAGGGAATAAAATCCCTCTTTTATTTTTAT	-9.6
JA1T_19	7855-7822	GGAGGGATTTAATTTCCCTCTTTTTTTATTTTAG	-10.4
JA1T_30	12789-12754	ACACCTATTAATTTAATAGGTGTTTTTTTATTGACT	-9.9
JA1T_41	18033-18000	GCAGACTTTTAATAAGTCTGCTTTTCTCTTATAT	-11.6
JA1T_47	21654-21611	TACCTTACCCTATGTTAAGTTATAGGTGTAAGGTATTTTTTTTT	-17.4
JA1T_49	22462-22425	CACCTTGCTTGTAGCCAAGCAGGGTGTTTTTTTTATAT	-16.9
JA1T_51	25371-25338	GAAGGACTTTAAAAAGTTCTTCTTTTTTTGTTGA	-9.3
JA1T_66	31442-31407	GACTAAGATTAATTTCTTAGTCtTTTTTTGTATATT	-9.3
JA1T_68	33022-31990	CCACCTATTGACATAGGTGGTTTTTTATGCTAT	-10.5
JA1T_70	34232-32198	AGACGGATTTTAAATCCGTCTaTTTTTTTTGCAAA	-10.8
JA1T_89	47829-47858	GAGGAGTAATTACTCCTCTTTTTTTGTTTG	-10.6
JA1T_92	48657-50348	AGCCTAGAATAAATCTAGGCTTTGTTTATTTTTT	-11
JA1T_95	53829-53864	GGGATAAACTTAGGGTTTATCCCTTTTTTATTAAAA	-12.8
JA1T_103	53829-53864	GACCAACTAAAAAGTTGGTCTTTTTTTATTGA	-11.3
JA1T_109	62491-62526	GGGTGGTAGGTGATACTACCATCCTTATTTTTTTAA	-15.4
JA1T_113	72559-72591	AGACCTATTAATTTAGGTCTTTTTTTAGTTGTA	-8.7
JA1T_120	82935-82966	GAGGGGTTGATTGACCCCTCTTTATTTAATAA	-14.2
JA1T_124	86840-86873	GACTAGGAGAAATTTCCTAGTCTTTTTTTTTCTT	-12.3
JA1T_137	100929-100962	TTGGGAGCAAGGAATCTCCCAATTTTGGACTCCT	-9.1
JA1T_142	103641-103674	GAAGAGAAATAATTCTCTTCtTTTTTTATTGACA	-9.1
JA1T_150	111001-111030	GAGTGCCTTAGAGCACTCTTTTATTTGAGA	-9
JA1T_158	116228-116259	GACCAACTAAAAAGTTGGTCTTTTTTTATTGA	-11.3
JA1T_163	119258-119291	GAGTCAAGTCTTTACTTGACTCTTTTTACTATAT	-12
JA1T_172	125251-125284	GAGTCAAGTTAATTCTTGACTCTCTTTTTGTTTT	-11.5
JA1T_178	126969-127005	GAAGGTAGAGAATAAGCTACCTTCTTCTACTCCTATT	-11.2
JA1T_197	132720-132762	TACCTGTTGACAGCCTGTTGACAGCAGGTATTTTTTATAGTAT	-14.2
JA1T_203	136282-136319	AACTCCCTATTGACAAAGGGAGTTtTTTATTATATAGT	-10.8

Table S7: Percentage similarity based on BLASTN of broad host range *Staphylococcus* phages that form commercial phage cocktails to that of *Staphylococcus* phage K.

Phage	Accession	Identity vs phage K, %
Phage K	KF766114	100
Team 1	KC012913	95
fRuSau02	MF398190	95
Sb-1	HQ163896	85
ISP	FR852584	96

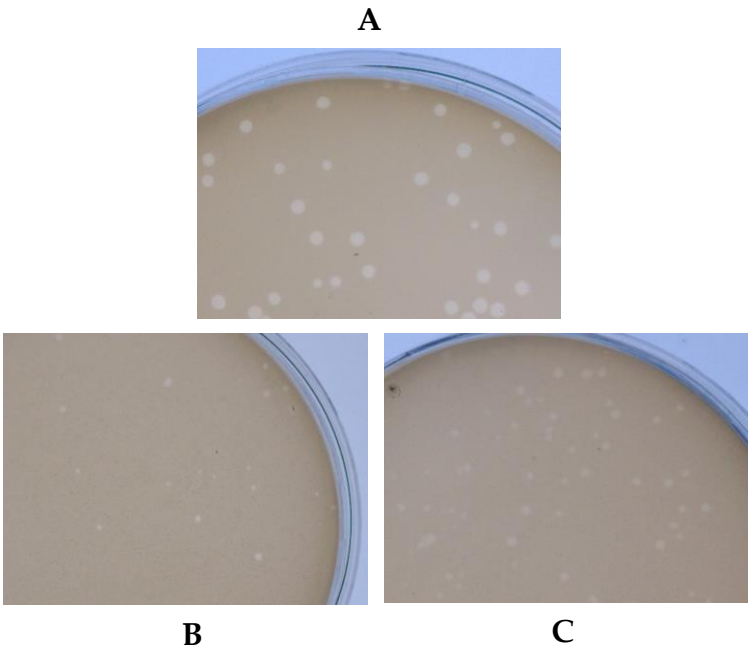


Figure S1. Plaque morphologies of phages B1, JA1 and K with common morphology types encountered in their host range study to include plaques sizes of 2mm (A), 0.5mm (B) and 1.0mm (C). Plaque morphologies with halos were encountered but were not clearly seen as photographs. Faint plaques were also encountered and these could only be clearly seen in direct path of light (C).

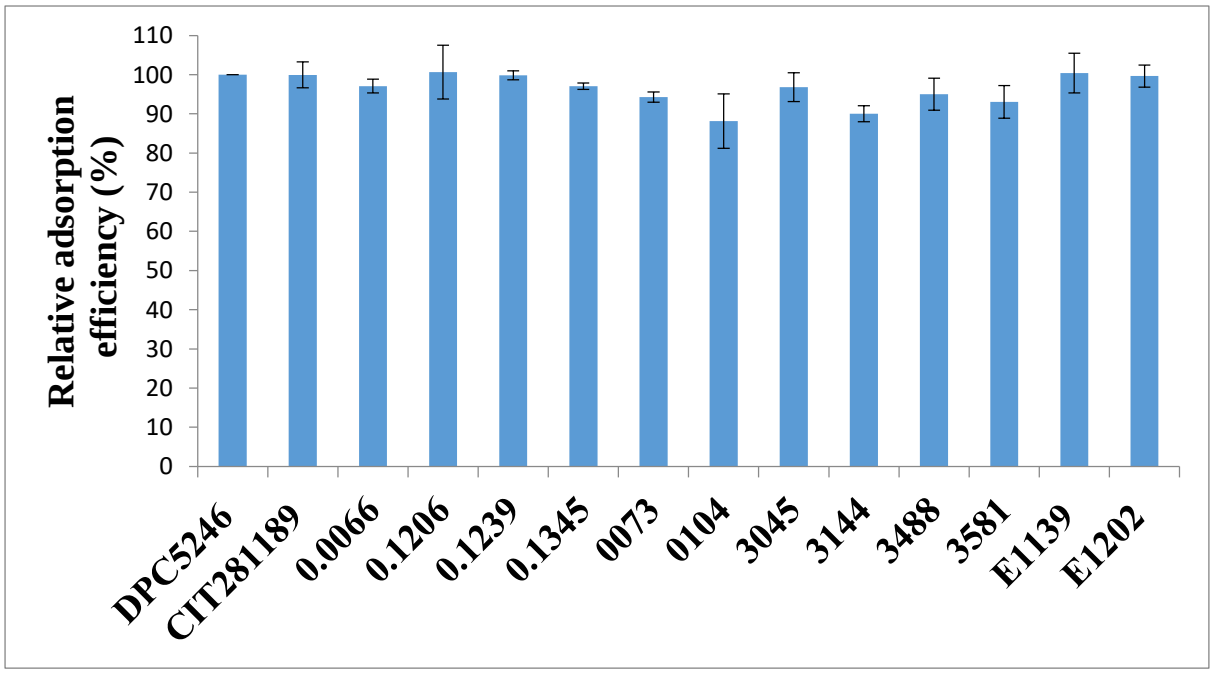


Figure S2. *Staphylococcus* phage K adsorption to strains of *Staphylococcus aureus* resistant to infection by, in comparison host strain DPC5246.

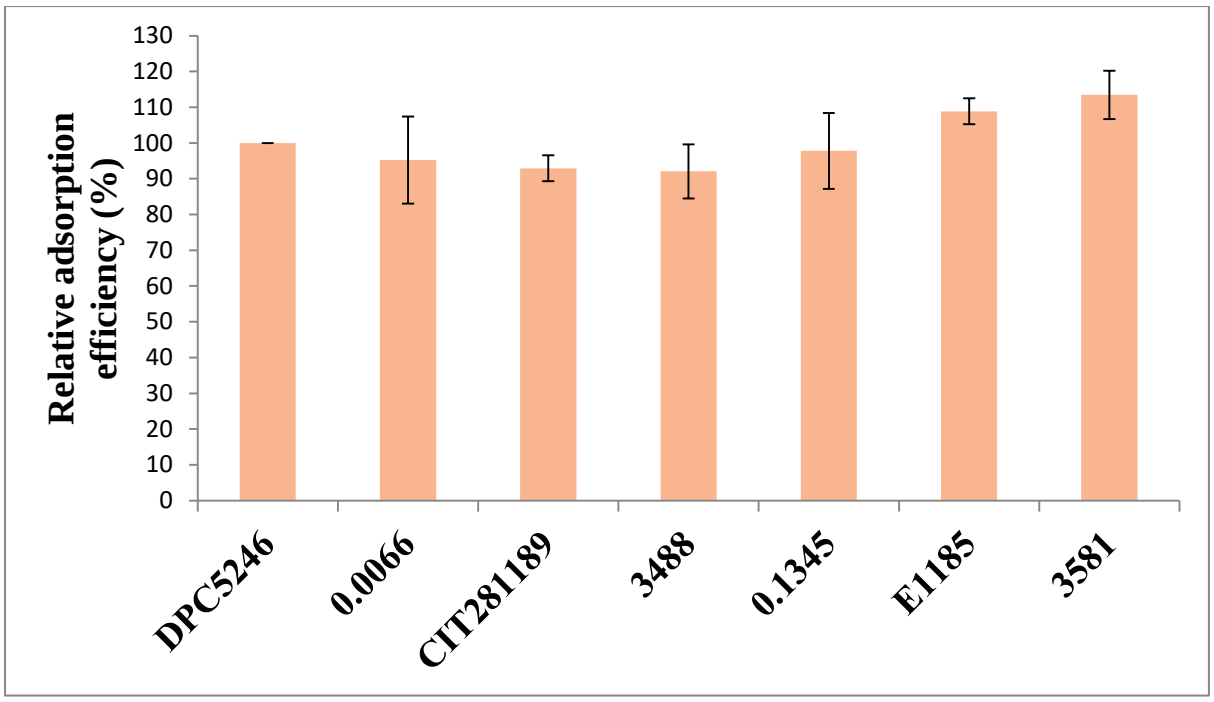


Figure S3. *Staphylococcus* phage B1 adsorption to strains of *Staphylococcus aureus* resistant to infection by, in comparison host strain DPC5246.

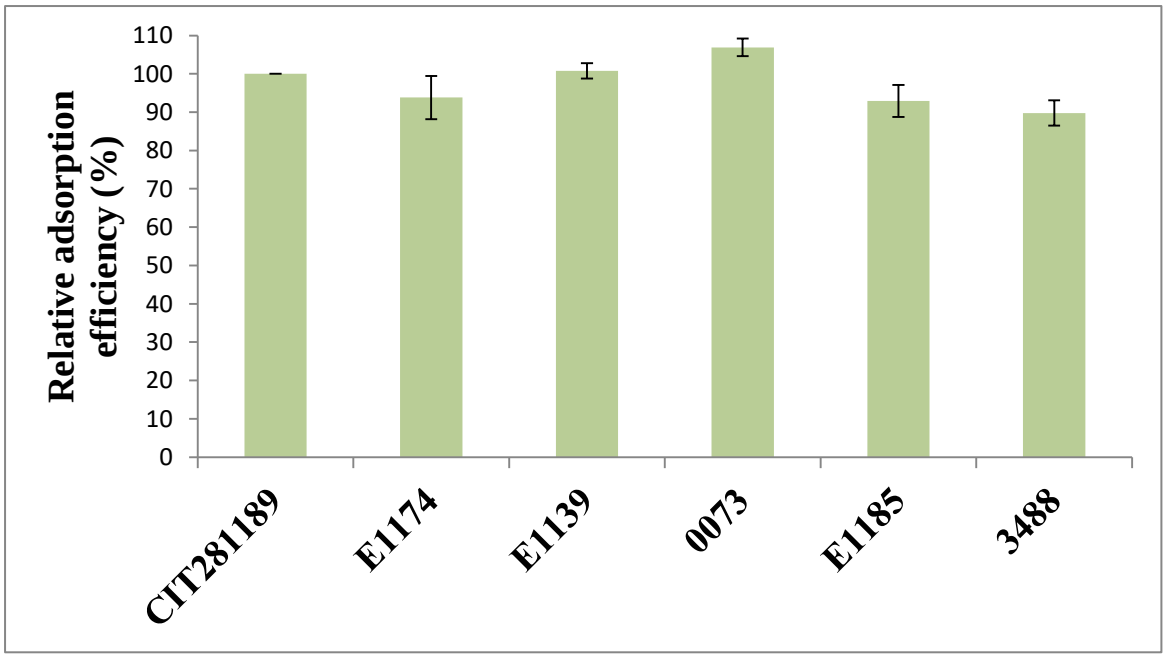


Figure S4. *Staphylococcus* phage JA1 adsorption to strains of *Staphylococcus aureus* resistant to infection by, in comparison host strain CIT281189.

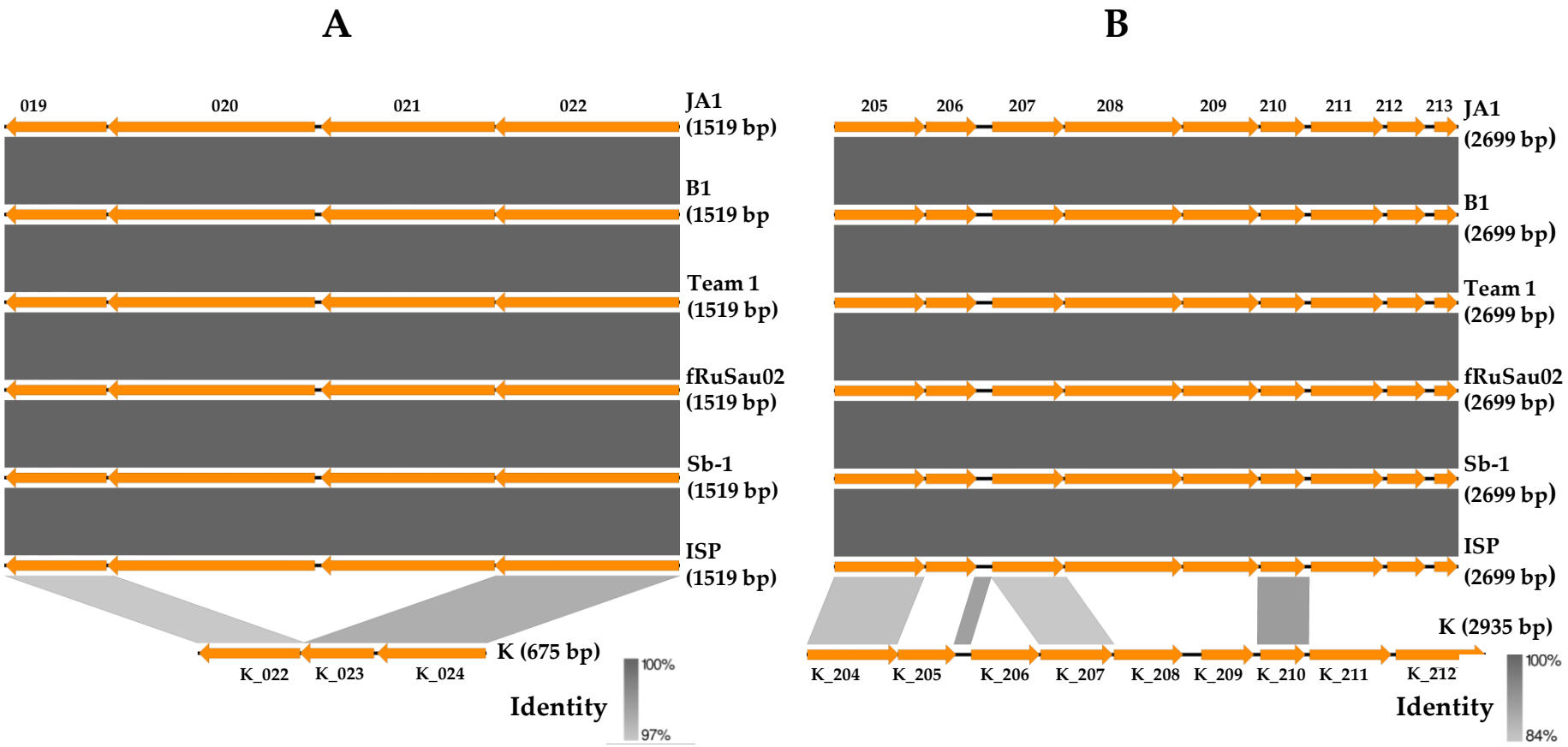


Figure S5. Comparison of regions within the genome of phage K to closely related staphylococcal phages (B1, JA1, Team 1, fRuSau02, Sb-1 and ISP) commonly employed as commercial phage mixtures using currently available annotations employing BLASTN and visualized with Easyfig. Several ORFs absent in phage K but present in both B1 and JA1 [phageJA1_020 & phageJA1_021 (A) and phage_206, phageJA1_208, phageJA1_209, phageJA1_211, phageJA1_212 & phageJA1_213 (B)] were also encountered in these closely related staphylococcal phages with reported wide host range.

Chapter III

Genomic characterisation of vB_SauM_B1 and vB_SauM_JA1, two *kay-like* bacteriophages isolated from a therapeutic phage mixture

3.1 Abstract

The genomes of the staphylococcal phages vB_SauM_B1 and vB_SauM_JA1 were characterised in detail. These phages belong to the *Myoviridae* family of viruses infecting bacteria based on their morphology. Using a combination of BLASTP, Interproscan and HHpred, these phages could be assigned functions involved in DNA replication, nucleotide metabolism, virion structure, morphogenesis and cell wall degradation. These phages possess genome sizes of 140,808bp and 139,484bp, respectively for B1 and JA1 with 219 and 215 predicted ORFs. Phylogenetic analysis revealed that both phages cluster within the same species and are also members of the genera *Kayvirus*. Both phages lack the restriction site GATC in their genome, making them insensitive to restriction enzymes such as *Sau*3A1 encoded by *Staphylococcus aureus*. In addition, the receptor binding protein candidates from phage vB_SauM_K was cloned, expressed and purified. These proteins were subjected to an agglutination assay involving *S. aureus* cell suspension. Successful agglutionation involving the receptor binding protein and an *S. aureus* target suggests that these proteins may be exploited for detection of specific bacteria in food products.

3.2 Introduction

Bacteriophages (phages) are the most abundant biological entities on earth, with an estimated global population of 10^{31} viral particles (Whitman *et al.,* 1998; Hendrix, 2002). Phages of the order *Caudovirales* (tailed phages) are the most studied of all the bacterial viruses examined since 1959 (Ackermann, 1996). This order of phages can be divided into four families, which includes *Myoviridae, Siphoviridae, Podoviridae* and the recently included classification "*Ackermannviridae"* (to this phage order) (Kropinski *et al.*, 2017). *Myoviridae* possess a characteristic long contractile tail with an icosahedral capsid (Łobocka *et al.*, 2012) and can be further divided into six subfamilies. These subfamilies are *Eucampyvirinae, Ounavirinae, Peduovirinae, Tevenvirinae, Vequintavirinae* and *Spounavirinae* (Lefkowitz *et al.*, 2018). Myoviruses infecting *Staphylococcus aureus* are classified within the genera *Kayvirus*, *Silviavirus* or *Twortvirus* (Lefkowitz *et al.*, 2018). These phages are strictly lytic, possessing genome sizes ranging from 127-141 kb, and they also have long terminal repeats (LTR) at their genome termini that may undergo homologous recombination enabling circularization of these phages (Łobocka *et al.*, 2012).

In the previous chapter, we reported the isolation of two phages from the Fersisi therapeutic phage mixture and designated them B1 (vB_SauM_B1) and JA1 (vB_SauM_JA1). Both possessed a wide host range against the Irish National MRSA bank consisting of twenty-one MLST isolates in addition to the relevant controls. Both phages were also similar to the well-known phage K at the nucleotide sequence level exhibiting a percentage identity of 95% and 94% to phages B1 and JA1 respectively. Based on the International Committee on the Taxonomy of Viruses (ICTV), all three phages are members of the recently described genus *Kayvirus* (Adriaenssens and Brister, 2017). We herein report the detailed characterisation of the genome sequences of both B1 and JA1, providing more insight into their genome genetics and classification.

3.3 Material and Methods

3.3.1 Phage propagation, purification and sequencing

Phages B1 and JA1 were propagated according to the method described in Chapter II. B1 was propagated on *S. aureus* host DPC5246 and JA1 on host CIT281189 to high titre (>1 x 10^9 plaque forming units [PFU/ml]). These high titre phages were then purified by isopycnic centrifugation through CsCl gradients and their genomic DNA extracted according to the protocol described in Chapter II. Genome sequencing of these phages was outsourced to GATC Biotech (Konstanz, Germany) and the sequencing method used is also described in Chapter II.

3.3.2 Bioinformatic analysis of phages B1 and JA1 genomes

The phage genomes of B1 and JA1 were autoannotated using Rapid Annotation using Subsystem Technology (RAST) and manually verified. The predicted ORFs were assigned a putative function using BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins), HHpred (https://toolkit.tuebingen.mpg.de/#/tools/hhpred; Söding *et al.*, 2005) and InterProscan (http://www.ebi.ac.uk/interpro/search/sequence-search; Mitchell *et al.*, 2015). ORFs possessing a transmembrane domain were identified with the use of TMHMM v.2 (http://www.cbs.dtu.dk/services/TMHMM/; Krogh *et al.*, 2001). Phylogenetic analysis was carried out with the Virus Classification and Tree Building Online Resource (VICTOR) using the Genome-BLAST Distance Phylogeny (GBDP) method (Meier-Kolthoff and Göker, 2017), under the recommended setting for prokaryotic viruses. The resulting intergenomic distances were used to infer a balanced minimum evolutionary tree with branch support via FASTME, including Subtree Pruning and Regrafting (SPR) post processing (Lefort, Desper and Gascuel, 2015) for the formula D6. Branch support was inferred from 100 pseudo-bootstrap replicates each. The tree was rooted at midpoint (Farris, 1972) and visualized with

FigTree (http://tree.bio.ed.ac.uk/software/figtree). Taxon boundaries at the species, genus and family level were estimated with the OPTSIL program (Göker *et al.*, 2009), the recommended clustering thresholds (Meier-Kolthoff and Göker, 2017) and an F value (fraction of links required for cluster fusion) of 0.5 (Meier-Kolthoff *et al.*, 2014).

3.3.3 SDS-PAGE analysis on phage structural proteins

Phage samples (B1 and JA1) were prepared for SDS-PAGE analysis by adding 30µl of high titre phage (>1 x 10^9) to 20µl of sample buffer (10% (w/v) SDS; 0.5% (w/v) bromophenol blue; 0.5M Tris-Hcl, pH 6.8; 2.5ml of glycerol, made up to 9.5ml with deionized water; 50µl of β-marcaptoethanol was added to the 950µl of this solution prior to use). These samples were then boiled for 5mins before being loaded onto a polyacrylamide gel. Phage structural proteins were separated using a 12% (w/v) resolving gel (30% (w/v) bisacrylamide; 1.5M Tris-Hcl, pH 8.8; 10% (w/v) SDS; 10% (w/v) ammonium persulphate; 0.25% (w/v) tetramethylethylenediamine) and electrophoresis conducted in Tris-Glycine buffer at 200V for 1hr in the BioRad Mini-Protean gel apparatus (BioRad, USA). Proteins were visualized by a colloidal Coomassie staining (0.02% (w/v) Coomassie Brilliant Blue G-250; 5% (w/v) aluminium sulphate-(14-18)-hydrate; 10% (v/v) ethanol; 2% (v/v) orthophosphoric acid) and destained in a solution consisting of 10% (v/v) ethanol and 2% (v/v) orthophosphoric acid.

3.3.4 Cloning, expression and purification of putative receptor binding proteins of phage K

DNA regions encoding the putative receptor binding proteins K_120 and K_122 were amplified by polymerase chain reaction (PCR) using phage DNA as template. For K_120, the forward primer 5'-AGATCGGATCCATGGCATTTAACTACACG-3' and the reverse primer 5'- CTACTCGAGTTATCCTCTATTAATTCCCAT-3' was used in its amplification. In the case of K_122, the forward primer 5'- TAGAGGATCCGCATTAAATTTTACTAC-3' and

the reverse primer 5'- CTACTCGAGTTACTATGGCATATTAATAC-3' were used. Both PCR products were digested with the restriction enzymes *BamH*I and *Xho*I and ligated to the *E. coli* shuttle vector pET28a. The resulting plasmids were transformed into *E. coli* JM109 (DE3). Protein expression was carried out by growing the recombinant *E. coli* at 37ºC to an OD_{600nm} of 0.5 and isopropyl β-D_1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5mM. The culture was incubated at 30ºC overnight, harvested and then resuspended in buffer A (20mM Tris HCl pH 8.5, 0.5M NaCl). Cells were lysed using Bugbuster protein extraction reagent (Novagen) and the cellular debris removed by centrifugation, together with filter sterilization of the supernatant through a 0.2µm filter. Protein purification was performed by affinity chromatography using the AKTA *Start* system (GE healthcare). The purified protein was concentrated and desalted by buffer exchange using an Amicon Ultra-15 Centrifugal Filter Unit with a molecular weight cut-off (MWCO) of 10kDa (Merck Millipore, Madrid Spain).

3.3.5 Phage adsorption on cell coated with receptor binding protein and agglutination assay

The adsorption efficiency of phage K on *S. aureus* DPC5246 coated with purified receptor binding protein (RBP) was performed to determine if the RBPs "K_120 and K_122" were capable of inihibiting phage infection. This involved establishing a one percent inoculum of overnight *S. aureus* culture and growing at 37ºC to an OD_{600nm} of 0.2 (approximately 10^8 CFU/ml). Afterwards, 100µl of *S. aureus* cells was pre-incubated at room temperature with 100µl of purified RBP (57mg/ml) for 1hr. The resulting mixture was added to 100µl of phage at a titre of 10^6 PFU/ml and the entire content incubated at room temperature for 5mins. Bound phage was separated from unbound phage by centrifugation and the titre of unbound phage was enumerated by plaque assay technique. A sample containing no RBP was also set up as a control.

Agglutination assay was performed according to the protocol described in Javed *et al.*, (2013) with some modification. Briefly, an overnight of DPC5246 was harvested by centrifugation and resuspended in PBS to the same cell density. A 50µl aliquot of the cell suspension was mixed with 1µl of RBP (57mg/ml) on a glass slide and the mixture was examined for agglutination. As a control, PBS was added to the cell suspension on a glass slide.

3.4 Results and discussion

3.4.1 General genomic features of phages B1 and JA1

Phages B1 and JA1 had genome sizes of 139,831bp and 139,484bp respectively, with an overall G+C content of 30.3% for both phages. These phages were predicted to contain a total of 219 and 215 open reading frames (ORFs) for B1 and JA1, respectively, some of which encoded proteins with little to no homology to protein sequences in the databases, and others encoding proteins with a high degree of homology to known phage proteins. The majority of the ORFs (201 for B1 and 198 for JA1) initiated translation with an AUG start codon, whereas few ORFs (11 for both B1 and JA1) initiated translation with a UUG codon and even fewer ORFs (7 for B1 and 6 for JA1) with a GUG codon. Likewise, the majority of the ORFs (159 for B1 and 157 for JA1) terminated translation with a TAA stop codon, whereas several ORFs (46 for B1 and JA1) terminated translation with a TAG codon and few ORFs (14 for B1 and 12 for JA1) with a TGA codon. Both phages also possess four tRNA genes (tRNA-Met, tRNA-Trp, tRNA-Phe and tRNA-Asp) in their genomes, the same four tRNA genes encountered in the genomes of other staphylococcal phages such as K, ISP and MSA6 (O'Flaherty *et al.*, 2004; Vandersteegen *et al.*, 2011; Łobocka *et al.*, 2012). Comparing the genomes of B1 and JA1 revealed that while both phages were closely related to each other with a percentage identity of 99%. The major difference between both comprised of small insertions and deletions (Figure 1), resulting in four additional hypothetical proteins in B1 (PhageB1_009, 016, 202, 203).

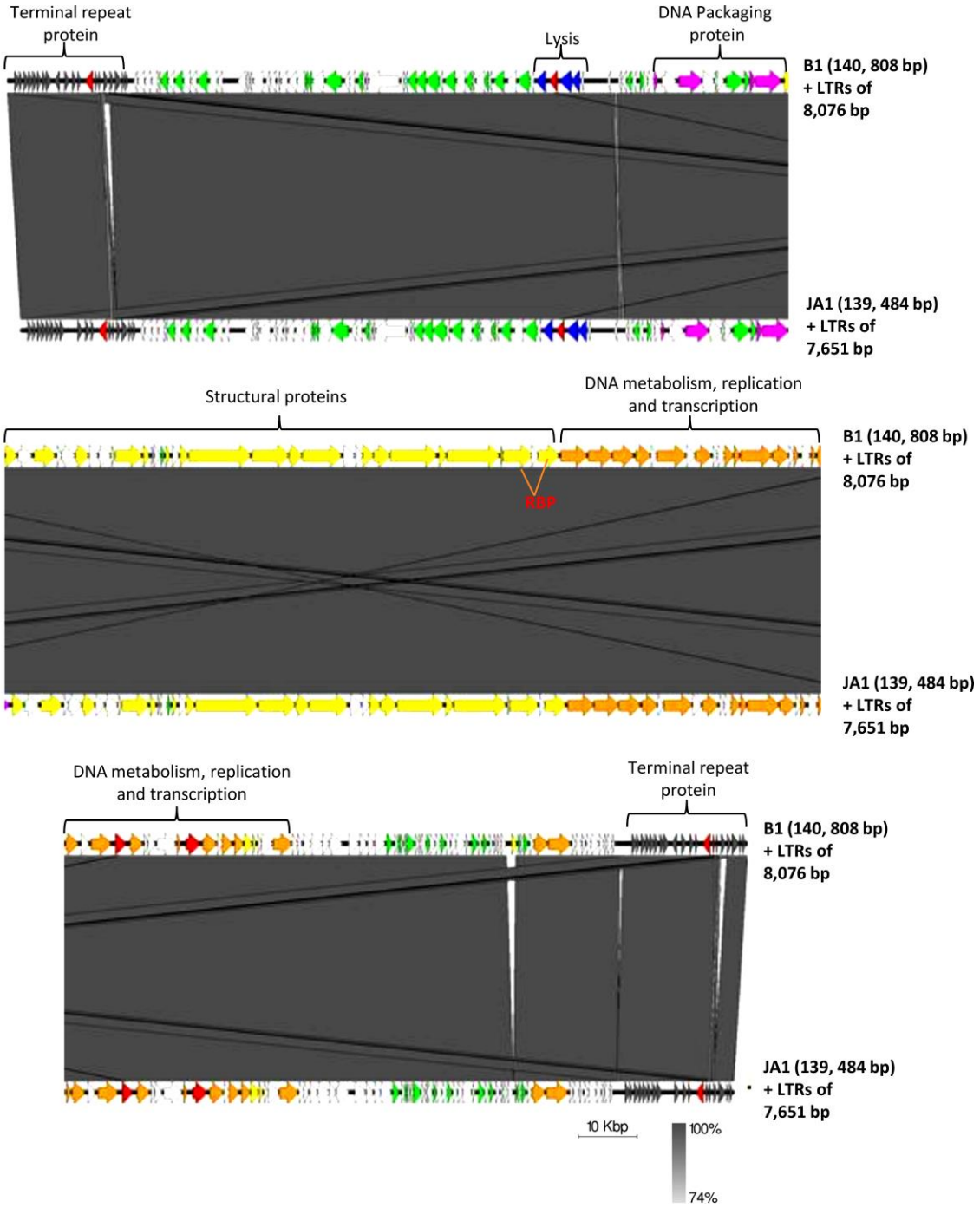


Figure 1. Genome comparison of phages B1 to JA1 employing BLASTN and visualized with Easyfig. Genome map shows ORF organisation, with arrows indicating location of genes on both phage genomes. Black arrows represent the genes encoding the terminal repeat proteins; Blue arrows represent the genes encoding the lysis proteins; Pink arrows represent the genes encoding the DNA packaging proteins; Yellow arrows represent the genes encoding the structural/morphogenesis proteins, Location of the receptor binding protein (RBP) shown; Orange arrows represent the gene encoding the DNA metabolism, replication and transcription proteins; Green arrows represent the genes encoding proteins with putative function; White arrows represent hypothetical proteins. Arrows are drawn to scale with the gene orientation portrayed by the arrow direction. B1 and JA1 are 99% identical but differed with respect to four hypothetical genes present in phage B1 but absent in JA1.

3.4.2 Modular organisation of B1 and JA1

The genomes of phages B1 and JA1 are organized into modules, similar to those of other virulent staphylococcal phages (Łobocka *et al.*, 2012). These functional modules are divided into modules for DNA replication/transcription, structural/morphogenesis, DNA packaging and lysis.

3.4.2.1 DNA replication and transcription module

Sequence-based prediction identified several genes in the genomes of B1 and JA1 involved in the metabolism and synthesis of DNA, suggesting that both phages are capable of replicating their DNA with less reliance on host machinery. Genes identified in this module encodes proteins such as DNA helicases (B1_128/130 and JA1_125/127), which are known to be responsible for unwinding DNA at the chromosomal replication fork (Jones *et al.*, 2001); a primase (B1_135 and JA1_132) possessing a Toprim_DnaG domain, similar to those encountered in *Bacillus subtilis* SPP1 primase. This protein interacts with the helicase leading to a more stable complex between the helicase and ssDNA, leading to increased helicase activity (Ayora *et al.*, 1998). This module also includes the resolvase (B1_138 and JA1_135), like those of the well-characterized T4 endonuclease VII and T7 endonuclease I enzymes, which are involved in the de-branching of DNA structures prior to packaging, in addition to their main role in resolving four-way intermediates that form during DNA recombination and repair events (Wyatt and West, 2014).

Phages of the *Spounavirinae* subfamily of which both B1 and JA1 are members, do not encode their own RNA polymerase but instead rely on host RNA polymerase for the transcription of their genes (Łobocka *et al.*, 2012). JA1 and B1 encodes an RNA polymerase

sigma factor (B1_158 and JA1_155) that binds to host core RNA polymerase redirecting it to recognise phage promoters for the transcription of phage genes (Dehbi *et al.*, 2009). These phages also encode the ribonucleotide reductase (B1_139/140/141 and JA1_136/137/18), responsible for the formation of deoxyribonucleotide from ribonucleotide (Dwivedi *et al.*, 2013) as well as a DNA polymerase (B1_146/148/150 and JA1_143/145/147) and a DNA repair protein (B1_154/156 and JA1_151/153), both of which are interrupted by introns encoding two proteins (a hypothetical protein and an endonuclease) in the case of the DNA polymerase and a single protein (endonuclease) for the DNA repair protein. Other proteins involved in the replication and transcription of B1/JA1 phage DNA includes the replication protein (B1_129 and JA1_126), recombination exonuclease (B1_131/133 and JA1_128/130), thioredoxin (B1_143 and JA1_140) and integration host factor (B1_145 and JA1_142).

3.4.2.2 DNA packaging module

The packaging of the DNA into the bacteriophage capsid is usually dependent on a powerful machinery that comprises of the terminase and portal proteins (Oliveira *et al.*, 2013). It was found that the ORFs encoding the terminase proteins (B1_086/088 and JA1_083/085) for both B1 and JA1 were interrupted by an intron, encoding a hypothetical protein (PhageB1_087 and JA1_084), of unknown function. These ORFs (B1_086/088 and JA1_083/085) could be assembled into the large terminase subunit, which are known to be involved in the site-specific binding, cutting and translocation of DNA during the initial and late stages of packaging (Duffy and Feiss, 2002). PhageB1_094/095 and JA1_091/092 on the other hand encode the portal protein, which is known to form a cleft and act as passage through which the viral DNA is packaged and ejected. It may also act as connector between the phage head and the tail proteins (Dröge *et al.*, 2000; Moore and Prevelige, 2002).

3.4.2.3 Structural/morphogenesis module

The structural/morphogenesis module of phages B1 and JA1 is located between the modules for DNA packaging and the module for DNA replication, transcription and metabolism, similar to those of other *Staphylococcus* phages like K and Stau2 (O'Flaherty *et al.*, 2004; Hsieh *et al.*, 2016). The structural genes in this module are occasionally interspaced by hypothetical proteins (Figure 1) and some of the gene products in this module are conserved in the *Spounavirinae* subfamily of myoviruses (Kwan *et al.*, 2005; Łobocka *et al.*, 2012).

Structural proteins from both B1 and JA1 were analysed by SDS-PAGE (Figure 2) and several polypeptide bands ranging from 28 to 143kDa were identified for both phages. The most abundant virion protein was identified as the major capsid protein (B1_098 and JA1_095) based on its migration to a position on the gel matching its predicted molecular weight. This protein was very similar to the major capsid protein of the *staphylococcus* phage 812, whose structural function was confirmed in a study by Eyer *et al.*, (2007). Together with the prohead protease (B1_096 and JA1_93) and scaffold protein (B1_125 and JA1_ 122), these proteins play a role in the development of the phage head. The second most intense band was identified as the major tail sheath protein (B1_105 and JA1_102) and is known to make tail contraction possible during phage infection process (Eyer *et al.*, 2007). Other proteins encoded by genes responsible for the morphogenesis of the phage tail in both B1 and JA1 include the tail tube (B1_106 and JA1_103), tail morphogenetic protein (B1_112 and JA1_109), DNA transfer protein (B1_113 and JA1_110), tail lysin (B1_114 and JA1_111) and the adsorption-associated tail protein (B1_123 and JA1_120). In addition, both B1 and JA1 also encode baseplate proteins (B1_119/120/122 and JA1_116/117/119), which possesses a huge degree of similarity to the *Staphylococcus* phage 812 baseplates. These proteins, upon binding to the *Staphylococcus* host, undergo conformational changes into a

two layered structure parallel to the host cell wall in events leading to contraction of the tail sheath and eventual release of the viral genome (Nováček *et al.*, 2016).

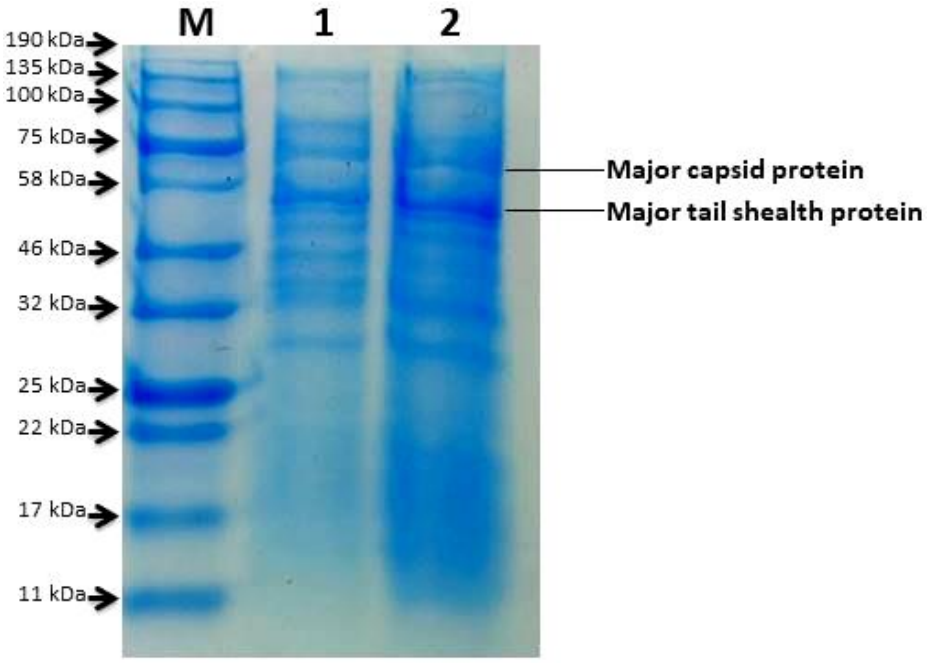


Figure 2. Separation of B1 and JA1 structural proteins by 12% (w/v) SDS-PAGE assay. Lane M: molecular weight marker, Lane 1: phage B1 and Lane 2: phage JA1.

3.4.2.3.1 Receptor binding proteins (RBP) of B1 and JA1

The protein B1_127/JA1_124 showed homology to the receptor binding protein (gp108) of the *Listeria* phage A511, with sequence alignment of gp108 and its three ortholoques in *Staphylococcus* phages B1, JA1 and K (B1_127, JA1_124 and K_122) showing 32% sequence identity over 151 aminoacids in its N-terminus (Figure 3). B1_127 and JA1_124 were very similar to K_122, possessing over 99% identity between their aminoacids sequences. Receptor binding proteins between related phages have been reported to possess a conserved N-terminal region responsible for attaching the tail receptor binding element to the phage (Duplessis and Moineau, 2001) as well as a non-conserved C-terminal region allowing for flexibility in the type of host cell and ligand recognized (Habann *et al.*, 2014).

A conserved domain search on B1_125/JA1_122, which are highly similar to K_120 revealed that these ORFs possess a carbohydrate binding domain and thus may encode another receptor binding protein in phages B1 and JA1. Phages possessing two receptor binding proteins have been reported in the literature and this has been attributed to result in the wide host range of such phages (Takeuchi *et al.*, 2016).

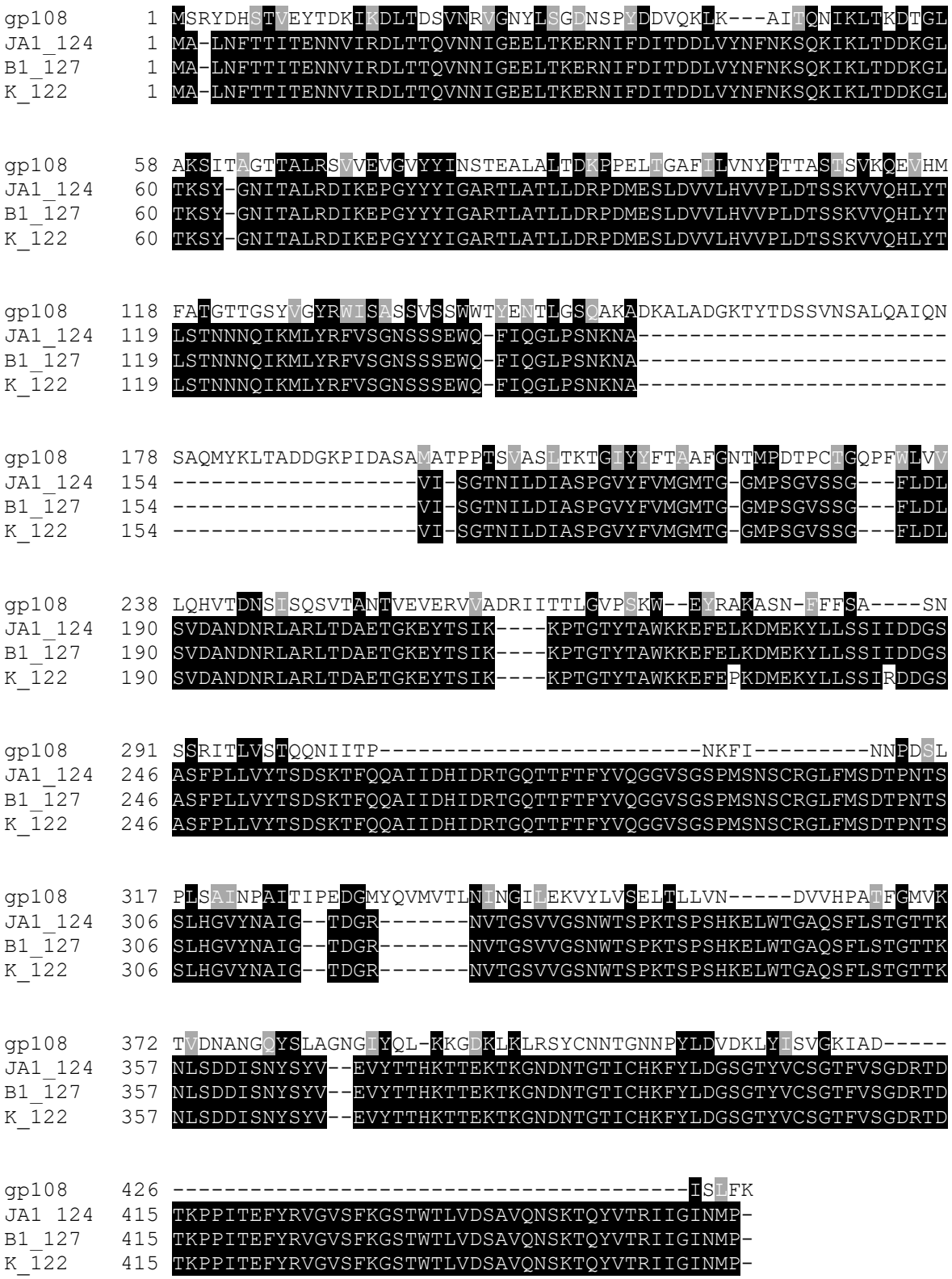


Figure 3. Amino acid sequence alignment of the receptor binding proteins for *Listeria* phage A511 (gp108), *Staphylococcus* phages B1 (B1_127), JA1 (JA1_124) and K (K_122).

Owing to the similarities between the receptor binding proteins of B1, JA1 and K, the ORFs encoding the two putative receptor binding proteins from phage K (protein K_122) were cloned and expressed in *E. coli*. The extent of adsorption blocking of phage K to the *S. aureus* host cells was investigated. For the control, no protein was added and the resulting average titre of unadsorbed phage was found to be $4.2x10^4$ PFU/ml and was used for comparison with titres calculated for samples where the receptor binding proteins had been added (Table 1). Following the addition of purified K_122, the average titre of unadsorbed phage was calculated to be $3.57x10^4$ PFU/ml, which is similar to the control titre. On the other hand, when K_120 was used, the average titre of unadsorbed phage was calculated to be $6.1x10^4$ PFU/ml (Table 1). This slightly higher value than the control titre was consistently observed in all replicates, thus suggesting a slight adsorption inhibition of phage K by the protein K_120. It has previously been reported that wall techoic acids (WTA) serve as receptors for *Staphylococcus* phages (Xia *et al.*, 2011) and these carbohydrate molecules are quite abundant in the staphylococcal cell wall (Baur *et al.*, 2014). This suggests that the protein K_120 is possibly blocking access to these moities, thus leading to the slight inhibition encountered with K_120. Both K_120 and K_122 are homologues of SA012_103 and SA012_105 respectively, which have been experimentally confirmed as receptor binding proteins for ϕSA012 (Takeuchi *et al.*, 2016). Interestingly, the addition of K_122 to staphylococcal cells resulted in agglutination as observed on a microscope slide (Figure 4) indicating some interaction of this protein with the staphylococcal cell surface. This phenomenon has also been encountered with other receptor binding proteins upon incubation with host cells (Javed *et al.*, 2013; Habann *et al.*, 2014). According to this scientific literature, RBP-based agglutination can be exploited for detection of specific bacteria in food products, and thus further investigation of the protein K_122 may well be warrented.

Table 1. Adsorption blocking of phage K to *S. aureus* cells coated with purified receptor binding protein.

Samples	Titre1 (PFU/ml)	Titre2 (PFU/ml)	Titre3 (PFU/ml)	Average titre (PFU/ml)
Original phage titre	$1.13x10^6$	$1.17x10^6$	$1.06x10^6$	$1.12x10^6$
No RBP (Control)	$4.9x10^4$	$3.9x10^4$	$3.8x10^4$	$4.2x10^4$
K_122	$3.2x10^4$	$3.6x10^4$	$3.9x10^4$	$3.57x10^4$
K_120	$6.3x10^4$	$5.5x10^4$	$6.5x10^4$	$6.1x10^4$

*Data is represented in PFU/ml of phage K that did not adsorb to RBP-coated host cells based on three independent experiments

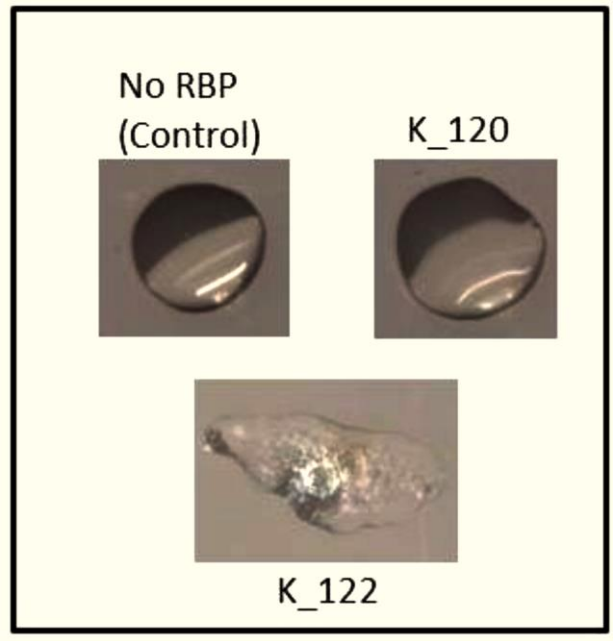


Figure 4. *Staphylococcus aureus* DPC5246 cell suspensions on microscope slides incubated in the presence of purified receptor binding proteins K_120 and K_122. Agglutinatination of staphylococcal cells was evident in the case of protein K_122.

3.4.2.4 Lysis module

The Gram-positive phage lysis cassette usually comprises of a holin and an endolysin, which are necessary for the release of mature phages from their bacterial host (Ajuebor *et al.*, 2016). Holins form lesions in the cell membrane of the host at a specific time point, thereby paving the way for endolysins to access their cell wall substrate. The lysis module for phages B1 and JA1 is located within the divergently transcribed region (Figure 1) similar to that of *Staphylococcus* phage K (O'Flaherty *et al.*, 2004). B1_076 and JA1_073 encode the holin in both phages, possessing a phage_holin_1 superfamily as well as two transmembrane domains (TMDs) located at aminoacids 24-46 and 56-73 (Figure 5). Functional analysis on a similar phage holin (HolGH15) revealed that both TMDs were essential for its lytic activity (Song *et al.*, 2016). On the other hand, B1_072/075 and JA1_070/072 encode the endolysin in both phages, possessing a CHAP, Amidase_2 and SH3b domain identical to LysK (O'Flaherty *et al.*, 2005). Functional analysis on this enzyme revealed that only the CHAP domain exhibited lytic activity on the cell substrate.

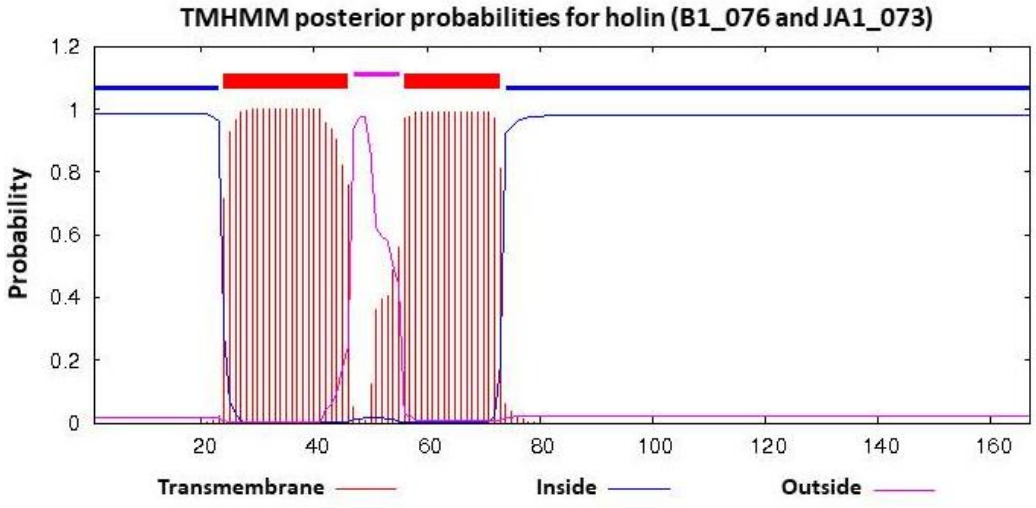


Figure 5. Amino acid sequence of the holin protein B1_076 (and JA1_073) applied to the TMHMM program for predicting the transmembrane helices. The red lines represent the transmembrane region, blue lines represent the cytoplasmic region and the pink line represents the non-cytoplasmic region of the holin protein. Two transmembrane helices at aminoacids 24-46 and 56-73 are evident.

3.4.3 Terminal repeat region

The genomes of both B1 and JA1 possess long terminal repeats of 8,076bp and 7,651bp respectively, with both encoding 20 and 18 terminal repeat proteins, respectively. The proteins in these regions are small molecular weight proteins (less than 20kDa) involved in the host take-over event similar to the terminal repeat proteins of phiIPLA-RODI (Gutiérrez *et al.*, 2015). In addition, the gene for a group I homing HNH endonuclease (B1_013 and JA1_012) was also identified in the terminal repeat region for both phages.

3.4.4 Phages B1 and JA1 lack restriction sites for common staphylococcal host-encoded endonucleases

S. aureus is known to encode the Type II restriction-modification systems *Sau*3A1 and *Sau*96I, which recognize 5'-GATC-3' and 5'-GGNCC-3' sequences, respectively (Sadykov, 2016). Analysis on the genomes of both B1 and JA1 revealed that it lacked the GATC site and possessed a single GGTCC site, similar to the genomes of phages K and fRuSau02 (O'Flaherty *et al.*, 2004; Leskinen *et al.*, 2017). In addition, no gene with significant homology to any DNA methylases were encountered in B1 or JA1. This suggests that these phages, through elimination of restriction endonuclease-susceptible nucleotide sequences, have evolved an efficient phage defence strategy against these specific endonucleases.

3.4.5 Phylogenetic relationship of B1 and JA1 to other phages within the *Myoviridae* family

Based on the morphologies of B1 and JA1, as shown in the previous chapter, these phages have been classified into the *Spounavirinae* subfamily of *Myoviridae*, possessing an icosahedral capsid and a long contractile tail. To better understand the position of these phages within the *Myoviridae* family, a whole-genome phylogenetic analysis of B1, JA1 and thirty other similar phages was done using the bioinformatic program VICTOR (Meier-

Kolthoff and Göker, 2017). This analysis yielded an average support of 63% and an OPSTIL clustering leading to 20 clusters at the species level, 3 clusters at the genus level and a single cluster at the family level. B1 and JA1 cluster within the same species as does vB_SauM_fRuSau02, ISP and Team 1, with more than 95% percentage nucleotide identity between their genome. These phages are also closely related to phage K, GH15, vB_Sau_Clo6, vB_Sau_S24 and pSco_10. These phages make up the genus *Kayvirus* with their closest evolutionary relationship appearing to be phages of the genera *Silviavirus*.

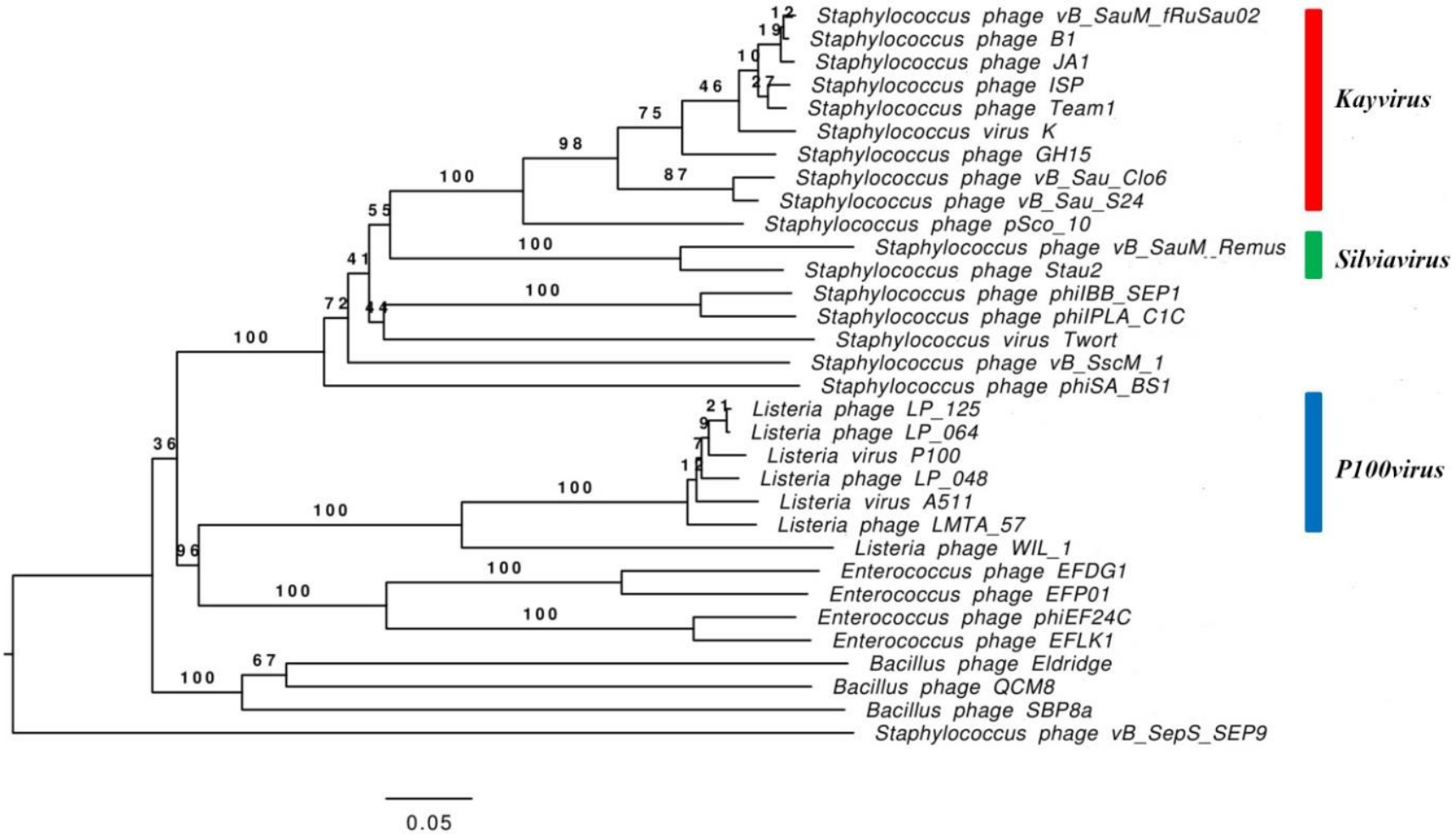


Figure 6. Genome-wide phylogenetic analysis of 32 phages in the myoviridae family. This analysis was performed using VICTOR; Virus Classification and Tree Building Online Resource (Meier-Kolthoff and Göker, 2017).

3.5 Conclusion

B1 and JA1 are virulent bacteriophages, which infect a wide range of MRSA sequence types in the Irish National MRSA collection. Genetic characterisation of these phages revealed that both phages contain several introns in their genome which interrupts several essential genes. B1 and JA1 are 99% identical to each other on the nucleotide level, with differences between both phages comprising of small insertions or deletions resulting in the presence of four hypothetical proteins in B1 that are absent in JA1. Interestingly, both phages lack the restriction site GATC in their genome, suggesting these phages have evolved a defence strategy against host-encoded restriction-modification systems. These phages are members of the genus *Kayvirus* and are thus closely related to phage K. In addition, the receptor binding protein candidates from phage K was cloned, expressed and purified. These proteins formed agglutination with *S. aureus* cell, suggesting that purified receptor binding proteins may be exploited for detection of specific bacteria in food products.

3.6 References

1. Ackermann, H. W. (1996) ‘Frequency of morphological phage descriptions in 1995.’, *Archives of Virology*, 141(2), pp. 209–218. doi: 10.1007/s007050170120.
2. Adriaenssens, E. and Brister, J. R. (2017) ‘How to name and classify your phage: An informal guide’, *Viruses*, 9(4), p. 70. doi: 10.3390/v9040070.
3. Ajuebor, J., McAuliffe, O., O’Mahony, J., Ross, R. P., Hill, C. and Coffey, A. (2016) ‘Bacteriophage endolysins and their applications.’, *Science Progress*, 99(2), pp. 183–199. doi: 10.3184/003685016X14627913637705.
4. Ayora, S., Langer, U. and Alonso, J. C. (1998) ‘*Bacillus subtilis* DnaG primase stabilises the bacteriophage SPP1 G40P helicase-ssDNA complex’, *FEBS Letters*, 439(1–2), pp. 59–62. doi: 10.1016/S0014-5793(98)01337-4.
5. Baur, S., Rautenberg, M., Faulstich, M., Grau, T., Severin, Y. and Weidenmaier, C. (2014) ‘A nasal epithelial receptor for *Staphylococcus aureus* WTA governs adhesion to epithelial cells and modulates nasal colonization’, *PloS One*. P10(5). p. e1004089. doi: 10.1371/journal.ppat.1004089.
6. Dehbi, M., Moeck, G., Arhin, F. F., Bauda, P., Bergeron, D., Kwan, T., Liu, J., McCarty, J., DuBow, M. and Pelletier, J. (2009) ‘Inhibition of transcription in *Staphylococcus aureus* by a primary sigma factor-binding polypeptide from phage G1’, *Journal of Bacteriology*, 191(12), pp. 3763–3771. doi: 10.1128/JB.00241-09.
7. Dröge, A., Santos, M. A., Stiege, A. C., Alonso, J. C., Lurz, R., Trautner, T. A. and Tavares, P. (2000) ‘Shape and DNA packaging activity of bacteriophage SPP1 procapsid: Protein components and interactions during assembly’, *Journal of Molecular Biology*, 296(1), pp. 117–132. doi: 10.1006/jmbi.1999.3450.
8. Duffy, C. and Feiss, M. (2002) ‘The large subunit of bacteriophage λ’s terminase plays a role in DNA translocation and packaging termination’, *Journal of Molecular Biology*, 316(3), pp.

547–561. doi: 10.1006/jmbi.2001.5368.

9. Duplessis, M. and Moineau, S. (2001) ‘Identification of a genetic determinant responsible for host specificity in *Streptococcus thermophilus* bacteriophages’, *Molecular Microbiology*, 41(2), pp. 325–336. doi: 10.1046/j.1365-2958.2001.02521.x.
10. Dwivedi, B., Xue, B., Lundin, D., Edwards, R. A. and Breitbart, M. (2013) ‘A bioinformatic analysis of ribonucleotide reductase genes in phage genomes and metagenomes’, *BMC Evolutionary Biology*, 13(1), pp. 1–17. doi: 10.1186/1471-2148-13-33.
11. Eyer, L., Pantůček, R., Zdráhal, Z., Konečná, H., Kašpárek, P., Růžičková, V., Hernychová, L., Preisler, J. and Doškař, J. (2007) ‘Structural protein analysis of the polyvalent staphylococcal bacteriophage 812’, *Proteomics*, 7(1), pp. 64–72. doi: 10.1002/pmic.200600280.
12. Farris, J. S. (1972) ‘Estimating phylogenetic trees from distance matrices’, *The American Naturalist*, 106(951), pp. 645–668. doi: 10.1086/282802.
13. Göker, M., García-Blázquez, G., Voglmayr, H., Tellería, M. T. and Martín, M. P. (2009) ‘Molecular taxonomy of phytopathogenic fungi: a case study in Peronospora.’, *PloS One*, 4(7), p. e6319. doi: 10.1371/journal.pone.0006319.
14. Gutiérrez, D., Vandenheuvel, D., Martínez, B., Rodríguez, A., Lavigne, R. and García, P. (2015) ‘Two phages, phiIPLA-RODI and phiIPLA-C1C, lyse mono- and dual-species staphylococcal biofilms’, *Applied and Environmental Microbiology*, 81(10), pp. 3336–3348. doi: 10.1128/AEM.03560-14.
15. Habann, M., Leiman, P. G., Vandersteegen, K., Van den Bossche, A., Lavigne, R., Shneider, M. M., Bielmann, R., Eugster, M. R., Loessner, M. J. and Klumpp, J. (2014) ‘*Listeria* phage A511, a model for the contractile tail machineries of SPO1-related bacteriophages’, *Molecular Microbiology*, 92(1), pp. 84–99. doi: 10.1111/mmi.12539.
16. Hendrix, R. W. (2002) ‘Bacteriophages: Evolution of the majority’, *Theoretical Population*

Biology, 61(4), pp. 471–480. doi: 10.1006/tpbi.2002.1590.

17. Hsieh, S. E., Tseng, Y. H., Lo, H. H., Chen, S. T. and Wu, C. N. (2016) 'Genomic analysis of *Staphylococcus* phage Stau2 isolated from medical specimen', *Virus Genes*. Springer US, 52(1), pp. 107–116. doi: 10.1007/s11262-015-1276-8.
18. Javed, M. A., Poshtiban, S., Arutyunov, D., Evoy, S. and Szymanski, C. M. (2013) 'Bacteriophage receptor binding protein based assays for the simultaneous detection of *Campylobacter jejuni* and *Campylobacter coli*', *PLoS One*, 8(7). p. e69770. doi: 10.1371/journal.pone.0069770.
19. Jones, C. E., Mueser, T. C., Dudas, K. C., Kreuzer, K. N. and Nossal, N. G. (2001) 'Bacteriophage T4 gene 41 helicase and gene 59 helicase-loading protein: a versatile couple with roles in replication and recombination.', *Proceedings of the National Academy of Sciences of the United States of America*, 98(15), pp. 8312–8318. doi: 10.1073/pnas.121009398.
20. Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. L. (2001) 'Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.', *Journal of Molecular Biology*, 305(3), pp. 567–580. doi: 10.1006/jmbi.2000.4315.
21. Kropinski, A. M., Anany, H., Kuhn, J. H., Tolstoy, I., Kutter, E. and Adriaenssens, E. M. (2017) 'To create a new bacteriophage family, *Ackermannviridae*, containing two (2) new subfamilies including four (4) genera', *ICTV Report*, pp. 1–12. doi: 10.13140/RG.2.2.29173.88800.
22. Kwan, T., Liu, J., DuBow, M., Gros, P. and Pelletier, J. (2005) 'The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages.', *Proceedings of the National Academy of Sciences of the United States of America*, 102(14), pp. 5174–5179. doi: 10.1073/pnas.0501140102.

23. Lefkowitz, E. J., Dempsey, D. M., Hendrickson, R. C., Orton, R. J., Siddell, S. G. and Smith, D. B. (2018) 'Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV)', *Nucleic Acids Research*. Oxford University Press, 46(D1), pp. D708–D717. doi: 10.1093/nar/gkx932.
24. Lefort, V., Desper, R. and Gascuel, O. (2015) 'FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program', *Molecular Biology and Evolution*, 32(10), pp. 2798–2800. doi: 10.1093/molbev/msv150.
25. Leskinen, K., Tuomala, H., Wicklund, A., Horsma-Heikkinen, J., Kuusela, P., Skurnik, M. and Kiljunen, S. (2017) 'Characterization of vB_SauM-fRuSau02, a twort-like bacteriophage isolated from a therapeutic phage cocktail', *Viruses*, 9(9), pp. 1–19. doi: 10.3390/v9090258.
26. Łobocka, M., Hejnowicz, M. S., Dąbrowski, K., Gozdek, A., Kosakowski, J., Witkowska, M., Ulatowska, M. I., Weber-Dąbrowska, B., Kwiatek, M., Parasion, S., Gawor, J., Kosowska, H. and Głowacka, A. (2012) 'Genomics of staphylococcal Twort-like phages--potential therapeutics of the post-antibiotic era.', *Advances in Virus Research*, 83, pp. 143–216. doi: 10.1016/B978-0-12-394438-2.00005-0.
27. Meier-Kolthoff, J. P. and Göker, M. (2017) 'VICTOR: genome-based phylogeny and classification of prokaryotic viruses', *Bioinformatics*, 33(July 2017), pp. 3396–3404. doi: 10.1093/bioinformatics/btx440.
28. Meier-Kolthoff, J. P., Hahnke, R. L., Petersen, J., Scheuner, C., Michael, V., Fiebig, A., Rohde, C., Rohde, M., Fartmann, B., Goodwin, L. A., Chertkov, O., Reddy, T. B. K., Pati, A., Ivanova, N. N., Markowitz, V., Kyrpides, N. C., Woyke, T., Goker, M. and Klenk, H. P. (2014) 'Complete genome sequence of DSM 30083T, the type strain (U5/41T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy', *Standards in Genomic Sciences*, 9(1), pp. 1–19. doi: 10.1186/1944-3277-9-2.
29. Mitchell, A., Chang, H. Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C.,

McMenamin, C., Nuka, G., Pesseat, S., Sangrador-Vegas, A., Scheremetjew, M., Rato, C., Yong, S. Y., Bateman, A., Punta, M., Attwood, T. K., Sigrist, C. J. A., Redaschi, N., Rivoire, C., Xenarios, I., Kahn, D., Guyot, D., Bork, P., Letunic, I., Gough, J., Oates, M., Haft, D., Huang, H., Natale, D. A., Wu, C. H., Orengo, C., Sillitoe, I., Mi, H., Thomas, P. D. and Finn, R. D. (2015) 'The InterPro protein families database: the classification resource after 15 years.', *Nucleic Acids Research*, 43(Database issue), pp. D213-221. doi: 10.1093/nar/gku1243.

30. Moore, S. D. and Prevelige, P. E. (2002) 'DNA packaging: a new class of molecular motors.', *Current Biology : CB*, 12(3), pp. R96-98. doi: 10.1016/S0960-9822(02)00670-X.
31. Nováček, J., Šiborová, M., Benešík, M., Pantůček, R., Doškař, J. and Plevka, P. (2016) 'Structure and genome release of *Twort-like Myoviridae* phage with a double-layered baseplate', *Proceedings of the National Academy of Sciences*, 113(33), pp. 9351–9356. doi: 10.1073/pnas.1605883113.
32. O'Flaherty, S., Coffey, A., Edwards, R., Meaney, W., Fitzgerald, G. F. and Ross, R. P. (2004) 'Genome of staphylococcal phage K: A new lineage of *Myoviridae* infecting Gram-positive bacteria with a low G+C Content', *Journal of Bacteriology*, 186(9), pp. 2862–2871. doi: 10.1128/JB.186.9.2862-2871.2004.
33. O'Flaherty, S., Ross, R. P., Meaney, W., Fitzgerald, G. F., Elbreki, M. F. and Coffey, A. (2005) 'Potential of the polyvalent Anti-*Staphylococcus* bacteriophage K for control of antibiotic-resistant staphylococci from hospitals', *Applied and Environmental Microbiology*, 71(4), pp. 1836–1842. doi: 10.1128/AEM.71.4.1836-1842.2005.
34. Oliveira, L., Tavares, P. and Alonso, J. C. (2013) 'Headful DNA packaging: Bacteriophage SPP1 as a model system', *Virus Research*. Elsevier B.V., 173(2), pp. 247–259. doi: 10.1016/j.virusres.2013.01.021.
35. Sadykov, M. R. (2016) 'Restriction-modification systems as a barrier for genetic

manipulation of *Staphylococcus aureus*.', *Methods in Molecular Biology (Clifton, N.J.)*, 1373(1341), pp. 9–23. doi: 10.1007/7651_2014_180.

36. Söding, J., Biegert, A. and Lupas, A. N. (2005) 'The HHpred interactive server for protein homology detection and structure prediction.', *Nucleic Acids Research*, 33(Web Server issue), pp. W244-248. doi: 10.1093/nar/gki408.
37. Song, J., Xia, F., Jiang, H., Li, X., Hu, L., Gong, P., Lei, L., Feng, X., Sun, C., Gu, J. and Han, W. (2016) 'Identification and characterization of HolGH15: The holin of *Staphylococcus aureus* bacteriophage GH15', *Journal of General Virology*, 97(5), pp. 1272–1281. doi: 10.1099/jgv.0.000428.
38. Takeuchi, I., Osada, K., Azam, A. H., Asakawa, H., Miyanaga, K. and Tanji, Y. (2016) 'The presence of two receptor-binding proteins contributes to the wide host range of staphylococcal *Twort-like* phages.', *Applied and Environmental Microbiology*, 82(19), pp. 5763–5774. doi: 10.1128/AEM.01385-16.
39. Vandersteegen, K., Mattheus, W., Ceyssens, P. J., Bilocq, F., de Vos, D., Pirnay, J. P., Noben, J. P., Merabishvili, M., Lipinska, U., Hermans, K. and Lavigne, R. (2011) 'Microbiological and molecular assessment of bacteriophage ISP for the control of *Staphylococcus aureus*', *PLoS One*, 6(9). p. e24418. doi: 10.1371/journal.pone.0024418.
40. Whitman, W. B., Coleman, D. C. and Wiebe, W. J. (1998) 'Prokaryotes: the unseen majority.', *Proceedings of the National Academy of Sciences of the United States of America*, 95(12), pp. 6578–6583. doi: 10.1073/pnas.95.12.6578.
41. Wyatt, H. D. M. and West, S. C. (2014) 'Holliday junction resolvases', *Cold Spring Harbor Perspectives in Biology*, 6(9), pp. 1–30. doi: 10.1101/cshperspect.a023192.
42. Xia, G., Corrigan, R. M., Winstel, V., Goerke, C., Gründling, A. and Peschel, A. (2011) 'Wall teichoic acid-dependent adsorption of staphylococcal siphovirus and myovirus', *Journal of Bacteriology*, 193(15), pp. 4006–4009. doi: 10.1128/JB.01412-10.

Chapter IV

Construction of a *Staphylococcus* phage endolysin secretion system in *Lactococcus* and evaluation of lytic activity of endolysin against *Staphylococcus aureus* in milk

4.1 Abstract

Staphylococcus aureus, often associated with bovine mastitis, frequently contaminates milk and many products derived from raw milk. Studies of bacteriophage-derived endolysins highlighting their ability to rapidly eliminate *S. aureus* have been reported in recent years. The aim of this study was to develop a secretion system using a *Lactococcus*-based host for delivery of recombinant phage endolysin as therapeutic model against *S. aureus* contamination in dairy product. This was constructed using the lactococcal cloning vector pNZ8048, and the resulting recombinant plasmid transformed into *L. lactis*. Endolysin secretion by the recombinant *L. lactis* strain was verified by SDS-PAGE and zymographic analysis with concentrated supernatant from an overnight recombinant culture. Purified $CHAP_k$ were also directly applied to milk contaminated with *S. aureus,* resulting in the elimination of this pathogen. These results suggest that the staphylococci endolysin $CHAP_k$ can be used in the development of a therapeutic model targeting *S. aureus* contamination in milk and milk-derived products.

4.2 Introduction

Staphylococcus aureus is a frequent contaminant in raw milk, and in this respect it may end up in high numbers in specific milk based products such as raw milk cheeses. The origin of these milk-contaminating staphylococci is usually in cases of sub-clinical mastitis in cattle herds (Barrett *et al.*, 2005), often leading to economic losses from reduced milk production, veterinary treatment, discarded milk and culling (Geary *et al.*, 2011). These contaminated milk and milk-based products usually result in food poisoning, especially when the affected dairy product ingested is contaminated with an enterotoxigenic staphylococcal strain (Hennekinne *et al.,* 2012). The onset of this illness is usually abrupt and symptoms may include hypersalivation, nausea, vomiting, and abdominal cramping with or without diarrhoea (Kadariya *et al.,* 2014). Vaughan and Sternberg first reported an incident involving staphylococcal food poisoning (SFP) in 1884 (Hennekinne *et al.*, 2012). Over the years numerous cases of SFP have been documented (Dolman, 1934; De Buyser *et al.*, 2001; Hennekinne *et al.,* 2012). These SFPs usually get resolved within 24 to 48hrs after onset but in some cases it can lead to severe illness requiring hospitalization (Mossong *et al.*, 2015).

The use of bacteriophages (phages), which are naturally occurring viruses capable of infecting and killing food spoilage pathogens, has several advantages over chemical preservatives (Greer, 2005; Hudson *et al.*, 2005). And the use of bacteriophage-encoded endolysins (lysins) has been implicated in the control of these food-borne pathogens (Schmelcher *et al.*, 2012; Zhang *et al.*, 2012). These enzymes are peptidoglycan hydrolases involved in the degradation of bacterial cells “from within” at the end of their replication cycle. Endolysins targeting Gram-positive bacteria usually possess two distinct functional domains; a cell wall binding domain, which confer specificity by recognising and binding to specific ligand within the cell wall, and a catalytic domain responsible for catalyzing the

breakdown of peptidoglycan (Schmelche *et al.,* 2012). When applied exogenously, these enzymes have been successfully used to eliminate *Listeria monocytogenes* in soya milk at refrigeration temperature (Zhang *et al.*, 2012). These enzymes have also been used in synergy with nisin to control *S. aureus* in pasteurized milk (García *et al.*, 2010). The specific nature of these enzymes makes them very attractive candidates in the control of food spoilage organism. Secretion of these enzymes in lactic acid bacteria (LAB), organisms with the ability to produce lactic acid during homo-fermentative or hetero-fermentative metabolism (Klaenhammer *et al.*, 2002), have been reported (Gaeng *et al.*, 2000; Turner *et al.*, 2007). LABs have useful applications in the dairy industry and *Lactococcus lactis* is one of the most frequently used organisms in dairy fermentation (Beresford *et al.*, 2001). Growth in milk requires the presence of specific plasmids in the lactococcal strains (Mills *et al.*, 2006; Gasson, 1983; Tarazanova *et al.*, 2016) and of these, pLP712 is responsible for the organism's ability to metabolise lactose and undergo casein proteolysis in milk (Gasson, 1983; Wegmann *et al.*, 2012).

The endolysin $CHAP_k$ (truncated derivative of LysK) from the *S. aureus* bacteriophage K has been previously isolated and characterised (Horgan *et al.*, 2009; Fenton *et al.*, 2010). This 18.6-kDa protein is a cysteine histidine-dependent amidohydrolase/peptidase (CHAP) capable of eliminating staphylococcal cells including those embedded in biofilms when applied exogenously (Fenton *et al.*, 2013). This enzyme could also eliminate *S. aureus* in the nares of mice models (Fenton *et al.*, 2010). The crystal structure of this enzyme has also been resolved (Sanz-Gaitero *et al.*, 2013) and the enzyme possesses a conserved proteolytic triad of Cys54, His119 and Glu134 aminoacid residue, which are necessary for $CHAP_k$'s activity. $CHAP_k$ has also been exploited in the development of a wound dressing incorporating a nanoparticle-based thermal trigger release system for controlled release of antimicrobials

including endolysins into infected wounds for the control of *S. aureus* (Hathaway *et al.*, 2017).

A secretion vector encoding the $CHAP_k$ endolysin was inserted into an *L. lactis* dairy starter culture with the aim of developing a therapeutic model for controlling *S. aureus* contamination in milk. Accordingly, the secretion vector was constructed and $CHAP_k$ was cloned and expressed under the control of a *Lactobacillus slpA* promoter and signal peptide to drive the secretion of active $CHAP_k$ from *L. lactis* cells.

4.3 Materials and Methods

4.3.1 Bacterial strains and growth conditions

Strains and plasmids used in this study are listed in Table 1. *Escherichia coli* strains were grown on Luria-Bertani medium (Sigma-Aldrich, St. Louis, MO, USA), with shaking at 37ºC. Both *Lactococcus lactis* NZ9000 and NCDO712 strains were grown on M17 medium (Sigma-Aldrich, USA) supplemented with 0.5% (w/v) glucose at 30ºC, without shaking. *Lactobacillus brevis* DSM 20556 was grown on MRS medium (Sigma-Aldrich, USA) at 37ºC and without shaking. *Staphylococcus aureus* DPC5246 was grown on Brain Heart Infusion (Sigma-Aldrich, USA) with shaking at 37ºC, unless otherwise stated. All strains were either grown on liquid medium or supplemented with 1.5% (w/v) agar.

4.3.2 DNA manipulations

Plasmid DNA was isolated from *E. coli* using the High pure plasmid isolation kit (Roche Applied Science, Germany). In the case of *L. lactis* and *Lb. brevis*, plasmid isolation required treatment of cells with protoplast buffer (20mM Tris-Hcl; 5mM EDTA; 0.75M sucrose; 10mg/ml lysozyme and 50units/ml of mutanolysin) at 37ºC for 30mins, before the use of the isolation kit (Roche Applied Science, Germany). Chromosomal DNA was isolated from *Lb.*

brevis using QIAamp DNA mini kit (Qiagen, Hilden, Germany). Restriction enzymes and other DNA-manipulation enzymes used in this study were all acquired from Roche and Merck Millipore, and were used according to the manufacturer's instructions. Analysis of nucleotides and amino acid sequences *in silico* was done using SnapGene bioinformatics software. Transformation into *E. coli* and *L. lactis* was carried out by electroporation using an Electroporator (Eppendorf, Germany). DNA sequencing services as well as synthesis of oligonucleotides for this study was done with Eurofins MWG (Germany).

Table 1. List of bacterial strains used in this study

Bacterial strain or plasmid	**Details**	**Source or reference**
Strains		
E. coli XL1-Blue	Cloning host: *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI*^{q}*ΔM15* Tn*10*(Tet^{r})]	Agilent Technologies
E. coli (pQE60-CHAP$_k$)	Staphylococcal phage lysin containing CHAP domain cloned into *E. coli* XL1-Blue	Horgan *et al.*, 2009
L. lactis NZ9000	MG1363 *pep*N::nisRK	Kuipers *et al.*, 1998
L. lactis NCDO712	*L. lactis* dairy isolate harbouring plasmids pLP712, pSH71, pSH72, pSH73, pSH74 and pNZ712	Gasson, 1983; Tarazanova *et al.*, 2016
L. brevis DSM 20556	Same as ATCC 8287	DSM
S. aureus DPC5246	Bovine *S. aureus*	O'Flaherty *et al.*, 2005
NCDO712 (pNZ8048-$_{SP}$*slpA*CHAP$_k$)	*L. lactis* NCDO712 strain secreting the staphylococcal phage lysin CHAP$_k$	This study
NZ9000 (pNZ8048-$_{SP}$*slpA*CHAP$_k$)	*L. lactis* NZ9000 strain secreting the staphylococcal phage lysin CHAP$_k$	This study

4.3.3 Construction of endolysin secretion vector and cloning in *L. lactis*

An endolysin secretion vector was created using the PCR technique "splicing by overlap extension" (SOE) PCR, which was previously described by Horton *et al* (1989). Briefly, the *slpA* (surface layer protein A) signal sequence and promoter sequence were amplified using *Lb-slpA*_F and *Lb-slpA*_R primers (Table 2), using chromosomal DNA from *L. brevis* as template. DNA encoding the staphylococcal endolysin $CHAP_k$ (Horgan *et al.*, 2009) was also amplified using the primers *Ec-CHAPk*_F and *Ec-CHAPk*_R (Table 2). The resulting two PCR fragments were spliced using the primers *Lb-slpA*_F and *Ec-CHAPk*_R (Table 2). The inserts generated were digested with *Bgl*II and *Nco*I restriction enzymes and ligated with pNZ8048 plasmid to generate the recombinant vector pNZ8048-$_{SP}$*slpA*$CHAP_k$. This vector was transformed by electroporation using an Eppendorf eporator into *L. lactis* NZ9000 and subsequently into NCDO712; a lactose-utilizing *L. lactis* strain.

Table 2. List of primers used in this study

Oligonucleotides	**Sequence (5'→3')**	**Details**
*Lb-slpA*_F	TTAAGATCTTTCAATCCAACGACAATCAGAG	Amplication of *slpA* promoter and leader sequence
*Lb-slpA*_R	TTAGCCATAGCTGAAGCAGTCGTTGAAA	Amplication of *slpA* promoter and leader sequence
*Ec-CHAPk*_F	CTTCAGCTATGGCTAAGACTCAAGCAGA	Amplication of $CHAP_k$
*Ec-CHAPk*_R	TTACCATGGCTATGCTTTTACAGGTATTTCAATG	Amplication of $CHAP_k$

Restriction sites are underlined

Table 3. List of plasmids used in this study

Plasmids	Details	Source or reference
pQE60	Cloning and expression vector, Amp^r	Qiagen
pQE60-$CHAP_k$	$CHAP_k$ (lysin of phage K origin) inserted into the *Nco*I/*Bgl*II site of pQE60	Horgan *et al.* 2009
pNZ8048	High-copy-number *E. coli-L. lactis* overexpression vector, Cm^r	De Ruyter et al. 1996
pNZ8048-$_{SP}$*slpA*$CHAP_k$	$_{SP}$*slpA*$CHAP_k$ fusion inserted into *Nco*I/*Bgl*II site of pNZ8048, *nis*A promoter replaced with *slpA* promoter (CHAP secretion vector)	This study

4.3.4 Analysis of secreted protein by SDS-PAGE and zymogram

An overnight culture of *L. lactis* NZ9000 carrying pNZ8048-$_{SP}$*slpA*$CHAP_K$ plasmid was harvested by centrifugation at 4ºC. The supernatant was filter sterilized, and the proteins concentrated by ultrafiltration (10kDa cutoff amicon filter; Merck Millipore, Darmstadt, Germany). The concentrated supernatant were then analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) assay using 12% polyacrylamide gel similar to the method described in Chapter III.

Zymogram assay was used detect the lytic activity of the secreted CHAP. Briefly, *S. aureus* cells were grown, autoclaved and used as cell substrates in 12% polyacrylamide gel. The protein samples were separated on the zymogram gel using Tris-Glycine buffer. The gel was soaked in distilled water at room temperature for 30mins and subsequently transferred into renaturing buffer (50mM Tris-HCl, pH 8.0; 1% (w/v) Triton-x) and shaken gently for 1hr. The protein sample containing lytic activity appeared as clear translucent bands on an opaque background.

4.3.5 Endolysin activity plate test

Endolysin secretion from *L. lactis* was assessed using an endolysin activity plate test according to the method described by Gaeng *et al* (2000). Briefly, heat-inactivated *S. aureus* cells were suspended in GM17 agar and *L. lactis* clones secreting $CHAP_k$ endolysin were then plated on the medium and incubated at 30ºC for 24hrs. After incubation, the agar medium was examined for the formation of clear halos surrounding the lactococcal clones.

4.3.6 Co-culture of *S. aureus* with *L. lactis* secreting $CHAP_k$

S. aureus DPC5246 and *L. lactis* NCDO712 secreting $CHAP_k$ were grown for 16hrs. These cells were harvested by centrifugation, washed several times in sterile ringers and resuspended to the same cell density in BHI broth and GM17 broth respectively. Cocultures were performed in 10% (w/v) skimmed milk containing 0.2M sodium phosphate buffer (pH 7) at 30ºC for 24hrs. These cocultures consisted of *L. lactis* NZ9000 (pNZ8048-$_{SP}$*slpA*$CHAP_k$) at 1 x 10^6 CFU/ml mixed with *S. aureus* DPC5246 (between 8 x 10^3 and 9 x 10^3 CFU/ml). To determine viable *Lactococcus*, dilutions were plated on GM17 agar containing chloramphenicol. Also, to determine viable *Staphylococcus*, dilutions were plated on Baird-Parker agar (Sigma-Aldrich, USA).

4.3.7 Expression and purification of $CHAP_k$

E. coli clones containing pQE60-$CHAP_k$ plasmid were grown in superbroth (3.2% (w/v) tryptone, 2% (w/v) yeast extract, 0.5% (w/v) NaCl) containing 200µg/ml of ampicillin to the mid-exponential phase of growth. Cells were induced with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Sigma-Aldrich, USA) for 14hrs at 26ºC. After induction, cells were harvested by centrifugation and then resuspended in 50mM sodium acetate. Cell lysis was achieved by using Bugbuster protein extraction reagent (Novagen) and the cellular debris was removed by centrifugation, together with filter sterilization of the supernatant through a

0.2µm filter. The sterilized cell lysate was then subject to purification. This was done by cation-exchange chromatography using the AKTA *Start* system (GE healthcare) to achieve > 90% homogeneity. The purified $CHAP_k$ was then quantified using the Bradford assay (Bradford, 1976).

4.3.8 *S. aureus* challenge with $CHAP_k$ in milk

Skim milk powder (Sigma-Aldrich, UK) was reconstituted in deionised water to a concentration of 10% (w/v) and sterilized. The sterilized milk was inoculated with approximately 10^6 CFU/ml of *S. aureus*, briefly shaken and $CHAP_k$ (745µg/ml) was added. Incubation at 30ºC was performed without shaking and samples were taken at hourly intervals for 3hrs. Survival of *S. aureus* was determined by plate decimal dilutions (in ringers) on Baird Parker selective plates (Sigma-Aldrich, USA), which were incubated at 37ºC for 24hrs. The *S. aureus* challenge with the $CHAP_k$ experiment was repeated in triplicate.

4.4 Results

4.4.1 Construction of CHAP expression and secretion system

To allow for a constitutive expression and secretion of the staphylococcal endolysin $CHAP_k$ (Horgan *et al.*, 2009), a secretion vector was constructed using the lactococcal plasmid pNZ8048 plasmid. This was achieved by amplifying a 253-bp fragment of *L. brevis* ATCC 8287's *slpA* promoter and leader sequence (Vidgren *et al.*, 1992), subsequently splicing it with a 495-bp fragment of staphylococcal lysin $CHAP_k$ (Horgan *et al.*, 2009) using the PCR technique "slicing by overlap extension PCR" (Horton *et al.*, 1989). The resulting 748-bp fragment was cloned into pNZ8048, resulting in a vector whose inducible *nis*A promoter (de Ruyter *et al.,* 1996) was replaced with an *slpA* promoter (Vidgren *et al.*, 1992), ensuring

constitutive expression of $CHAP_k$. The *slpA* leader sequence was also positioned upstream of the promoter (Figure 1) to ensure membrane translocation of $CHAP_k$. The resulting plasmid, designated pNZ8048-$_{sp}$*slpA*-$CHAP_{k,}$ was transformed directly into *L. lactis* NZ9000 as every attempt to use *E. coli* as an intermediate recipient resulted in deletions.

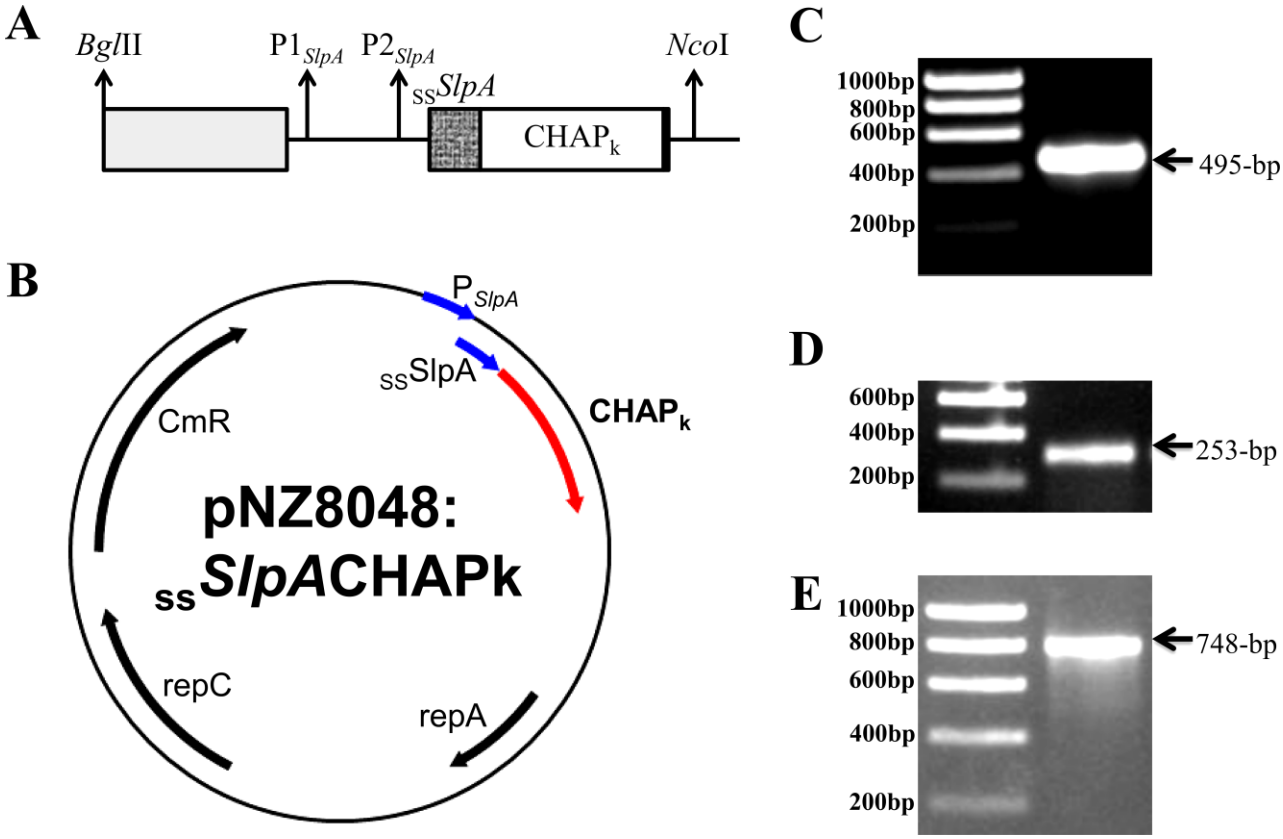


Figure 1. Schematic illustration of $CHAP_k$ expression construct (A) together with map illustrating the construction of the $CHAP_k$ secretion vector (B). The gel electrophoresis photo shows the splicing of a 495-bp DNA fragment of $CHAP_k$ (C) with a 253-bp DNA fragment of *L. brevis slpA* signal peptide and promoter (D) to create a 748-bp DNA fragment (E), which was used together with pNZ8048 plasmid in the creation of the $CHAP_k$ secretion vector.

4.4.2 *SlpA* signal peptide enables membrane translocation of active $CHAP_k$ lysin

Expression and secretion of $CHAP_k$ was detected using an SDS-PAGE assay on recombinant *L. lactis*. Supernatant from the *L. lactis* strain carrying pNZ8048-$_{sp}$*slpA*-$CHAP_k$ plasmid and cell extract from the recombinant *E. coli* carrying pQE60-$CHAP_k$ plasmid were both tested. The cell extract recovered from the recombinant *E. coli* was used as a positive control. For

the negative control, supernatant from *L. lactis* (pNZ8048) was used. A 10kDa amicon filter was used to concentrate the spent supernatant to ensure detection of secreted protein on the SDS gel. The presence of a band on *L. lactis* (pNZ8048-$_{sp}$*slpA*-CHAP$_k$) at the same position as the positive control (Figure 2A) demonstrated CHAP$_k$ secretion by the recombinant *L. lactis*. To demonstrate that the secreted CHAP$_k$ was still active, a zymogramic assay was performed using heat-inactivated *S. aureus* DPC5246 cells. Concentrated spent supernatant from pNZ8048-$_{sp}$*slpA*-CHAP$_k$ exhibited clearing on the zymogram gel (Figure 2B) thus confirming the CHAP$_k$ secreted by the recombinant *L. lactis* was indeed active.

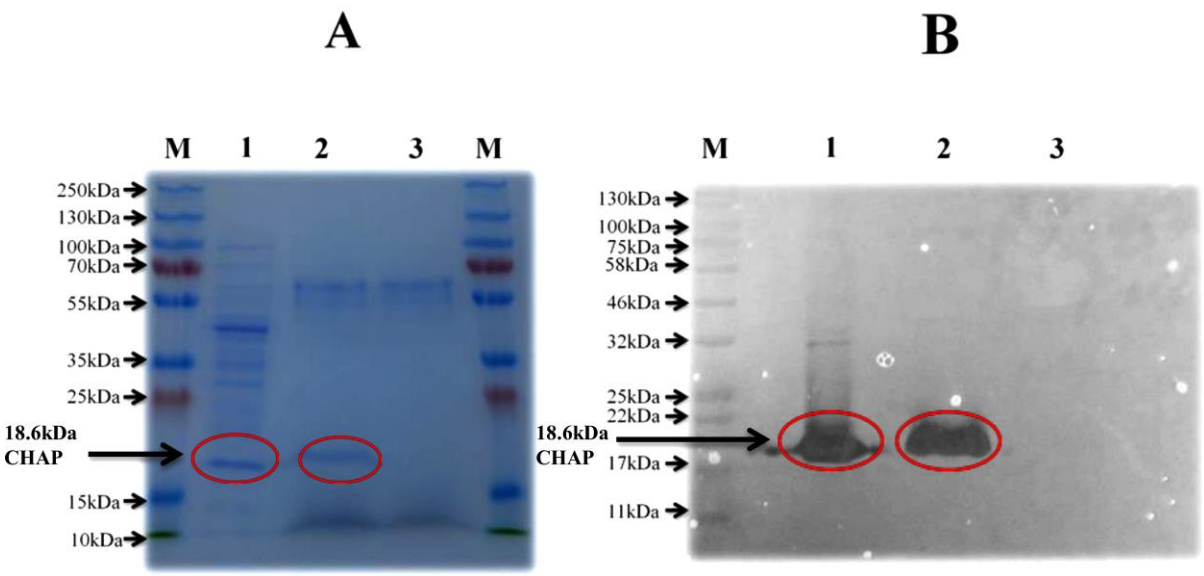


Figure 2. Detection of active CHAP$_k$ (18.6kDa) secreted by recombinant *L. lactis* using SDS-PAGE (A) and zymographic analysis with heat-inactivated *S. aureus* as cell substrates (B). Lane M: Pre-stained protein markers (PageRuler prestained protein ladder from ThermoFisher Scientific for A and Blue prestained protein standard from New England BioLabs for B); Lane 1: Cell lysate from *E. coli* containing pQE60-CHAP$_k$ as positive control (both gels); Lane 2: Concentrated spent supernatant from *L. lactis* NZ9000 containing CHAP secretion vector (both gels); Lane 3: Concentrated spent supernatant from *L. lactis* NZ9000 containing pNZ8048 as negative control (both gels). Band positioned at 18.6kDa in SDS gel (lane 2) demonstrates secretion of CHAP, while clearing in zymogram shows the secreted CHAP is active.

4.4.3 CHAP secretion in lactose utilising *L. lactis* NCDO712 strain

For growth in milk, lactose utilization is a key property of lactic acid bacteria used in the dairy industry. For this reason, *L. lactis* NCDO712, a strain originally isolated from a dairy starter culture in 1983 (Gasson, 1983) was used as the host for the pNZ8048-$_{sp}$*slpA*-CHAP$_k$ construct. As with the plasmid-free laboratory strain NZ9000, recombinant NCDO712 clones containing pNZ8048-$_{sp}$*slpA*-CHAP$_k$ were capable of exporting active CHAP$_k$ into their surrounding environment. This was demonstrated using an endolysin activity plate test, where secretion of CHAP$_k$ from NCDO712 (pNZ8048-$_{sp}$*slpA*-CHAP$_k$) was detected as clearing on turbid GM17 agar embedded with heat-inactivated staphylococcal cells (Fig. 3B). No clearing was detected from the control strain (Figure 3A).

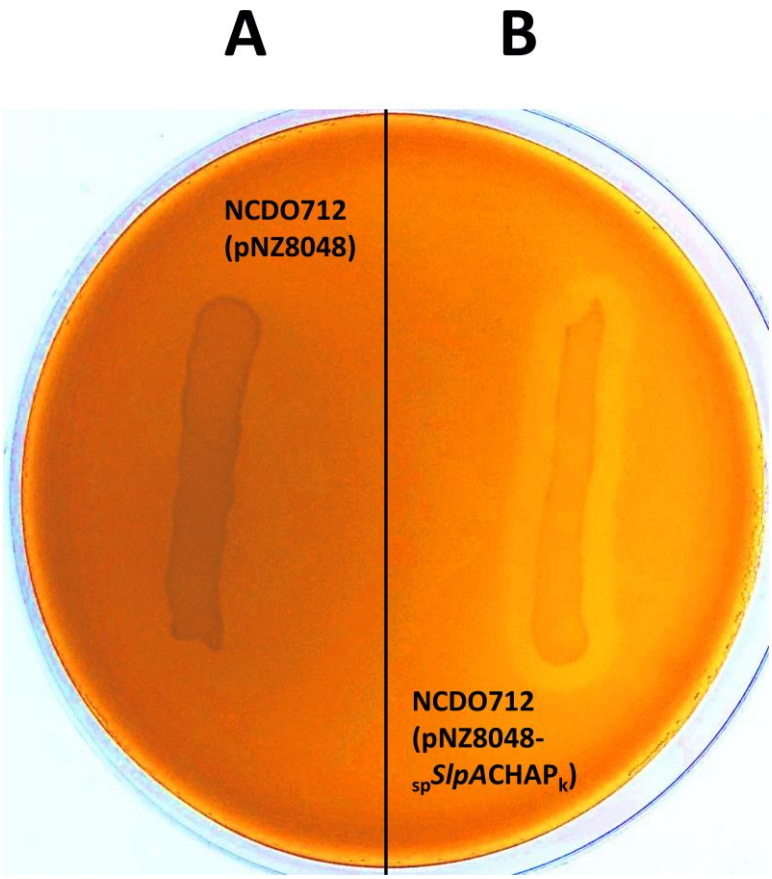


Figure 3. Recombinant *L. lactis* grown on GM17 agar medium embedded with heat-inactivated *S. aureus* cells. The control strain NCDO712 (pNZ8048) shows no effect (A), whereas NCDO712 (pNZ8048-$_{sp}$*slpA*-CHAP$_k$) (B) secreted active CHAP$_k$ endolysin showing clear zones of lysis around the *L. lactis* streak.

4.4.4 *S. aureus* survival in milk in the presence of *L. lactis* NCDO712-secreting-$CHAP_k$

The antimicrobial activity of *L. lactis* NCDO712 (pNZ8048-$_{sp}$*slpA*-$CHAP_k$) was assessed in co-culture with *S. aureus* DPC5246 in milk over a 24-hr period. The milk was buffered with a phosphate buffer as earlier attempts at co-culture resulted in *S. aureus* inhibition due to production of lactic acid by the *L. lactis* strain. The pH of the milk, over the 24hr period after inoculating the unbuffered milk with *L. lactis* NCDO712 (pNZ8048-$_{sp}$*slpA*-$CHAP_k$) was recorded as 4.25. *S. aureus* and *L. lactis* levels in milk were initially recorded as 10^3 CFU/ml and 10^6 CFU/ml respectively. However, no inhibition of growth was encountered in the *S. aureus* levels, with the titre increasing from $8.3x10^3$ to $8.2x10^7$ CFU/ml over the 24-hr period (Table 4).

Table 4. Survival of *S. aureus* DPC5246 in co-culture with $CHAP_k$-secreting *L. lactis* in a milk environment

Time (hrs)	*S. aureus* count (CFU/ml)		*L. Lactis* count (CFU/ml)	
	pNZ8048	pNZ8048 ($CHAP_k$)	pNZ8048	pNZ8048 ($CHAP_k$)
0	$8.9x10^3$	$8.3x10^3$	$1.27x10^6$	$1.28x10^6$
6	$1.92x10^5$	$4.2x10^5$	$1.12x10^8$	$1.23x10^8$
24	$1.41x10^8$	$8.2x10^7$	$1.23x10^9$	$1.10x10^9$

4.4.5 *S. aureus* challenge in milk containing purified $CHAP_k$ endolysin

Owing to the inability of $CHAP_k$-secreting *L. lactis* NCDO712 in reducing *S. aureus* numbers in milk, the effect of purified $CHAP_k$ in milk spiked with *S. aureus* at a contamination level of 10^6 CFU/ml was subsequently investigated. This was done to determine if $CHAP_k$ was capable of eliminating *S. aureus* in the milk matrix. Addition of $CHAP_k$ resulted in a 1.5 log reduction of the *S. aureus* titre in the first hour when incubated at 30ºC, after-which the titre

levelled out. When the experiment was repeated with a higher dosage, the activity still disappeared after the first hour (Figure 4).

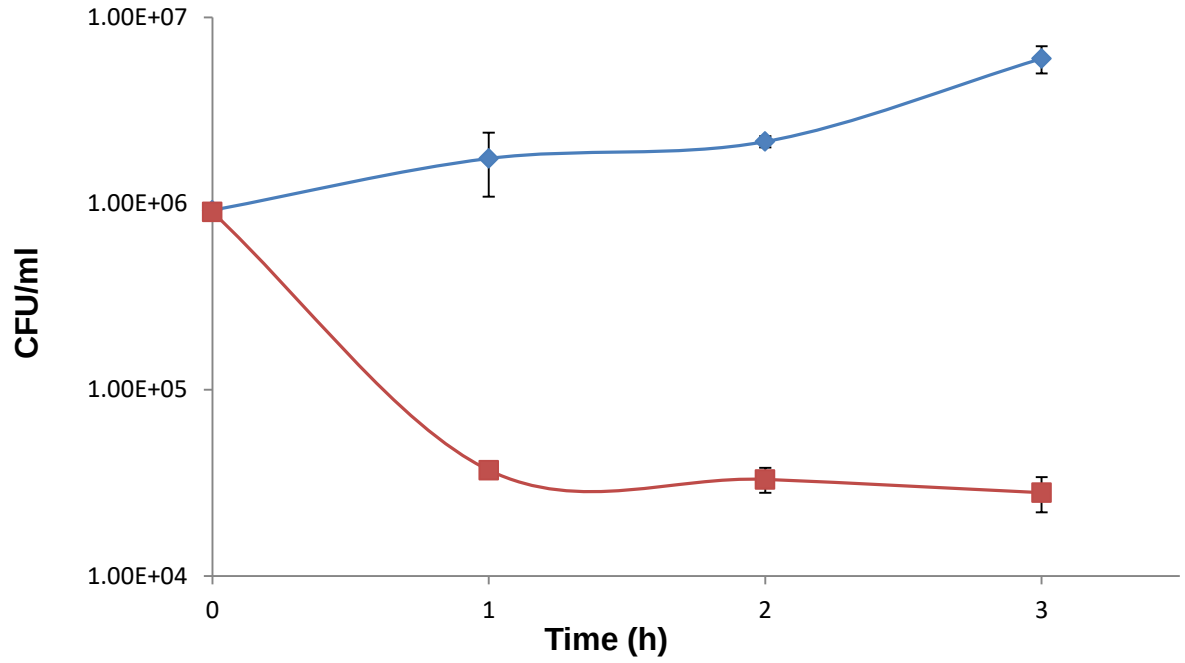


Figure 4: Killing of *S. aureus* DPC5246 with purified $CHAP_k$ in reconstituted skimmed milk; ◆ *S. aureus* plus 50mM sodium acetate (control); ■ *S. aureus* plus $CHAP_k$ (745µg/ml). Values are the averages of three independent experiments with standard error indicated by vertical bars.

4.5 Discussion

In this study, the staphylococcal endolysin $CHAP_k$ was cloned into the food-grade *L. lactis* bacterium in order to achieve secretion of this lytic enzyme into its surrounding environment. It was previously shown that $CHAP_k$, the deleted derivative of the staphylococcal phage K endolysin LysK, possesses as much lytic activity against *S. aureus* cells as LysK (contains CHAP and amidase enzymatic domains and also a cell-wall-binding domain) (Horgan *et al.*, 2009). In addition, the significantly smaller size of $CHAP_k$ and the lack of a cell-wall binding domain rendered it a more straightforward protein to work with. For this reason, $CHAP_k$ was used in the construction of an *L. lactis* delivery system that could serve as a therapeutic

model against *S. aureus* contamination in milk. Secretion of $CHAP_k$ was driven with the aid of the *Lactobacillus* expression and secretion signal "*slpA*", which normally drives the expression and secretion of surface (S)-layer proteins in *Lactobacillus brevis*. These secretion signals have previously been shown to be compactible for secreting proteins such as β-lactamase, (Savijoki *et al.,* 1997), β-glucuronidase, luciferase (Kahala and Palva, 1999) and *Listeria* endolysin (Gaeng *et al.,* 2000) in the literature. For these reasons, the *slpA* expression and secretion signals (promoter and signal peptide) and the lactococcal plasmid pNZ8048 (Kuipers *et al.*, 1998) were employed for the construction of the $CHAP_k$ secretion vector (pNZ8048-$_{SP}$*slpA*$CHAP_k$) (Figure 1).

Several attempts to introduce the $CHAP_k$ secretion vector into *E. coli* resulted in deletions within the *slpA* secretion sequence and/or the sequence encoding $CHAP_k$, effectively preventing the expression and secretion of intact endolysin. As a result, direct transformation into *L. lactis* was chosen as the better approach. Cytoplasmic production of $CHAP_k$ without a secretion signal has previously been shown to be possible in *E. coli* (Horgan *et al.*, 2009), suggesting that the deletions may have resulted from difficulties with membrane translocation and proteolytic processing of $CHAP_k$ by *E. coli*. Although $CHAP_k$ does not result in lysis of *E. coli* cells, there is a possibility that export of this enzyme across the *E. coli* cell wall may be detrimental to functions vital for its cell growth and division. A similar finding and hypothesis were made by Gaeng *et al* (2000) in their attempt to secrete the *Listeria* endolysin Ply118 in a lactococcal host.

L. lactis uses the Sec pathway for secretion of proteins bearing an N-terminal signal peptide (Morello *et al.*, 2007). This involves recognition of the precursor protein thereby targeting it to the membrane translocation machinery, subsequently leading to the translocation of the precursor protein across the cytoplasmic membrane. The late stages of secretion in *L. lactis*

involve cleavage of signal peptide by the leader peptidase enzyme as well as release and folding of the mature protein. The enzyme $CHAP_k$ was successfully secreted in this manner and its export by *L. lactis* to its extracellular environment was detected by SDS-PAGE assay (Figure 2A). Enzymatic activity of the secreted $CHAP_k$ was verified by zymographic analysis.

The $CHAP_k$ secretion vector was also introduced into an industrial strain of *L. lactis* (NCDO712) possessing essential genes for lactose catabolism and casein hydrolysis encoded within its native plasmid content, thereby allowing it to grow in milk (Wegmann *et al.*, 2012). The resulting *L. lactis* NCDO712 (pNZ8048-$_{SP}$*slpA*$CHAP_k$) model was capable of exporting active $CHAP_k$ to its external environment as detectable CHAP was verified by the endolysin activity plate assay (Figure 3). The ability of this model secretion system to reduce *S. aureus* in milk was also investigated. The NCDO712 (pNZ8048-$_{SP}$*slpA*$CHAP_k$) model was however unable to cause any significant reduction to the *S. aureus* cells that had been inoculated into milk due to insufficient secretion of CHAP by the model secretion system.

Owing to the inability of NCDO712 (pNZ8048-$_{SP}$*slpA*$CHAP_k$) model to reduce *S. aureus* in milk, purified $CHAP_k$ was investigated as an alternative model additive to eliminate *S. aureus* in milk. $CHAP_k$ was indeed effective at eliminating *S. aureus* in milk (Figure 4).

Conclusion

The data shown here demonstrates the potential for staphylococcal phage endolysin to be used as a therapeutical model in the elimination of staphylococci in milk. Furthermore, cloning of pNZ8048-$_{SP}$*slpA*CHAP$_k$ into wildtype *L. lactis*, commercially used as starter culture demonstrates a potential application in food. However, the antibiotic resistant vetor used in this study would have to be replaced with a food-grade vector (Dickely *et al.*, 1995; Froseth and McKay, 1991; MacCormick *et al.,* 1995; Ross *et al.,* 1990) for successful food application. The secretion model developed in this study could successfully secrete detectable quantities of CHAP$_k$ into its surround environment. Purified CHAP$_k$ was also successfully used in the elimination of *S. aureus* pathogen in milk. These results suggest that the staphylococci endolysin CHAP$_k$ is a promising enzyme in the development of a therapeutic model that could elimate *S. aureus* contamination in milk and milk-derived products.

4.7 References

1. Barrett, D. J., Healy, A. M., Leonard, F. C. and Doherty, M. L. (2005) 'Prevalence of pathogens causing subclinical mastitis in 15 dairy herds in the Republic of Ireland', *Irish Veterinary Journal*, 58(6), pp. 333–337. doi: 10.1186/2046-0481-58-6-333.
2. Becker, S. C., Foster-Frey, J. and Donovan, D. M. (2008) 'The phage K lytic enzyme LysK and lysostaphin act synergistically to kill MRSA', *FEMS Microbiology Letters*, 287(2), pp. 185–191. doi: 10.1111/j.1574-6968.2008.01308.x.
3. Beresford, T. P., Fitzsimons, N. A., Brennan, N. L. and Cogan, T. M. (2001) 'Recent advances in cheese microbiology', *International Dairy Journal*, 11(4–7), pp. 259–274. doi: 10.1016/S0958-6946(01)00056-5.
4. Bradford, M. M. (1976) 'A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding', *Analytical Biochemistry*, 72(1–2), pp. 248–254. doi: 10.1016/0003-2697(76)90527-3.
5. De Buyser, M. L., Dufour, B., Maire, M. and Lafarge, V. (2001) 'Implication of milk and milk products in food-borne diseases in France and in different industrialised countries', *International Journal of Food Microbiology*, 67(1–2), pp. 1–17. doi: 10.1016/S0168-1605(01)00443-3.
6. Dickely, F., Nilsson, D., Hansen, E. B. and Johansen, E. (1995) 'Isolation of *Lactococcus lactis* nonsense suppressors and construction of a food-grade cloning vector', *Molecular Microbiology*, 15(5), pp. 839–847. doi: 10.1111/j.1365-2958.1995.tb02354.x.
7. Dolman, E. (1934) 'Ingestion of *Staphylococcus* exotoxin by human volunteers: With special reference to staphylococcic food poisoning', *The Journal of Infectious Diseases*, 55(2), pp. 172–183.
8. Fenton, M., Casey, P. G., Hill, C., Gahan, C. G., Ross, R. P., McAuliffe, O., O'Mahony, J., Maher, F. and Coffey, A. (2010) 'The truncated phage lysin CHAP(k) eliminates

Staphylococcus aureus in the nares of mice.', *Bioengineered Bugs*, 1(6), pp. 404–407. doi: 10.4161/bbug.1.6.13422.

9. Fenton, M., Keary, R., McAuliffe, O., Ross, R. P., O'Mahony, J. and Coffey, A. (2013) 'Bacteriophage-derived peptidase CHAP(K) eliminates and prevents staphylococcal biofilms.', *International Journal of Microbiology*, 2013, p. 625341. doi: 10.1155/2013/625341.
10. Froseth, B. R. and McKay, L. L. (1991) 'Development and application of pFM011 as a possible food-grade cloning vector', *Journal of Dairy Science*, 74(5), pp. 1445–1453. doi: 10.3168/jds.S0022-0302(91)78302-1.
11. Gaeng, S., Scherer, S., Neve, H. and Loessner, M. J. (2000) 'Gene cloning and expression and secretion of *Listeria monocytogenes* bacteriophage-lytic enzymes in *Lactococcus lactis*.', *Applied and Environmental Microbiology*, 66(7), pp. 2951–2958. doi: 10.1128/AEM.66.7.2951-2958.2000.
12. García, P., Martínez, B., Rodríguez, L. and Rodríguez, A. (2010) 'Synergy between the phage endolysin LysH5 and nisin to kill *Staphylococcus aureus* in pasteurized milk', *International Journal of Food Microbiology*. Elsevier B.V., 141(3), pp. 151–155. doi: 10.1016/j.ijfoodmicro.2010.04.029.
13. Gasson, M. J. (1983) 'Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing', *Journal of Bacteriology*, 154(1), pp. 1–9. doi: 10.1.1.625.2641.
14. Geary, U., Begley, N., McCoy, F., O'Brien, B., O'Grady, L. and Shalloo, L. (2011) 'Estimating the impact of mastitis on the profitability of Irish dairy farms', in *Udder Health and Communication*. Wageningen: Wageningen Academic Publishers, pp. 221–228. doi: 10.3920/978-90-8686-742-4_36.
15. Greer, G. G. (2005) 'Bacteriophage control of foodborne bacteria.', *Journal of Food*

Protection, 68(5), pp. 1102–1111. doi: 10.4315/0362-028X-68.5.1102.

16. Hathaway, H., Ajuebor, J., Stephens, L., Coffey, A., Potter, U., Sutton, J. M. and Jenkins, A. T. A. (2017) 'Thermally triggered release of the bacteriophage endolysin $CHAP_K$ and the bacteriocin lysostaphin for the control of methicillin resistant *Staphylococcus aureus* (MRSA)', *Journal of Controlled Release*. 245, pp. 108–115. doi: 10.1016/j.jconrel.2016.11.030.
17. Hennekinne, J. A., De Buyser, M. L. and Dragacci, S. (2012) '*Staphylococcus aureus* and its food poisoning toxins: Characterization and outbreak investigation', *FEMS Microbiology Reviews*, 36(4), pp. 815–836. doi: 10.1111/j.1574-6976.2011.00311.x.
18. Horgan, M., O'Flynn, G., Garry, J., Cooney, J., Coffey, A., Fitzgerald, G. F., Paul Ross, R. and McAuliffe, O. (2009) 'Phage lysin LysK can be truncated to its CHAP domain and retain lytic activity against live antibiotic-resistant staphylococci', *Applied and Environmental Microbiology*, 75(3), pp. 872–874. doi: 10.1128/AEM.01831-08.
19. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. and Pease, L. R. (1989) 'Engineering hybrid genes without the use of restriction enzymes: Gene splicing by overlap extension', *Gene*, 77, pp. 61–68. doi: 10.1016/0378-1119(89)90359-4.
20. Hudson, J. A, Billington, C., Carey-Smith, G. and Greening, G. (2005) 'Bacteriophages as biocontrol agents in food', *Journal of Food Protection*, 68(2), pp. 426–437. doi: 10.4315/0362-028X-68.2.426.
21. Kadariya, J., Smith, T. C. and Thapaliya, D. (2014) '*Staphylococcus aureus* and staphylococcal food-borne disease: An ongoing challenge in public health', *BioMed Research International*. 2014. doi: 10.1155/2014/827965.
22. Kahala, M. and Palva, A (1999) 'The expression signals of the *Lactobacillus brevis slpA* gene direct efficient heterologous protein production in lactic acid bacteria', *Applied Microbiology*, 57, pp. 71–78.

23. Klaenhammer, T., Altermann, E., Arigoni, F., Bolotin, A., Breidt, F., Broadbent, J., Cano, R., Chaillou, S., Deutscher, J., Gasson, M., van de Guchte, M., Guzzo, J., Hartke, A., Hawkins, T., Hols, P., Hutkins, R., Kleerebezem, M., Kok, J., Kuipers, O., Lubbers, M., Maguin, E., McKay, L., Mills, D., Nauta, A., Overbeek, R., Pel, H., Pridmore, D., Saier, M., van Sinderen, D., Sorokin, A., Steele, J., O'Sullivan, D., de Vos, W., Weimer, B., Zagorec, M. and Siezen, R. (2002) 'Discovering lactic acid bacteria by genomics', *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 82(1–4), pp. 29–58. doi: 10.1023/A:1020638309912.

24. Kuipers, O. P., De Ruyter, P. G. G. A., Kleerebezem, M. and De Vos, W. M. (1998) 'Quorum sensing-controlled gene expression in lactic acid bacteria', *Journal of Biotechnology*, 64(1), pp. 15–21. doi: 10.1016/S0168-1656(98)00100-X.

25. Maccormick, C. A., Griffin, H. G. and Gasson, M. J. (1995) 'Construction of a food-grade host/vector system for *Lactococcus*-based on the lactose operon', *FEMS Microbiology Letters*. 127, pp. 105–109.

26. Mills, S., McAuliffe, O. E., Coffey, A., Fitzgerald, G. F. and Ross, R. P. (2006) 'Plasmids of lactococci-genetic accessories or genetic necessities?', *FEMS Microbiology Reviews*, 30(2), pp. 243–273. doi: 10.1111/j.1574-6976.2005.00011.x.

27. Morello, E., Bermúdez-Humarán, L. G., Llull, D., Solé, V., Miraglio, N., Langella, P. and Poquet, I. (2007) '*Lactococcus lactis,* an efficient cell factory for recombinant protein production and secretion', *Journal of Molecular Microbiology and Biotechnology*, 14(1–3), pp. 48–58. doi: 10.1159/000106082.

28. Mossong, J., Decruyenaere, F., Moris, G., Ragimbeau, C., Olinger, C. M., Johler, S., Perrin, M., Hau, P. and Weicherding, P. (2015) 'Investigation of a staphylococcal food poisoning outbreak combining case-control, traditional typing and whole genome sequencing methods, Luxembourg, June 2014', *Eurosurveillance*, 20(45), pp. 1–7. doi: 10.2807/1560-

7917.ES.2015.20.45.30059.

29. O'Flaherty, S., Ross, R. P., Meaney, W., Fitzgerald, G. F., Elbreki, M. F. and Coffey, A. (2005) 'Potential of the polyvalent anti-*Staphylococcus* bacteriophage K for control of antibiotic-resistant staphylococci from hospitals', *Applied and Environmental Microbiology*, 71(4), pp. 1836–1842. doi: 10.1128/AEM.71.4.1836-1842.2005.
30. Ross, P., O'Gara, F. and Condon, S. (1990) 'Thymidylate synthase gene from *Lactococcus lactis* as a genetic marker: An alternative to antibiotic resistance genes', *Applied and Environmental Microbiology*, 56(7), pp. 2164–2169.
31. de Ruyter, P. G., Kuipers, O. P. and de Vos, W. M. (1996) 'Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin.', *Applied and Environmental Microbiology*, 62(10), pp. 3662–3667. doi: 0099-2240/96/$04.00 1 0.
32. Sanz-Gaitero, M., Keary, R., Garcia-Doval, C., Coffey, A. and van Raaij, M. J. (2013) 'Crystallization of the CHAP domain of the endolysin from *Staphylococcus aureus* bacteriophage K', *Acta Crystallographica Section F Structural Biology and Crystallization Communications*, 69(Pt 12), pp. 1393–1396. doi: 10.1107/S1744309113030133.
33. Savijoki, K., Kahala, M. and Palva, A. (1997) 'High level heterologous protein production in *Lactococcus* and *Lactobacillus* using a new secretion system based on the *Lactobacillus brevis* S-layer signals', *Gene*, 186, pp. 255–262. doi: 10.1016/S0378-1119(96)00717-2.
34. Schmelcher, M., Donovan, D. M. and Loessner, M. J. (2012) 'Bacteriophage endolysins as novel antimicrobials', *Future Microbiology*, 7(10), pp. 1147–1171. doi: 10.2217/fmb.12.97.
35. Schmelcher, M., Powell, A. M., Becker, S. C., Camp, M. J. and Donovan, D. M. (2012) 'Chimeric phage lysins act synergistically with lysostaphin to kill mastitis-causing *Staphylococcus aureus* in murine mammary glands', *Applied and Environmental Microbiology*, 78(7), pp. 2297–2305. doi: 10.1128/AEM.07050-11.
36. Tarazanova, M., Beerthuyzen, M., Siezen, R., Fernandez-Gutierrez, M. M., De Jong, A., Van

Der Meulen, S., Kok, J. and Bachmann, H. (2016) 'Plasmid complement of *Lactococcus lactis* NCDO712 reveals a novel pilus gene cluster', *PLoS One*, 11(12), pp. 1–21. doi: 10.1371/journal.pone.0167970.

37. Turner, M. S., Waldherr, F., Loessner, M. J. and Giffard, P. M. (2007) 'Antimicrobial activity of lysostaphin and a *Listeria monocytogenes* bacteriophage endolysin produced and secreted by lactic acid bacteria', *Systematic and Applied Microbiology*, 30(1), pp. 58–67. doi: 10.1016/j.syapm.2006.01.013.
38. Vidgren, G., Palva, I., Pakkanen, R., Lounatmaa, K. and Palva, A. (1992) 'S-layer protein gene of *Lactobacillus brevis*: Cloning by polymerase chain reaction and determination of the nucleotide sequence', *Journal of Bacteriology*, 174(22), pp. 7419–7427. doi: 10.1128/JB.174.22.7419-7427.1992.
39. Wegmann, U., Overweg, K., Jeanson, S., Gasson, M. and Shearman, C. (2012) 'Molecular characterization and structural instability of the industrially important composite metabolic plasmid pLP712', *Microbiology (United Kingdom)*, 158(12), pp. 2936–2945. doi: 10.1099/mic.0.062554-0.
40. Zhang, H., Bao, H., Billington, C., Hudson, J. A. and Wang, R. (2012) 'Isolation and lytic activity of the *Listeria* bacteriophage endolysin LysZ5 against *Listeria monocytogenes* in soya milk', *Food Microbiology*. Elsevier Ltd, 31(1), pp. 133–136. doi: 10.1016/j.fm.2012.01.005.

Chapter V

Engineering a secretion system for delivery of a *Clostridium difficile* phage endolysin in Gram-positive bacteria (*Lactococcus lactis*)

5.1 Abstract

Clostridium difficile is a Gram-positive spore-forming bacterium and is an aetiological cause of antibiotic-associated pseudomembranous colitis in humans. Its pathogenicity is attributed to the production of exotoxins, which have a cytotoxic effect on mammalian cells. Owing to the emergence of hypervirulent strains with increased resistance to antibiotics, new control strategies are needed. Endolysins have useful potential in the elimination of antibiotic resistant infectious bacteria, and this has been demonstrated in the literature. In this work, the *C. difficile* endolysin gene *lyscd* was cloned and expressed in *E. coli* for intracellular production. Furthermore, a model secretion system for this endolysin was developed in *L. lactis* based on the pNZ8048 plasmid including a *L. brevis* surface layer protein A (*slpA*) expression and secretion signal. Successful secretion of this endolysin was detected using an endolysin plate assay, which involved streaking the recombinant *L. lactis* strain on media seeded with heat-inactivated *C. difficile* cells. The effective secretion of *C. difficile* endolysin in the *Lactococcus* host suggests that it may be further applied in other lactic acid bacteria with the aid of delivering it to members of this group that inhabit the gastrointestinal tract.

5.2 Introduction

Clostridium difficile is a Gram-positive, spore-forming, anaerobic rod-shaped bacterium known to produce clinically significant toxins, which contributes to its pathogenicity. These toxins, designated toxin A and B, are encoded on a pathogenicity locus (Viswanathan *et al.,* 2010) and have cytotoxic effects on mammalian cells, subsequently causing diarrhoea and colonic inflammation (Monaghan *et al.,* 2009). They also disrupt the epithelial cell layer of the colon, resulting in an inflammatory response that contributes to the disease pathology, with symptoms ranging from mild diarrhoea to chronic pseudomembraneous colitis (Hargreaves and Clokie, 2014). *C. difficile* infection (CDI) mostly occurs following the use of broad spectrum antibiotics leading to the disruption of the colon's normal flora (Predrag, 2016). As such, without the colonization resistance brought about by these commensal flora, opportunistic *C. difficile* are thus left to multiply and eventually colonize the colon (Viswanathan *et al.*, 2010).

Antimicrobials with the ability to provide protection against *C. difficile* without causing any collateral damage to commensal bacteria in the colon are beneficial. The specific nature of endolysins makes them an attractive candidate for combating *C. difficile* without such damage (Mayer *et al.,* 2008). These enzymes typically consist of two domains: a peptidoglycan hydrolase domain involved in enzymatic degradation of bacterial cell (usually on the N-terminal) and a C-terminal cell wall binding domain responsible for binding to bacterial cell wall substrates (Fischetti, 2008). The use of phage encoded enzymes as antimicrobial therapy in eliminating infectious bacteria has effectively been demonstrated in literature (Jado *et al.*, 2003; Fenton *et al.*, 2010; Jun *et al.*, 2013; Díez-Martínez *et al.*, 2015; Wang *et al.*, 2015). To date, no bacterial resistance to these enzymes has been encountered, even with studies exposing these enzymes to events that would normally result in bacterial resistance (Schuch *et al.,* 2002; Rodríguez-Rubio *et al.*, 2013).

Lactic acid bacteria (LAB) are food grade organisms with the potential of delivering proteins of interest to sites in the digestive tract. *L. lactis*, a member of this group of bacteria is one of the most studied species and is considered a model lactic acid bacterium (Van Hylckama Vlieg *et al.*, 2006). They also possess GRAS (Generally regarded as safe) status and are involved in the production of certain metabolite and toxins with inhibitory effect on bacterial pathogens (Rolfe, 1984*;* Viswanathan *et al.,* 2010). Few *L. lactis* strains like *L. lactis* subsp. *lactis* CV56 (Gao *et al.*, 2011) and *L. lactis* subsp. *lactis* KLDS4.0325 (Yang *et al.*, 2014) have probiotic properties. These organism are well characterised lactic acid bacteria and are promising candidates for delivering heterologous proteins to the digestive tract (Le Loir *et al.,* 2005). Genetic tools for bio-engineering of this organism are widely available and its entire genome has also been sequenced (Le Loir *et al.*, 2005).

In this study, the amidase endolysin gene *lyscd* from the genome of the *C. difficile* bacteriophage phiCD6356 was cloned and expressed in *E. coli* for intracellular production. The sequence of the enzymatic domain of the endolysin (designated LysCD) was similar to the previously characterised prophage endolysin PlyCD (Wang *et al.*, 2015) and quite different from the other characterised *C. difficile* endolysin CD27L (Mayer *et al.,* 2008). This study also reports the development of a model secretion system for the *C. difficile* endolysin $LysCD_{1\text{-}175}$ (a truncated derivative of LysCD), using *L. lactis* as a host cell.

5.3 Materials and Methods

5.3.1 Bacterial strains and growth conditions

Strains and plasmids used in this study are listed in Table 1. *C. difficile* strain DSMZ 1296 was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) Braunschweig, Germany. This strain was routinely cultured anaerobically at 37ºC in Fastidious Anaerobe Agar (FAA) (Lab M, United Kingdom), supplemented with 6% (v/v) defibrinated horse blood (TCS Biosciences, UK), before inoculating a single colony into Fastidious Anaerobe Broth (FAB). *E. coli* strains were grown in LB media (Sigma-Aldrich, USA) with shaking at 37ºC. *L. lactis* NZ9000 was grown in M17 (Sigma-Aldrich, USA) media supplemented with 0.5% (w/v) glucose at 30ºC, and *L. brevis* DSM 20556 was grown in MRS medium at 37ºC.

Table 1. List of bacterial strains and plasmids used in this study.

Bacterial strain or plasmid	Relevant Features	Source or reference
Strain		
E. coli XL1-Blue	Cloning host: *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI*^{q}*ΔM15* Tn*10*(Tet^{r})]	Agilent Technologies
L. lactis NZ9000	MG1363 *pep*N::nisRK	Kuipers *et al.*, 1998
L. brevis DSM 20556	Same as ATCC 8287	DSM
C. difficile DSM 1296	Same as ATCC 9689, type strain	DSM
Plasmids		
pQE60	Cloning and expression vector, Amp^{r}	Qiagen
pQE60-*lyscd*	*C. difficile* endolysin gene *lyscd* inserted into the *Nco*I/*Bgl*II site of pQE60	This study
PQE60-*lyscd*$_{1\text{-}175}$	*lyscd*$_{1\text{-}175}$ (truncated derivative of *lyscd*) inserted into the *Nco*I/*Bgl*II site of pQE60	This study
pNZ8048	High-copy-number *E. coli-L. lactis* overexpression vector, P*nisA*;, Cm^{r}	de Ruyter *et al.*, 1996
pNZ8048-*lyscd*	*L. brevis slpA* signal sequence and *lyscd* fusion inserted into the *Nco*I/*Bgl*II site of pNZ8048, *nis*A promoter replaced with *slpA* promoter	This study
pNZ8048-*lyscd*$_{1\text{-}175}$	*L. brevis slpA* and *lyscd*$_{1\text{-}175}$ fusion inserted into the *Nco*I/*Bgl*II site of pNZ8048, *nis*A promoter replaced with *slpA* promoter	This study

5.3.2 Bio-informatic analysis of LysCD endolysin

Conserved domain search on the LysCD endolysin was performed using the NCBI conserved domain database. Sequence alignment of LysCD and the previously characterised *C. difficile* endolysins: CD27L and PlyCD was determined using T-coffee alignment tool (http://tcoffee.crg.cat/apps/tcoffee/do:regular).

5.3.3 Cloning of LysCD endolysin and its subdomain into *E. coli*

The *lyscd* gene was amplified from the bacteriophage phiCD6356 genomic DNA using the KOD hotstart DNA polymerase (Novagen). To ensure cloning into *Nco*I and *Bgl*II sites of the expression vector pQE60, primers synthesized by Eurofins genomics (U.K.) were designed to incorporate these restriction sites in their sequence. These primers (Table 2) were used in the amplification the *C. difficile* endolysins. Primer set *Ec-lyscd*_F and *Ec-lyscd*_R were used in the amplification of *lyscd*, and its enzymatic domain designated $lyscd_{1\text{-}175}$ was amplified using primers *Ec-lyscd*_F and *Ec-lyscdT*_R. Both Amplicons were purified with the PCR product purification kit (Roche) before restriction digest with *Nco*I and *Bgl*II enzymes (Roche). The restricted products were ligated to pQE60, generating pQE60-*lyscd* and pQE60-$lyscd_{1\text{-}175,}$ respectively. The resulting constructs were both transformed into chemically competent *E. coli* XL1-Blue (Agilent technologies, U.S.A) and selected with 200µg/ml ampicillin (Sigma-Aldrich, USA). Positive clones were verified by colony PCR and the integrity of the plasmid was validated by DNA sequencing (GATC, Germany).

5.3.4 Construction of endolysin secretion vector and subcloning in *L. lactis*

An endolysin expression and secretion vector was constructed using the lactococcal pNZ8048 vector. Splicing by overlap extension (SOE) PCR (Horton *et al.*, 1989) was used to fuse the gene *lyscd*, as well as its enzymatic domain $lyscd_{1\text{-}175}$, with the *Lactobacillus slpA* leader sequence using primers described in Table 2. Both genes encoding endolysins were placed

under the control of the *slpA* promotor. The SOE products were then digested with *NCO*1 and *Bgl*II restriction enzymes and ligated with the pNZ8048 plasmid. The resulting vectors pNZ8048-*lyscd* and pNZ8048-*lyscd*$_{1\text{-}175}$ were transformed by electroporation into *L. lactis* NZ9000. Transformants were selected with 10µg/ml chloramphenicol (Sigma-Aldrich, USA) with plasmids from positive clones verified by DNA sequencing.

Table 2. List of primers used in this study.

Primer name	**Sequence (5' > 3')''**
*Ec-lyscd*_F	5'- ATATCCATGGAGGTTGTACTAACAGCAG -3'
*Ec-lyscd*_R	5'- CCCAGATCTTTTCTTAATAAAATCTAATACT -3'
*Ec-lyscd*T_R	5'- AGAAGATCTATTATCTATATTTTTATTTAATATACCC -3
*Lb-slpA*_F	5'- TTAAGATCTTTCAATCCAACGACAATCAGA -3'
*Lb-slpA*_R	5'- CCATGAATTCAGATGAAGCAGTCGTTGA -3'
*Ll-lyscd*_F	5'- AGCTGAATTCATGGAGGTTGTACTAACAG -3'
*Ll-lyscd*_R	5'-TCACCATGGCTATTTCTTAATAAAATCTAATACTT -3'
*Ll-lyscd*T_R	5'-TTACCATGGCTAATTATCTATATTTTTATTTAATATACCC -3'

Restriction sites are underlined

5.3.5 Protein expression

Following nucleotide sequence verification of the clones by DNA sequencing, both *E. coli* (pQE60-*lyscd)* and *E. coli* (pQE60-*lyscd*$_{1\text{-}175}$) were grown in super broth (3.2% (w/v) tryptone, 2% (w/v) yeast extract, 0.5% (w/v) NaCl) containing 200µg/ml of ampicillin to a mid-exponential phase of growth. Endolysin expression in both cultures was performed by inducing the cells with 1mM IPTG for 14hrs at 28ºC. The cells were subsequently harvested by centrifugation and re-suspended in TN buffer (20mM Tris-Hcl, pH 8.0, 50mM NaCl). Cell lysis was achieved using Bugbuster protein extraction reagent (Novagen) and cellular debris was removed by centrifugation at 4ºC together with filter sterilization of the supernatant through a 0.2µm filter. Cellular lysates for both endolysins were analysed by SDS-PAGE assay with 12% (w/v) polyacrylamide gel.

5.3.6 Endolysin lysis assay

The lytic activity of the endolysin cell lysate recovered from recombinant *E. coli* was determined using a diffuse plate assay according to the method described by Ugorcakova *et al.*, (2015), with slight modification. Briefly, *C. difficile* cells were grown to late exponential phase of growth. The cells were heat inactivated by autoclaving for 20mins and afterwards cell substrates were harvested by centrifugation (4,000*g*, 10min, 4°C), washed with equilibration buffer (50mM Tris-HCl pH 7.4, 150mM NaCl) and re-suspended in the same buffer. Cells were either used immediately or stored at −80ºC. The cell substrates were then re-suspended in BHI agar and allowed to solidify. After solidification, 5mm wells were made with sterile tips and 50µl of endolysin crude lysate added to it. The plates were incubated at 37ºC overnight and examined for clear zones surrounding the wells. Crude lysate from *E. coli* containing the empty pQE60 was used as negative control.

Endolysin secretion from *L. lactis* using the vector pNZ8048 was determined using an endolysin activity plate test according to the method described by Gaeng *et al.*, (2000). Autoclaved *C. difficile* cell substrates were embedded in GM17 agar to ensure a clearly turbid medium. *L. lactis* clones secreting the *C. difficile* endolysin $LysCD_{1\text{-}175}$ were then plated on the medium and incubated at 30ºC for 48hrs. After incubation, the agar medium was examined for the formation of clear halos surrounding the lactococcal clones.

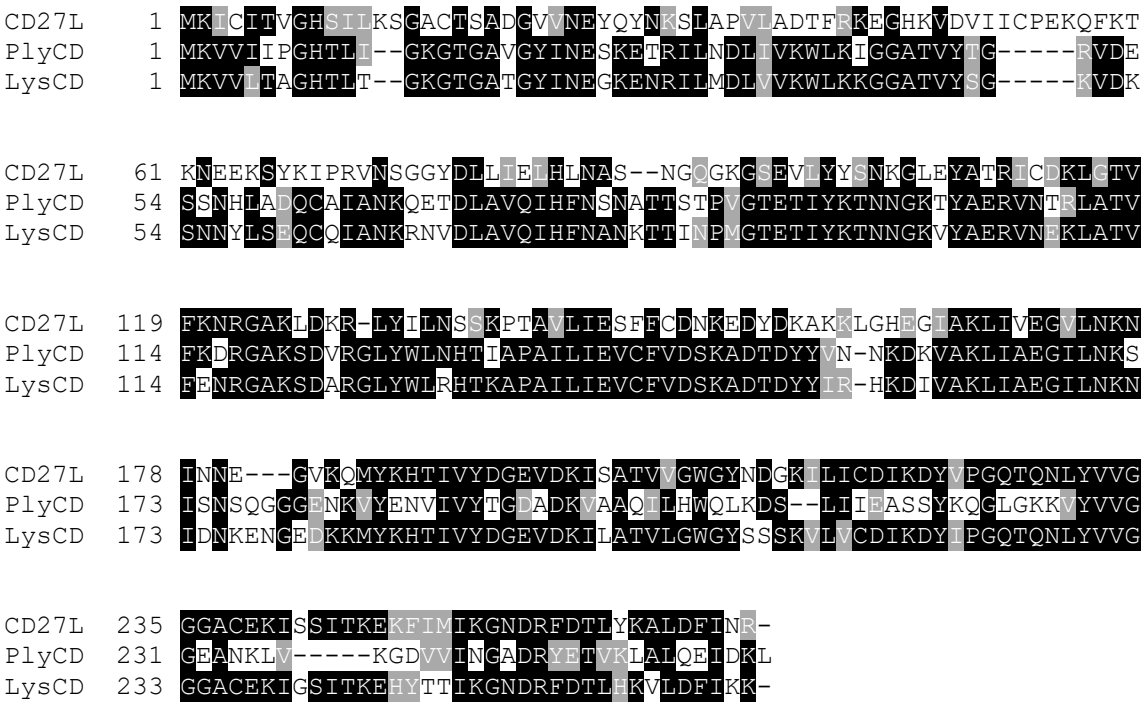
5.4 Results

5.4.1 Bioinformatic analysis of LysCD endolysin

The gene encoding a *C. difficile* endolysin and designated as *lyscd* was previously identified in the genome of the *C. difficile* bacteriophage phiCD6356 (Horgan *et al.*, 2010). The N-terminal region of this gene was predicted to encode an N-acetylmuramoyl-L-alanine amidase domain according to NCBI conserved domain search. However, no identifiable domain was observed on the C-terminal. A similar observation was made for the CD27L endolysin (Mayer *et al.*, 2011) and the truncated version of the endolysin still retained its specificity when tested in the literature. A conserved domain search on other *C. difficile* endolysin sequences (YP_006990512.1, Meessen-Pinard *et al,* 2012; AAZ32275.1, Govind *et al.*, 2006; YP_004508401.1, (Sekulovic *et al.*, 2011) and ABE99499.1, Goh *et al.*, 2007 also failed to yield a recognisable domain on their C-terminal. This suggests that the C-terminal domain for these groups of *Clostridium* endolysin may contain an uncharacterised and distinct domain.

Alignment of *lyscd* (full lysin) with other *C. difficile* endolysins (Figure 1A) that have been experimentally determined revealed a sequence identity of 52% against *cd27l* (Mayer *et al.*, 2008) and 62% against *plycd* (Wang *et al.*, 2015). Alignment on just the enzymatic domain ($lyscd_{1\text{-}175}$) (Figure 1B) revealed 39% sequence identity to $cd27l_{1\text{-}179}$ and 77% identity to $plycd_{1\text{-}174}$. This suggests that enzymatic domain of *lyscd* has an increased homology to that of *plycd* in comparison to *cd27l*.

A



B

Figure 1. Amino acid sequence alignment of *lyscd* with previously-cloned *C. difficile* lysins: *cd27l* (Genbank YP_002290910.1) (Mayer *et al.*, 2008) and *Plycd* (Genbank YP_001088405.1) (Wang *et al.*, 2015) for both the full endolysins (A) and the enzymatic domains (B) using T-coffee alignment tool. Identical amino acid residues in this alignment are shaded in black and grey.

5.4.2 Cloning and expression of LysCD and a truncated derivative $LysCD_{1\text{-}175}$

The bacteriophage phiCD6356-derived gene *lyscd* predicted as the endolysin was confirmed experimentally. It was found in the course of our experimentations that only a truncated (albeit active) version of the endolysin could be secreted in *L. lactis* as discussed later. Thus this derivative has been included in all experiments involving the full endolysin for reasons of comparison. The cloning of both variants of the endolysin involved PCR amplification of both *lyscd* and the truncated derivative $lyscd_{1\text{-}175.}$ These PCR products were digested with restriction enzymes and ligated into a pQE60 expression vector generating pQE60-*lyscd* and pQE60-$lyscd_{1\text{-}175,}$ respectively. This allowed their expression exploiting the inducible T5 promoter with IPTG induction. Both LysCD and $LysCD_{1\text{-}175}$ were expressed as a full length endolysin and a truncated derivative, respectively. Expression of both proteins was confirmed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Both bands were visible on the SDS gel as a 31kDa and 20.5kDa protein (Figure 2), which was absent on the *E. coli* (pQE60) crude lysate control.

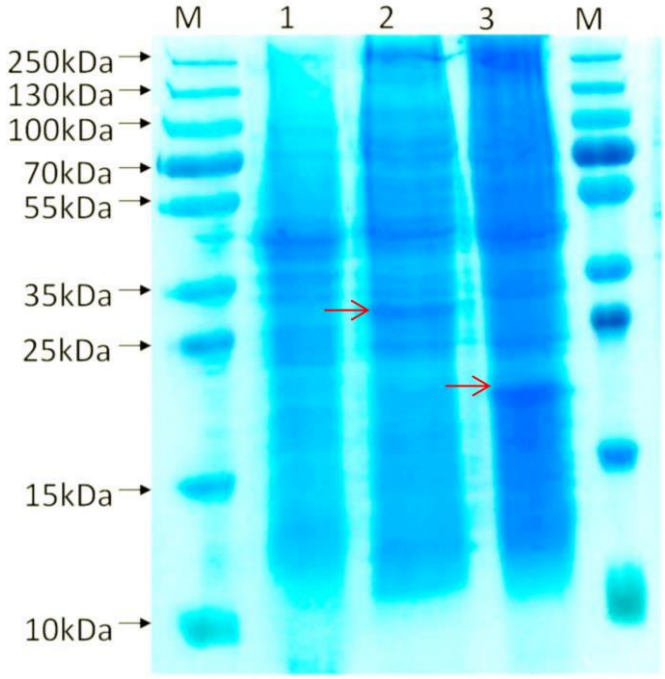


Figure 2. SDS-PAGE analysis of crudes lysates from *E. coli* expressing *lyscd* lysin. Lane M: Pre-stained protein ladder; Lane 1: Cell lysate from *E. coli* containing pQe60; Lane 2: Cell lysate from *E. coli* containing pQe60-*lyscd* with a His tag (full lysin, contains both enzymatic and cell-wall-binding domain); Lane 3: Cell lysate from *E. coli* containing pQe60-$lyscd_{1\text{-}175}$ (truncated lysin, limited to just enzymatic domain). Expression of *lyscd* and $lyscd_{1\text{-}175}$ visible in both lane 2 and 3 at the 31kDa and 20.5kDa cut-off mark respectively.

5.4.3 Enzymatic activity of LysCD and $LysCD_{1\text{-}175}$

Endolysin enzymatic activity was determined by a diffusion plate assay. The lytic activity of the endolysins on *C. difficile* cell substrates was displayed as clear zones around the well (Figure 3B&C). Crude lysate from the *E. coli* culture accommodating the plasmid pQE60 was used as a negative control and no zone of clearing was visible in the well harbouring the cell lysate (Figure 3A). $LysCD_{1\text{-}175}$ (Figure 3C) demonstrated a more effective lysis on the heat-inactivated *C. difficile* cell substrate compared to the full endolysin LysCD (Figure 3B).

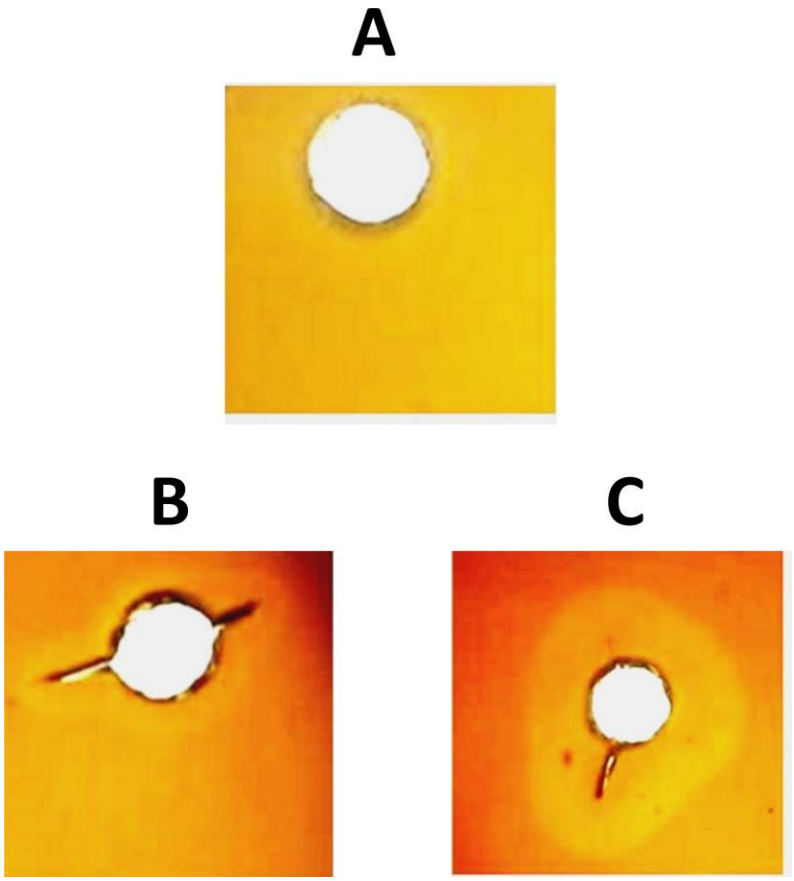


Figure 3. Diffusion plate assay demonstrating endolysin lytic activity from *E. coli* crude lysates containing (A) pQE60 (control), (B) pQE60-*lyscd* and (C) pQE60-*lyscd*$_{1\text{-}175}$ on media resuspended in heat-inactivated *C. difficile* cells substrate. Endolysin lytic activity was shown as zones of inhibition surrounding the well as a result of degradation of *C. difficile* cell substrate.

5.4.4 Secretion of LysCD$_{1\text{-}175}$ in a *Lactococcus lactis* host

The vector pNZ8048 was used for the development of a *C. difficile* endolysin secretion system. The native P*nisA* promotor of pNZ8048, which requires induction by extracellular nisin to bring about protein expression (de Ruyter *et al.*, 1996), was replaced with the *Lactobacillus slpA* (surface layer protein A) promoter to allow constitutive expression of the endolysin. Ligation reactions from the constructed secretion vector were directly transformed into *L. lactis* because recombinant plasmids previously recovered from *E. coli* clones of similar secretion vectors were observed to have deletions. Endolysin expression and secretion was demonstrated by growing a single streak of recombinant *L. lactis* on media plates

containing heat-inactivated *C. difficile* cells. *L. lactis* clones carrying the full length lysin gene (*lyscd*) failed to produce any appreciable zones of inhibition due to no lysin secretion. As a result of this observation, the active truncated derivative of the lysin ($lyscd_{1\text{-}175}$) was constructed for the secretion aspect of this work. *L. lactis* clones carrying this truncated lysin ($lyscd_{1\text{-}175}$) formed clear zones of inhibition on the agar plate containing heat-inactivated *C. difficile* cell substrate, further demonstrating the production and secretion of functionally active endolysin by the recombinant cells into its surrounding environment (Figure 5).

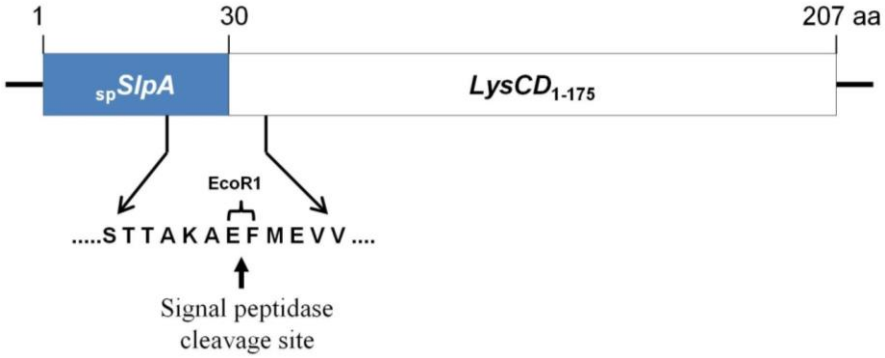


Figure 4. Schematic illustration of the genetic fusion between the surface layer protein A (*slpA*) signal sequence from *L. brevis* and the nucleotide sequence encoding the *C. difficile* endolysin ($lyscd_{1\text{-}175}$).

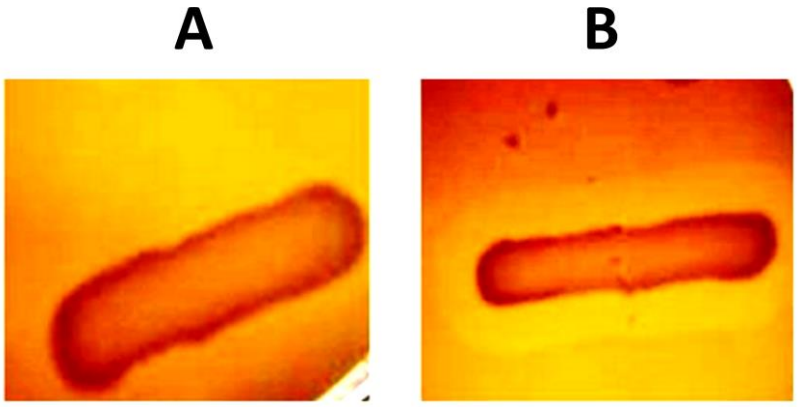


Figure 5. Endolysin secretion in *L. lactis* containing (A) pNZ8048 (control) and (B) pNZ8048-$lyscd_{1\text{-}175}$ on media resuspended in heat inactivated *C. difficile* cell substrate. The control strain has no detectable effect whereas the *L. lactis* strain containing $lyscd_{1\text{-}175}$ shows a clear halo around the bacteria.

5.5 Discussion

Clostridium difficile is the causative agent of pseudomembranous colitis, an inflammatory condition of the colon. It is a major problem in hospitals, especially with the elderly and the immuno-compromised. Current treatment options include the use of antibiotics such as metronidazole, fidaxamicin and vancomycin (Farooq *et al.*, 2015). However, due to the prevalent nature of antibiotic resistance, new treatment options are needed. Endolysins represent a potential alternative for treating these infectious bacteria. Endolysins that are active against several infectious bacteria have been characterised and are well-documented in the literature (Nelson *et al.*, 2001; Mayer *et al.*, 2008; Fenton *et al.*, 2010). Only two *C. difficile* endolysins have been characterised to date, the *cd27l* endolysin (Mayer *et al.*, 2008) and the *plycd* prophage endolysin (Wang *et al.*, 2015). In this study, the endolysin originating from the *C. difficile* bacteriophage phiCD6356 (Horgan *et al.*, 2010) was cloned and expressed in *E. coli*. The enzymatic activity of this enzyme was demonstrated on *C. difficile* cell substrate and a *C. difficile* secretion system was also developed for this enzyme using a *L. lactis* host.

LysCD is an amidase enzyme similar to the previously characterised lysins CD27L and PlyCD, which have been demonstrated to target *C. difficile* in the literature (Mayer *et al.*, 2008; Wang *et al.*, 2015). All three lysins, when examined bio-informatically, did not possess a recognisable C-terminal domain. This was quite unusual as Gram-positive phage lysins generally possess two distinct domain structures connected by a short linker: the N-terminal enzymatic domain responsible for cell lytic activity and a C-terminal cell wall binding domain associated with binding to the cell wall ligand (Fischetti, 2008). Sequence alignment on all three lysins did, however, reveal amino acid similarities between all three proteins. Similar to the CD27L endolysin (Mayer *et al.*, 2008), LysCD possessed the crucial amino acid methionine, at position 186, which was reported to be involved in an autoproteolytic

processing cleavage event leading to cleavage of the *C. difficile* endolysin CD27L at the stem of the connective linker between the enzymatic and cell wall binding domain (Dunne *et al.*, 2014). This crucial amino acid was however missing in the prophage endolysin PlyCD (Wang *et al.*, 2015).

Following expression in *E. coli*, both LysCD and $LysCD_{1-175}$ demonstrated enzymatic activities against *C. difficile* cell substrates. The *C. difficile* endolysin truncated to its enzymatic domain $LysCD_{1-175}$ had an increased activity towards *C. difficile* cell substrates in a diffusion plate assay in comparison with the full endolysin LysCD. This was possibly as a result of an increase in the rate of diffusion in the agar by $LysCD_{1-175}$ due to its significantly smaller size. The antibacterial properties of the phage-derived endolysins targeting *C. difficile* have useful potential to combat *C. difficile* associated diseases and has been demonstrated in a mouse *ex vivo* model (Wang *et al.*, 2015). The specific nature of a *C. difficile* endolysin has also been reported (Mayer *et al.*, 2008). Here, the endolysin was capable of lysing a wide range of *C. difficile* strains but failed to lyse a selection of gut commensal bacteria (Mayer *et al.*, 2008).

L. Lactis is an useful tool for the GRAS expression of recombinant proteins (Le Loir *et al.*, 2005). It has been successfully utilized in the delivery of antigens (Kasarello *et al.*, 2015), cytokines (Fernandez *et al.*, 2009) and antimicrobial peptides (Volzing *et al.*, 2013) to the gastrointestinal tract. For this purpose, this organism was chosen as suitable host for the construction of a *C. difficile* secretion system. The development of such secretion system necessitates the need of a signal peptide to bring about protein translocation across the cell wall of the host. Hence, the *Lactobacillus slpA* signal sequence was selected as a suitable candidate, as it has been used in the secretion of a *Listeria* endolysin to the surrounding environment of the host in the literature (Gaeng *et al.*, 2000). Moreover, replacement of

P*nisA* promoter with that of the *slpA* promoter also ensured constitutive production of the lysin, eliminating the need for specific induction by nisin. The *slpA* promoters (Vidgren *et al.*, 1992) has been successfully utilized for the expression of recombinant proteins such as β-lactamase (*Bla*) (Savijoki *et al.*, 1997), β-glucoronidase (*gusA*) (Kahala and Palva, 1999), luciferase (*luc*) (Kahala and Palva, 1999) and aminopeptidase (*pepN*) (Kahala and Palva, 1999), in lactococcal hosts. The first attempt at secreting the full endolysin (accommodating both the N-terminal enzymatic domain and its subsequent C-terminal) was unsuccessful resulting in no endolysin secretion. As a result of this observation, the *C. difficile* endolysin was truncated to the endolysin's enzymatic domain included in the first to 175th amino acid residue regions. This reduced size resulted in the successful secretion of active endolysin, as demonstrated by lysis assay (Figure 5). The use of *L. lactis* in the delivery of proteins to their intended target is an attractive option as it is safe, has relatively few secreted proteins and the laboratory strains do not produce any extracellular proteases that could degrade the secreted proteins (Nouaille *et al.*, 2003).

5.6 References

1. Begley, M., Gahan, C. G. M. and Hill, C. (2005) 'The interaction between bacteria and bile', *FEMS Microbiology Reviews*, 29(4), pp. 625–651. doi: 10.1016/j.femsre.2004.09.003.
2. Díez-Martínez, R., De Paz, H. D., García-Fernández, E., Bustamante, N., Euler, C. W., Fischetti, V. A., Menendez, M. and García, P. (2015) 'A novel chimeric phage lysin with high in vitro and in vivo bactericidal activity against *Streptococcus pneumoniae*.', *The Journal of Antimicrobial Chemotherapy*, 70(6), pp. 1763–73. doi: 10.1093/jac/dkv038.
3. Dunne, M., Mertens, H. D. T., Garefalaki, V., Jeffries, C. M., Thompson, A., Lemke, E. A., Svergun, D. I., Mayer, M. J., Narbad, A. and Meijers, R. (2014) 'The CD27L and CTP1L endolysins targeting clostridia contain a built-in trigger and release factor', *PLoS Pathogens*. Edited by G. Zhang, 10(7), p. e1004228. doi: 10.1371/journal.ppat.1004228.
4. Farooq, P. D., Urrunaga, N. H., Tang, D. M. and von Rosenvinge, E. C. (2015) 'Pseudomembranous colitis', *Disease-a-Month*. Elsevier, 61(5), pp. 181–206. doi: 10.1016/j.disamonth.2015.01.006.
5. Fenton, M., Casey, P. G., Hill, C., Gahan, C. G., Ross, R. P., McAuliffe, O., O'Mahony, J., Maher, F. and Coffey, A. (2010) 'The truncated phage lysin CHAP(k) eliminates *Staphylococcus aureus* in the nares of mice.', *Bioengineered Bugs*, 1(6), pp. 404–407. doi: 10.4161/bbug.1.6.13422.
6. Fernandez, A., Horn, N., Wegmann, U., Nicoletti, C., Gasson, M. J. and Narbad, A. (2009) 'Enhanced secretion of biologically active murine interleukin-12 by *Lactococcus lactis*', *Applied and Environmental Microbiology*, 75(3), pp. 869–871. doi: 10.1128/AEM.01728-08.
7. Fischetti, V. A. (2008) 'Bacteriophage lysins as effective antibacterials', *Current Opinion in Microbiology*, 11(5), pp. 393–400. doi: 10.1016/j.mib.2008.09.012.
8. Gaeng, S., Scherer, S., Neve, H. and Loessner, M. J. (2000) 'Gene cloning and expression and secretion of *Listeria monocytogenes* bacteriophage-lytic enzymes in *Lactococcus lactis*.',

Applied and Environmental Microbiology, 66(7), pp. 2951–2958. doi: 10.1128/AEM.66.7.2951-2958.2000.

9. Gao, Y., Lu, Y., Teng, K. L., Chen, M. L., Zheng, H. J., Zhu, Y. Q. and Zhong, J. (2011) 'Complete genome sequence of *Lactococcus lactis* subsp. *lactis* CV56, a probiotic strain isolated from the vaginas of healthy women.', *Journal of Bacteriology*, 193(11), pp. 2886–2887. doi: 10.1128/JB.00358-11.
10. Goh, S., Ong, P. F., Song, K. P., Rily, T. V. and Chang, B. J. (2007) 'The complete genome sequence of *Clostridium difficile* phage φC2 and comparisons to φCD119 and inducible prophages of CD630', *Microbiology*, 153(2007), pp. 676–685. doi: 10.1099/mic.0.2006/002436-0.
11. Govind, R., Fralick, J. A and Rolfe, R. D. (2006) 'Genomic organization and molecular characterization of *Clostridium dificile* bacteriophage PhiCD119', *Journal of Bacteriology*, 188(7), pp. 2568–2577. doi: 10.1128/JB.188.7.2568.
12. Hargreaves, K. R. and Clokie, M. R. J. (2014) '*Clostridium difficile* phages: Still difficult?', *Frontiers in Microbiology*, 5(184), pp. 1–14. doi: 10.3389/fmicb.2014.00184.
13. Horgan, M., O'Sullivan, O., Coffey, A., Fitzgerald, G. F., van Sinderen, D., McAuliffe, O. and Ross, R. P. (2010) 'Genome analysis of the *Clostridium difficile* phage φCD6356, a temperate phage of the *Siphoviridae* family', *Gene*. Elsevier B.V., 462(1–2), pp. 34–43. doi: 10.1016/j.gene.2010.04.010.
14. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. and Pease, L. R. (1989) 'Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension', *Gene*, 77, pp. 61–68. doi: 10.1016/0378-1119(89)90359-4.
15. Van Hylckama Vlieg, J. E. H. V., Rademaker, J. L. W., Bachmann, H., Molenaar, D., Kelly, W. J. and Siezen, R. J. (2006) 'Natural diversity and adaptive responses of *Lactococcus lactis'*, *Current Opinion in Biotechnology*, (17), pp. 183–190. doi:

10.1016/j.copbio.2006.02.007.

16. Jado, I., López, R., García, E., Fenoll, A., Casal, J., García, P., Pallares, R., de la Campa, A. G., Bouza, E., Baquero, F., Soriano, F., Prieto, J., Pallarés, R., Liñares, J., Garau, J., Martínez Lacasa, J., Latorre, C., Pérez-Trallero, E., García de Lomas, J. and Fleites, A. (2003) 'Phage lytic enzymes as therapy for antibiotic-resistant *Streptococcus pneumoniae* infection in a murine sepsis model', *Journal of Antimicrobial Chemotherapy*, 52(6), pp. 967–973. doi: 10.1093/jac/dkg485.
17. Jun, S. Y., Jung, G. M., Yoon, S. J., Oh, M. D., Choi, Y. J., Lee, W. J., Kong, J. C., Seol, J. G. and Kang, S. H. (2013) 'Antibacterial properties of a pre-formulated recombinant phage endolysin, SAL-1', *International Journal of Antimicrobial Agents*. Elsevier B.V., 41(2), pp. 156–161. doi: 10.1016/j.ijantimicag.2012.10.011.
18. Kahala, M. and Palva, A (1999) 'The expression signals of the *Lactobacillus brevis slpA* gene direct efficient heterologous protein production in lactic acid bacteria', *Applied Microbiology*, 57, pp. 71–78.
19. Kasarello, K., Kwiatkowska-Patzer, B., Lipkowski, A. W., Bardowski, J. K. and Szczepankowska, A. K. (2015) 'Oral administration of *Lactococcus lactis* expressing synthetic genes of myelin antigens in decreasing experimental autoimmune encephalomyelitis in rats', *Med Sci Monit*, 21, pp. 1587–1597. doi: 10.12659/msm.892764.
20. Kuipers, O. P., De Ruyter, P. G. G. A., Kleerebezem, M. and De Vos, W. M. (1998) 'Quorum sensing-controlled gene expression in lactic acid bacteria', *Journal of Biotechnology*, 64(1), pp. 15–21. doi: 10.1016/S0168-1656(98)00100-X.
21. Le Loir, Y., Azevedo, V., Oliveira, S. C., Freitas, D. A, Miyoshi, A., Bermúdez-Humarán, L. G., Nouaille, S., Ribeiro, L. A, Leclercq, S., Gabriel, J. E., Guimaraes, V. D., Oliveira, M. N., Charlier, C., Gautier, M. and Langella, P. (2005) 'Protein secretion in *Lactococcus lactis*: an efficient way to increase the overall heterologous protein production', *Microbial Cell*

Factories, 4(1), p. 2. doi: 10.1186/1475-2859-4-2.

22. Mayer, M. J., Garefalaki, V., Spoerl, R., Narbad, A. and Meijers, R. (2011) 'Structure-based modification of a *Clostridium difficile*-targeting endolysin affects activity and host range', *Journal of Bacteriology*, 193(19), pp. 5477–5486. doi: 10.1128/JB.00439-11.
23. Mayer, M. J., Narbad, A. and Gasson, M. J. (2008) 'Molecular characterization of a *Clostridium difficile* bacteriophage and its cloned biologically active endolysin', *Journal of Bacteriology*, 190(20), pp. 6734–6740. doi: 10.1128/JB.00686-08.
24. Meessen-Pinard, M., Sekulovic, O. and Fortier, L. C. (2012) 'Evidence of *in vivo* prophage induction during *Clostridium difficile* infection', *Applied and Environmental Microbiology*, 78(21), pp. 7662–7670. doi: 10.1128/AEM.02275-12.
25. Monaghan, T., Boswell, T. and Mahida, Y. R. (2009) 'Recent advances in *Clostridium difficile*-associated disease', *Postgraduate Medical Journal*, 85(1001), pp. 152–162. doi: 10.1136/gut.2007.128157.
26. Nelson, D., Loomis, L. and Fischetti, V. A (2001) 'Prevention and elimination of upper respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme.', *Proceedings of the National Academy of Sciences of the United States of America*, 98(7), pp. 4107–4112. doi: 10.1073/pnas.061038398.
27. Nouaille, S., Ribeiro, L. A, Miyoshi, A., Pontes, D., Le Loir, Y., Oliveira, S. C., Langella, P. and Azevedo, V. (2003) 'Heterologous protein production and delivery systems for *Lactococcus lactis*.', *Genetics and Molecular Research : GMR*, 2(1), pp. 102–111.
28. Predrag, S. (2016) 'Analysis of risk factors and clinical manifestations associated with *Clostridium difficile* disease in Serbian hospitalized patients', *Brazilian Journal of Microbiology*, 47(4), pp. 902-810. doi: 10.1016/j.bjm.2016.07.011.
29. Rodríguez-Rubio, L., Martínez, B., Rodríguez, A., Donovan, D. M., Götz, F. and García, P. (2013) 'The phage lytic proteins from the *Staphylococcus aureus* bacteriophage vB_SauS-

phiIPLA88 display multiple active catalytic domains and do not trigger staphylococcal resistance', *PLoS One*, 8(5), pp. 1–7. doi: 10.1371/journal.pone.0064671.

30. Rolfe, R. D. (1984) 'Role of volatile fatty acids in colonization resistance to *Clostridium difficile'*, *Infection and Immunity*, 45(1), pp. 185–191.
31. de Ruyter, P. G., Kuipers, O. P. and de Vos, W. M. (1996) 'Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin.', *Applied and Environmental Microbiology*, 62(10), pp. 3662–3667. doi: 0099-2240/96/$04.00 1 0.
32. Savijoki, K., Kahala, M. and Palva, A. (1997) 'High level heterologous protein production in *Lactococcus* and *Lactobacillus* using a new secretion system based on the *Lactobacillus brevis* S-layer signals', *Gene*, 186, pp. 255–262. doi: 10.1016/S0378-1119(96)00717-2.
33. Schuch, R., Nelson, D. and Fischetti, V. A (2002) 'A bacteriolytic agent that detects and kills *Bacillus anthracis*.', *Nature*, 418(6900), pp. 884–889. doi: 10.1038/nature01026.
34. Sekulovic, O., Meessen-Pinard, M. and Fortier, L. C. (2011) 'Prophage-stimulated toxin production in *Clostridium difficile* NAP1/027 lysogens', *Journal of Bacteriology*, 193(11), pp. 2726–2734. doi: 10.1128/JB.00787-10.
35. Sleator, R. D., Wemekamp-Kamphuis, H. H., Gahan, C. G. M., Abee, T. and Hill, C. (2005) 'A PrfA-regulated bile exclusion system (*Bil*E) is a novel virulence factor in *Listeria monocytogenes'*, *Molecular Microbiology*, 55(4), pp. 1183–1195. doi: 10.1111/j.1365-2958.2004.04454.x.
36. Ugorcakova, J., Medzova, L., Solteszova, B. and Bukovska, G. (2015) 'Characterization of a phiBP endolysin encoded by the *Paenibacillus polymyxa* CCM 7400 phage', *FEMS Microbiology Letters*, 362(13), pp. 1–9. doi: 10.1093/femsle/fnv098.
37. Vidgren, G., Palva, I., Pakkanen, R., Lounatmaa, K. and Palva, A. (1992) 'S-layer protein gene of *Lactobacillus brevis*: Cloning by polymerase chain reaction and determination of the nucleotide sequence', *Journal of Bacteriology*, 174(22), pp. 7419–7427. doi:

10.1128/JB.174.22.7419-7427.1992.

38. Viswanathan, V. K., Mallozzi, M. J. and Vedantam, G. (2010) '*Clostridium difficile* infection an overview of the disease and its pathogenesis, epidemiology and interventions', *Gut Microbes*, 1(4), pp. 234–242. doi: 10.4161/gmic.1.4.12706.
39. Volzing, K., Borrero, J., Sadowsky, M. J. and Kaznessis, Y. N. (2013) 'Antimicrobial peptides targeting Gram-negative pathogens, produced and delivered by lactic acid bacteria', *ACS Synthetic Biology*, 2(11), pp. 643–650. doi: 10.1021/sb4000367.
40. Wang, Q., Euler, C. W., Delaune, A. and Fischetti, V. A. (2015) 'Using a novel lysin to help control *Clostridium difficile* infections', *Antimicrobial Agents and Chemotherapy*, 59(12), pp. 7447-7457. doi: 10.1128/AAC.01357-15.
41. Watson, D., Sleator, R. D., Hill, C. and Gahan, C. G. M. (2008) 'Enhancing bile tolerance improves survival and persistence of *Bifidobacterium* and *Lactococcus* in the murine gastrointestinal tract.', *BMC microbiology*, 8, p. 176. doi: 10.1186/1471-2180-8-176.
42. Yang, X., Wang, Y., Zhou, Y., Gao, X. and Li, B. (2014) 'Complete genome sequence of *Lactococcus lactis* subsp. *lactis* KLDS4.0325, a bacterium newly isolated from Koumiss in Xinjiang , China', *Acta Microbiologica Sinica*, 54(12), pp. 1406–1418.

Chapter VI

Evaluation and optimisation of different *Clostridium difficile*-endolysin-secreting hosts for delivery of endolysin into the gastrointestinal tract environment

6.1 Abstract

Commensal gastrointestinal microorganisms are highly adapted to the harsh environment of the gastrointestinal tract. Accordingly, microorganisms employed as probiotics need to withstand or be adapted to withstand the same stresses. The development of probiotic organisms as oral delivery vehicles for the transport of therapeutic substances to the gastrointestinal tract is a worth-while endeavour. A case in point is their exploitation to deliver the *C. difficile* phage endolysin to target this pathogenic bacterium at the site of infection in the colon. In this study, the *C. difficile Lactococcus*-based lysin secretion system developed in the previous chapter was selected, and efforts were made to improve its tolerance to bile. Following limited success, the secretion system was translocated to *L. salivarius*, a naturally bile-tolerant strain. Again, following limited success with this system, another system was adapted to *E. coli*, with the intention of ultimately utilising the probiotic *E. coli* Nissle for delivery. The endolysin was successfully secreted from the *E. coli* host and easily detectable levels of secreted endolysin were observed.

6.2 Introduction

Oral delivery of therapeutic proteins for intestinal disorders has its advantages and challenges in modern pharmaceutical biotechnology. Its main advantage is that it possesses less risk of immunogenicity, discomfort and pain that are usually associated with the parenteral route of administration (Berlec *et al.,* 2012). Its obvious draw-back is degradation of these proteins by digestive enzymes in the upper gastrointestinal tract. Therefore, a more sophisticated means of delivering these therapeutic proteins to the lower GIT, whereby they are protected from destruction in the stomach and duodenum is of biotechnological interest.

Lactic acid bacteria (LAB) are promising vehicles for delivering therapeutic molecules to the GIT, given that many members of this group naturally reside there as part of the commensal flora. They are non-pathogenic, and possess GRAS status. These organisms are also used in the fermentation and preservation of dairy, meat and vegetable products and are thus safely consumed by humans (and animals in the case of silage). Lactococci and lactobacilli have previously been manipulated for delivering therapeutic molecules to the GIT (Bermúdez-Humarán *et al.*, 2011) and antigens such as Tetanus toxin Fragment C and *Bacillus anthracis* protective antigen have succesfully been delivered as oral vaccines using members of these LAB genera (Robinson *et al.*, 1997; Grangette *et al.,* 2002; Mohamadzadeh *et al.,* 2009). Similarly, therapeutic proteins such as Interleukin-10 and alpha-melanocyte stimulating hormone have also been delivered to the GIT using these organisms (Steidler *et al.*, 2000; Braat *et al.*, 2006; Yoon *et al.*, 2008). Their exploration for delivery of a *C. difficile*–targeting antimicrobial is thus a worth-while endeavour.

C. difficile is the causative agent for infectious diseases that range in severity from mild diarrhoea to a life-threatening pseudomembranous colitis (Burke and Lamont, 2014). In Chapter V, the design and construction of a *C. difficile* lysin delivery model system in an *L.*

lactis host, capable of secreting this peptidoglycan hydrolase to its external environment was carried out. This model system successfully secreted the *C. difficile* lysin into its external environment. This study improves upon the model system for potential delivery of *C. difficile* endolysin into the GIT by firstly, adaptating the *C. difficile* endolysin-secreting *L. lactis* to better tolerate bile. Secondly transferring the secretion system into a naturally bile-tolerant *L. salivarius* and thirdly, setting up a secretion system in a naturally bile-tolerant *E. coli* and all three systems were compared.

6.3 Materials and Methods

6.3.1 Bacterial strains, plasmids and growth condition

The bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* strains were grown in Luria-Bertani (LB) medium (Sigma-Aldrich, USA) at 37°C with shaking. *L. lactis* strains were grown in M17 medium (Sigma-Aldrich, USA) supplemented with 0.5% (w/v) glucose at 30°C with no shaking. *L. salivarius* strains were grown in de Man, Rogosa and Sharpe (MRS) medium at 37°C without shaking.

6.3.2 Construction of bile tolerant *C. difficile*-lysin secretion system in *L. lactis*

The primer pair *BilE*_F (5'-CATTCTGCAGGCGGAAACTTTGTTTGTAAG-3') and *BilE*_R (5'-AATGTCTAGATGGTTTTTACGCCACTTCG-3') was used in the amplification of the *Bil*E gene using the *L. monocytogenes* EDG-e (Sleator *et al.*, 2005) DNA template. The resulting 2.8-kb PCR product was digested with *Pst*I and *Xba*I restriction enzymes and ligated using T4 DNA ligase (Roche Applied Science, Germany) into the lactoccocal *C. difficile* lysin secretion vector (pNZ8048-*lyscd*$_{1\text{-}175}$), created in the Chapter V. The resulting plasmid, designated as pNZ8048-*lyscd*$_{1\text{-}175}$*bilE* was transformed by electroporation into *L. lactis* NZ9000. Transformants were selected with 10µg/ml chloramphenicol, with plasmids from insert positive clones recovered and their integrity confirmed by sequencing. Recombinant *L. Lactis* containing pNZ8048-*lyscd*$_{1\text{-}175}$*bilE* plasmid was tested for *C. difficile* lysin secretion by incorporating autoclaved *C. difficile* cell substrates in media followed by streaking of recombinant cell.

Table 1. Bacterial strains used in this study

Bacterial Strains	Details	Source or reference
E. coli DH5α	Cloning host: *F− supE44 ΔlacU169 Φ80lacZ ΔM15 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Invitrogen
E. coli K12 N3406 (pJL3)	Cloning host transformed with pJL3 vector containing bacteriocin-release-protein (BRP)	MoBiTec
E. coli (pQE60-*lyscd*$_{1\text{-}175}$)	*C. difficile* phage lysin containing amidase domain cloned into *E. coli* XL1-Blue	(Chapter V, This thesis)
E. coli (pQE60/pJL3)	Harbours pQE60 empty plasmid and pJL3 containing BRP	This study
E. coli (pQE60-*lyscd*$_{1\text{-}175}$/pJL3)	pQE60 containing *C. difficile* lysin and pJL3 containing BRP cloned into *E. coli*	This study
L. lactis NZ9000	MG1363 *pep*N::nisRK	Kuipers *et al.*, 1998
NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$)	pNZ8048 containing *C. difficile* lysin with signal peptide attached for secretion	(Chapter V, This thesis)
NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$*BilE*)	pNZ8048 containing *C. difficile* lysin with signal peptide and bile resistant gene	This study
L. monocytogenes EDG-e	Wild-type of serotype 1/2a for which the genome sequence is available	Sleator *et al.*, 2005
L. salivarius NRRL B-30514	Host strain, originally isolated from cecal contents of broiler chicken	Stern *et al.*, 2006
L. salivarius (pNZ8048)	Harbours the pNZ8048 empty plasmid	This study
L. salivarius (pNZ9530)	Harbours the pNZ9530 helper plasmid, Ery^{r}	This study
L. salivarius (pNZ8048-*lyscd*$_{1\text{-}175}$)	Harbours pNZ8048 containing *C. difficile* lysin with signal peptide attached for secretion	This study
L. salivarius (pNZ8048/pNZ9530)	Harbours pNZ8048 empty plasmid and pNZ9530 helper plasmid	This study
L. salivarius (pNZ8048-P*nisAlyscd*$_{1\text{-}175}$/pNZ9530)	Harbours pNZ9530 helper plasmid and pNZ8048 containing *C. difficile* lysin with signal peptide	This study

6.3.3 Resistance to bile

The ability of the recombinant *L. lactis* cell to withstand porcine bile was investigated. These cells were grown to an early stationary phase, harvested and re-suspended to the same cell density in GM17 medium containing 0.05% (w/v) porcine bile. Viable plate counts were performed after 24hrs. Additionally, these cells were grown overnight and diluted (1 in 100) in fresh media with and without porcine bile (0.01% w/v). Their growth was monitored for 8hrs using a micro-plate reader and *L. lactis* NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$), absent for the *BilE* gene was used as control.

6.3.4 Preparation of competent *Lactobacillus* cells

L. salivarius NRRL B-30514 and *L. salivarius* (pNZ9530) were made competent according to the method described by Park and Stewart, (1990). Briefly, an overnight culture of *L. salivarius* was diluted (1 in 100) in fresh media and grown at 37°C to an OD_{600nm} of 0.2. At this point, Penicillin G was added to a final concentration of 8µg/ml and grown for an additional hour. Cells were harvested by centrifugation and washed twice in ice cold 10mM $MgCl_2$, once in ice cold 0.5M sucrose/10% (v/v) glycerol buffer and re-suspended in 100µl of the same solution. These electro-competent cells were kept on ice and used within 30mins of preparation.

Electroporation was carried out using the Eppendorf Eporator (Eppendorf, Germany). 1µg/µl of plasmid DNA was added to 45µl of competent cells and the mixture was inserted into a chilled 0.2cm electrode gap electroporation cuvette (Biorad) and incubated on ice for 5mins. The cells within the cuvette were pulsed at 1,750kV and 1ml of pre-warm MRS-SM buffer (MRS broth, 300mM Sucrose, 80mM $MgCl_2$) was added to it. These cells were incubated for a minimum of 3hrs at 37°C and plated on MRS agar containing the appropriate antibiotics.

6.3.5 Construction of a *Lactobacillus*-secreting *C. difficile* lysin system

The *slpA* promoter used to establish the *C. difficile* lysin secretion vector for *L. lactis* in the previous chapter was replaced with the P*nisA* promoter. This involved amplification of the *C. difficile* lysin and *slpA* signal peptide with the primer pair *Lb-lyscd/pnisA*_F (5'-ATTCCATGGATGCAATCAAGTTTAAAGAAATCTC-3') and *Lb-lyscd/pnisA*_R (5'-TTATCTAGACTAATTATCTATATTTTTATTTAATATACCCTC-3') using the *C. difficile* lysin secretion vector constructed in Chapter V as template. The resulting 642-bp PCR product was digested with *Nco*I and *Xba*I restriction enzymes and subsequently ligated to pNZ8048 vector. The resulting plasmid designated pNZ8048-P*nisAlyscd*$_{1\text{-}175}$ was transformed into *L. lactis* NZ9000 and subsequently into *L. salivarius* (pNZ9530). Additionally, the *C. difficile* lysin secretion vector containing the *slpA* promoter was isolated from *L. lactis* (pNZ8048-*lyscd*$_{1\text{-}175}$) and introduced by transformation into *L. salivarius* NRRL B-30514. Transformants were selected with 5µg/ml chloramphenicol for transformed *L. salivarius* NRRL B-30514 and a combination of 5µg/ml chloramphenicol and 5µg/ml erythromycin for *L. salivarius* (pNZ8048/pNZ9530).

6.3.6 Construction of a *C. difficile* lysin delivery system in an *E. coli* host

Competent *E. coli* cells already harbouring the pJL3 vector were purchased from MoBiTec GmbH (Germany). The plasmids pQE60 and pQE60-*lyscd*$_{1\text{-}175}$ were introduced by transformation into these competent cells using the traditional heat shock method. Transformants were selected on LB plates containing ampicillin (200µg/ml) and chloramphenicol (34µg/ml) for pJL3/pQE60-*lyscd*$_{1\text{-}175}$ transformants.

6.3.7 Expression and activity of *C. difficile* lysin excreted by recombinant *E. coli*

E. coli cells harbouring pJL3/pQE60-*lyscd*$_{1\text{-}175}$ plasmids were grown in super broth (3.2% (w/v) tryptone, 2% (w/v) yeast extract, 0.5% (w/v) NaCl) containing ampicillin (200µg/ml) and chloramphenicol (34µg/ml) at 37°C and subcultured into 100mls of fresh super broth. The subcultured cells were grown at 37°C to an OD_{600nm} of 0.5 and induced with 20µM IPTG. After induction, the cells were grown for an additional 16hrs at 26°C and then harvested by centrifugation. The supernatant was concentrated by 100 fold using a 10kDa Amicon Ultra-15 centrifugal filters unit (Merck, Millipore, Darmstadt, Germany). Both cellular and extracellular fractions were detected for *C. difficile* lysin by SDS-PAGE analysis and the activity of the lysin analysed by diffuse plate assay as described in chapter V.

6.4 Results

6.4.1 Analysis on the ability of a model *C. difficile*-lysin-secreting *L. lactis* containing *Bil*E gene to tolerate porcine bile

The *bilE* system; a two-gene operon consisting of the *bilEA* and *bilEB* genes (Sleator *et al.*, 2005), were amplified from a preparation of *L. monocytogenes* template DNA and ligated to the *C. difficile* secretion vector (pNZ8048-*lyscd*$_{1\text{-}175}$ detailed in Chapter V). The 6.7kb plasmid (Figure 1) was introduced by electroporation into *L. lactis* NZ9000. The recombinant strain designated NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$*bilE*) was evaluated for its ability to survive the lowest concentration of bile salts encountered in the human intestine at 0.05% (w/v) under normal physiological conditions (Islam *et al.*, 2011; Ruiz *et al.*, 2013). There was no difference in the viable cell reduction brought about by porcine bile at 0.05% (w/v) between NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$*bilE*) and the control strain "NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$)" used in this study (Table 2). However, the *bilE* gene did have some improvement on our

recombinant *C. difficile*-lysin-secreting *L. lactis* at concentration of 0.01% (w/v) bile (Figure 2). Additionally, in the absence of the *bilE* gene, the *C. difficile*-lysin-secreting *L. lactis* grew poorly in GM17 medium containing 0.01% (w/v) bile. However, with the introduction of the *bilE* gene into the secretion vector, *C. difficile* lysin secretion was completely eliminated, possibly due to burden on the cell in the lactococcal system (discussed in detail later).

Table 2. *L. lactis* survival in medium containing 0.05% (w/v) porcine bile

Time (hrs)	**NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$*BilE*) cell count (CFU/ml)**	**NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$) cell count (CFU/ml)**
0	1.48 x 10^9	1.78 x 10^9
24	2.70 x 10^6	3.00 x 10^6

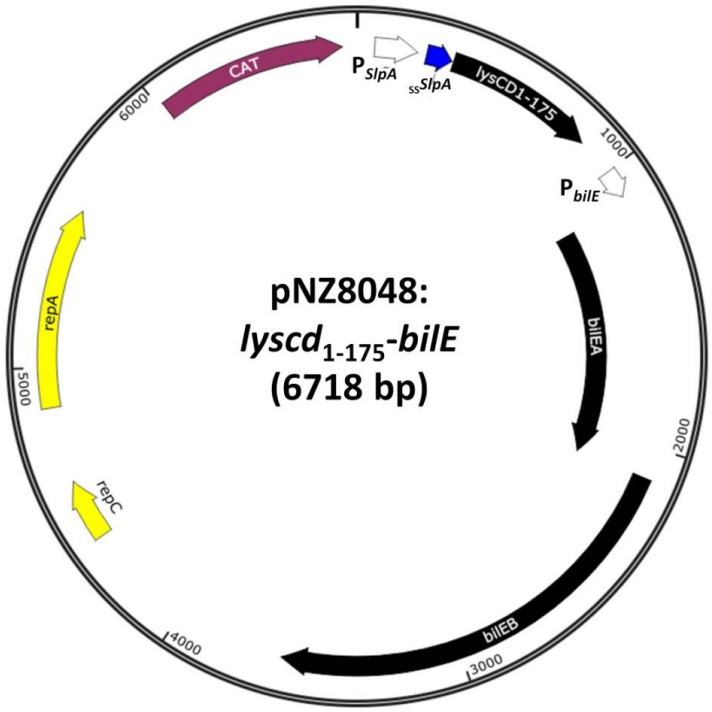


Figure 1. Schematic illustration of the *C. difficile* lysin secretion vector fused with the *bilE* gene.

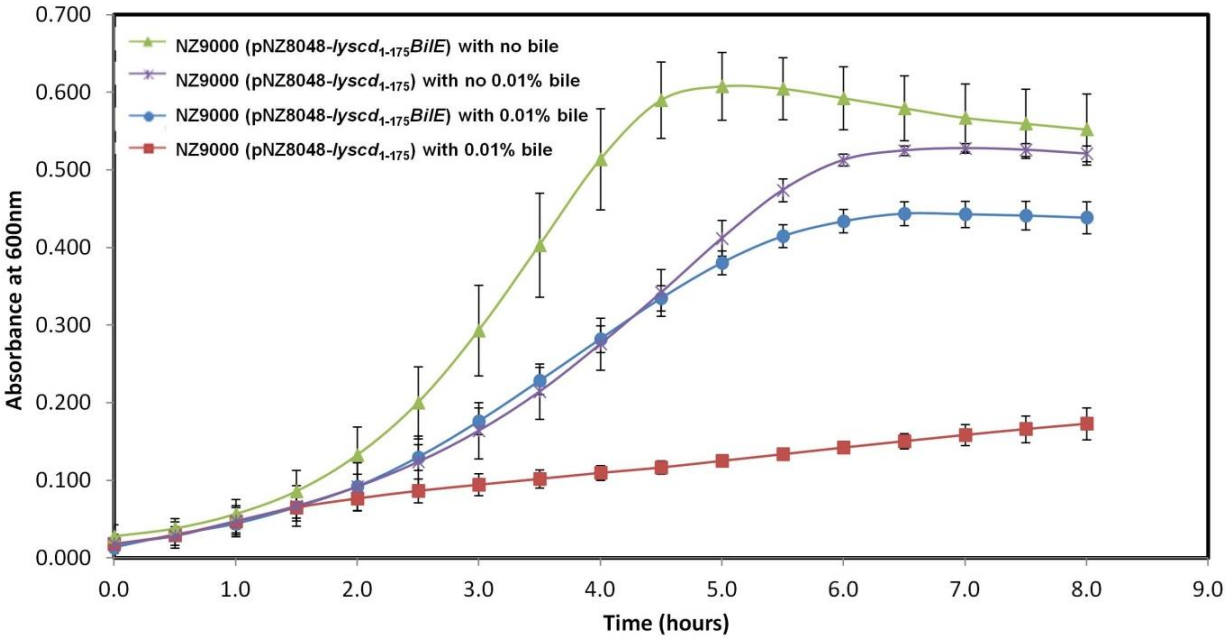


Figure 2. Growth profile of several recombinant *L. lacti*s strains in GM17 medium containing 0.01% (w/v) porcine bile as well as in the absence of bile.

6.4.2 Design/development of a *Lactobacillus* secreting *C. difficile*-lysin

The *C. difficile* lysin secretion vector pNZ8048-*lyscd*$_{1\text{-}175}$ was isoated from recombinant *L. lactis* NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$) and introduced by electroporation into *L. salivarius* NRRL B-30514 was attempted repeatedly. The transformation efficiency for the control plasmid pNZ8048 was quite low, and unfortunately no *Lactobacillus* transformants could be recovered for pNZ8048-*lyscd*$_{1\text{-}175}$. Thus, an alternative strategy exploiting the nisin inducible expression system P*nisA* was attempted. This strategy involved replacing the constitutive *slpA* promoter with the nisin inducible P*nisA* promoter. The resulting vector was then transformed into *L. lactis* NZ9000 and subsequently into *L. salivarius* NRRL B-30514. The *L. lactis* strain NZ9000 used in this study possesses the necessary regulatory genes *nisK* and *nisR* in the chromosome (Kuipers *et al.*, 1998). However, these genes are not present in the chromosome of *L. salivarius* NRRL B-30514. To this end, the helper plasmid pNZ9530 containing both regulatory genes was also introduced by electroporation into *L. salivarius* NRRL B-30514. The recombinant *L. salivarius* containing both the pNZ9530 helper plasmid

and pNZ8048-P*nisAlyscd*$_{1\text{-}175}$ plasmids was then assayed for *C. difficile* lysin secretion using the standard agar plate system where the plate was seeded with susceptible heat-inactivated *C. difficile* cells and also with sufficient nisin (5ng/ml) to induce endolysin expression and secretion by the lactobacilli. However, following nisin induction of the cloned endolysin in *L. salivarius* NRRL B-30514, no *C. difficile* could be detected around the *L. salivarius* streaks (Figure 3A) by comparison with control *Lactococcus* streaks containing the same pNZ8048-P*nisAlyscd*$_{1\text{-}175}$ (Figure 3B).

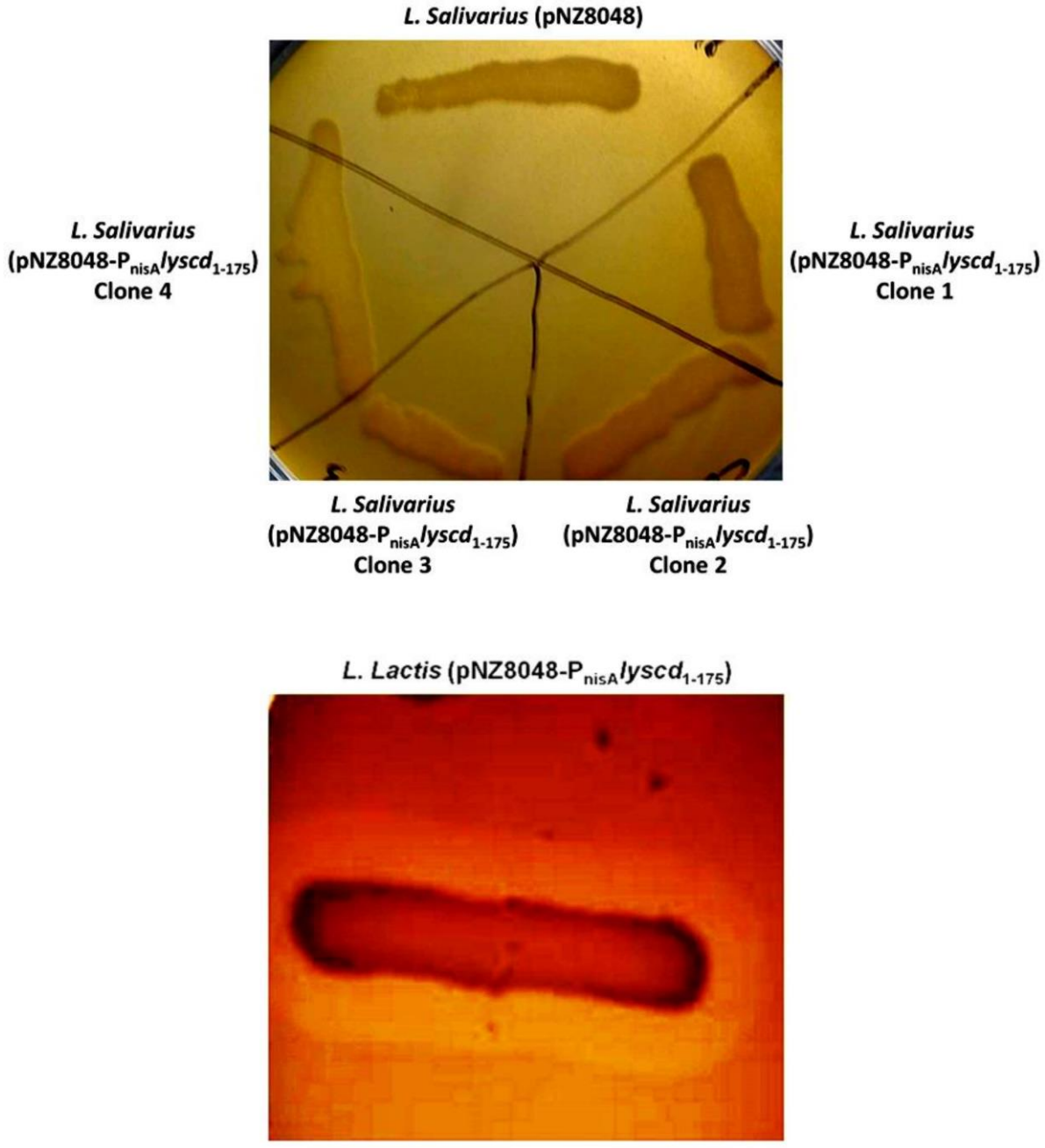


Figure 3. Lysin activity-plate-lysis assay of recombinant *L. salivarius* strains (A) and a recombinant *L. lactis* strain as positive control (B) on MRS medium containing heat-inactivated *C. difficile* cell substrate.

6.4.3 Construction of an *E. coli* secreting *C. difficile*-lysin

Following the experiences using the bile-tolerant *Lactococus lactis* and the *Lactobacillus salivarius* discussed above, *E. coli* was chosen as a possible superior secretion host for delivery of the *C. difficile* lysin into the lower gastrointestinal tract. This system involved co-expression of the bacteriocin-release-protein (BRP) gene encoded on the pJL3 vector, with *C. difficile* lysin gene encoded on the pQE60-*lyscd*$_{1\text{-}175}$ vector. BRP is a 28-amino acid lipoprotein, which is produced as a precursor containing a signal peptide allowing for its secretion across the cytoplasmic membrane (Van der Wal *et al.*, 1995). BRP activates the detergent-resistant phospholipase A, resulting in the formation of permeable pore in the cell envelope (Figure 4), allowing for protein release into the external environment of the cell (Choi and Lee, 2004). The *C. difficile* lysin is regulated by the pQE60-associated T5 expression signal, whereas the BRP is regulated by the lpp-lac tandem promoter/operator system (Hsiung *et al.*, 1989). Both the *C. difficile* lysin and BRP require induction by IPTG and the concentration range at which BRP could be induced (10µM to 40µM) without triggering complete cell lysis was investigated to determine the levels of *C. difficile* lysins produced at these concentrations. It was observed that the regulatory system in the expression of *C. difficile* lysin was not tightly controlled leading to leaky expression of the enzyme (Figure 5A). Upon induction with IPTG, the levels of *C. difficile* lysin increased concomitant with increasing IPTG concentrations. Examination of the extracellular fraction by SDS-PAGE revealed that the 20.5kDa *C. difficile* lysin was indeed released to the extracellular medium together with other cytosolic and periplasmic proteins (Figure 5B). This *C. difficile* lysin released into the extracellular matrix was also deemed to be active against heat-inactivated *C. difficile* cell substrate (Figure 6).

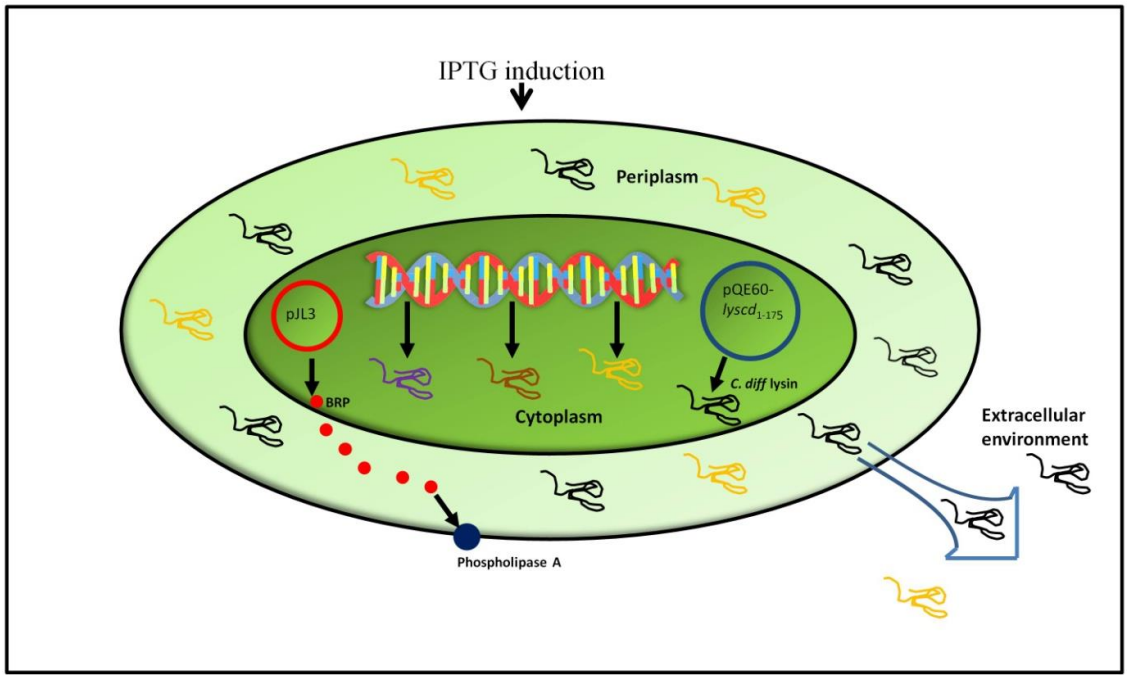


Figure 4. Schematic representation of the secretory strategy exploited in *E. coli* to release its proteins to the extracellular environment. The pQE60-*lyscd*$_{1\text{-}175}$ encoding the endolysin and the pJL3 encoding BRP are co-transformed into the *E. coli* cell. BRP activates phospholipase A, which causes pores in the membrane of cell. Proteins (host and recombinant) are then released throught these pores into extracellular medium.

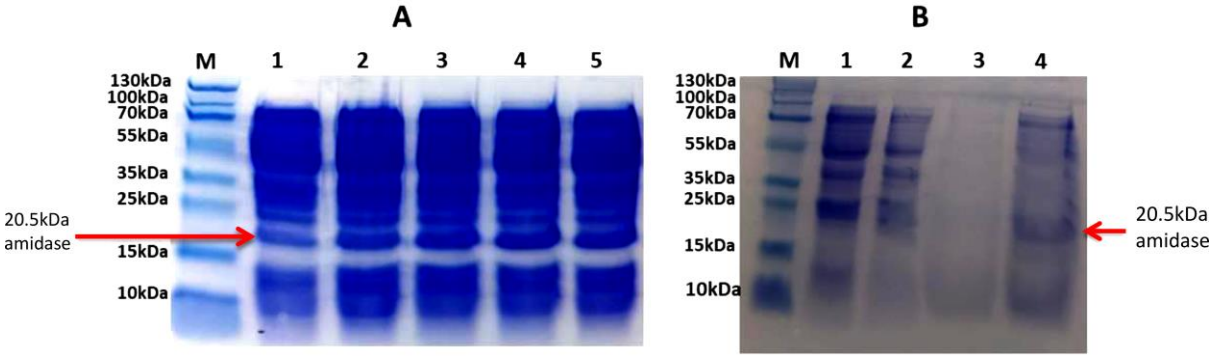


Figure 5. SDS-PAGE analysis of *C. difficile* endolysin produced by the recombinant *E. coli*. A: Increasing intracellular production of the endolysin (at 20.5kDa) with increasing IPTG concentrations. Lane M: molecular weight markers, Lane 1: no IPTG, Lane 2: 10µM IPTG, Lane 3: 20µM IPTG, Lane 4: 30µM IPTG, Lane 5: 40µM IPTG. B: Extracellular secretion of *C. difficile* endolysin from recombinant *E. coli* harbouring pQE60-*lyscd*$_{1\text{-}175}$ and pJL3; Lane M: molecular weight markers, Lane 1: non-induced intracellular fraction, Lane 2: IPTG-induced intracellular fraction, Lane 3: non-induced extracellular fraction, Lane 4: IPTG-induced extracellular fraction showing secreted proteins including the 20.5kDa endolysin.

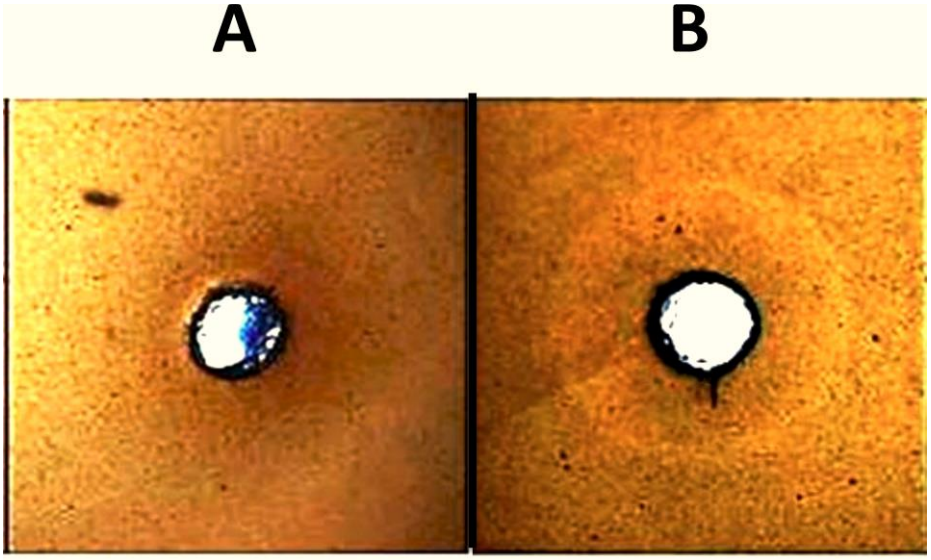


Figure 6. Diffusion plate assay demonstrating lysin activity from concentrated spent supernatant of *E. coli* host harbouring (A) pQE60/pJL3 (control) and (B) pQE60-*lyscd*$_{1\text{-}175}$/pJL3 on media containing heat-inactivated *C. difficile* cells. Endolysin secretion is evident in plate B.

6.5 Discussion

The *L. lactis*-secreting *C. difficile* lysin was shown in the previous chapter to efficiently secrete the 20.5kDa lysin into its surrounding environment. However, *L. lactis* does not colonize nor readily survive in the mammalian gastrointestinal tract (GIT) as it possesses poor natural tolerance to stresses encountered in the GIT (Li *et al.*, 2015) in comparison to other LABs such as *Bifidobacterium* and *Lactobacillus* (Ruiz *et al.*, 2013). As such, the aim of this study was to explore workable strategies for the construction of a *C. difficile* lysin delivery system that could potentially work in the presence of the various stresses encountered in the mammalian intestine.

The main role of bile in the intestine is the emulsification of fat but, in addition, it also has significant antimicrobial properties (Begley *et al.*, 2005; Kimoto *et al.*, 2003). A strategy employed in improving *L. lactis's* tolerance to bile involved the use of the bile exclusion

system (*bilE*) native to *L. monocytogenes*, which facilitates the exclusion of bile salts from the cell similar to that of the multidrug efflux pump of Gram-negatives bacteria (Sleator *et al.*, 2005). In this chapter, the *bilE* gene was incorporated into the *C. difficile*-lysin-secreting *L. lactis* with the aim of improving the tolerance of the recombinant *L. lactis* in bile. The bile concentration encountered in the human intestine under normal physiological conditions is within the ranges of 0.05% to 2.0% (w/v) (Islam *et al.*, 2011; Ruiz *et al.*, 2013) and thus 0.05% (w/v) was used in the assays. However, no difference in the titre of recovered viable cells of *L. lactis* NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$*bilE*) was observed in comparison to the titre of control *L. lactis* NZ9000 (pNZ8048-*lyscd*$_{1\text{-}175}$) cells. This is in contrast to the study by Watson *et al.*, (2008) where expression of *bilE* in *L. lactis* resulted in increased tolerance to 1.0% (w/v) porcine bile. We hypothesize that the expression of both the *bilE* gene as well as the gene encoding the *C. difficile* lysin within the same expression system and host may have resulted in an increased metabolic burden to the host cell. Metabolic burden brought about by expressing recombinant proteins have been reported in the literature (Bentley *et al.*, 1990; Rosano and Ceccarelli, 2014) where resources were reported to be drawn from the host metabolism for the expression and maintenance of the foreign DNA. It is possible that the increased metabolic burden in the expression and export of *C. difficile* lysin in the lactoccocal host may have resulted in lower expression of *bilE* as a result leading to its poor survival in 0.05% (w/v) bile. This metabolic burden on the lactococcal host may also be responsible for the abolished *C. difficile* lysin secretion also encountered here.

The naturally bile-tolerant *L. salivarius* (Messaoudi *et al.*, 2013) was then used as the delivery host for the *C. difficile* endolysin. However, the transformation of pNZ8048-*lyscd*$_{1\text{-}175}$ into the *Lactobacillus* host was unsuccessful despite several attempts. This may have been caused by lysin toxicity in this host. As a result, an alternative approach involving the use of the nisin inducible system was attempted to avoid constitutive expression of the lysin. The

nisin expression system employed includes the regulatory genes *nisK* and *nisR*. The *nisK* gene encodes the histidine-protein kinase, which upon binding with nisin undergoes autophosphorylation thereby transferring a phosphate group to the response regulator (*nisK*) activating it (Mierau and Kleerebezem, 2005). The activated response regulator induces the transcription of the gene downstream of the P*nisA* promoter. This approach resulted in the successful transformation of pNZ8048-P*nisAlyscd*$_{1\text{-}175}$ vector into the *Lactobacillus* strain. Unfortunately, no secretion was observed when the transformants were assayed for lysin activity by plate assay with heat-inactivated *C.difficle* cells. Factors influencing the unsuccessful secretion of *C. difficile* lysin in this heterologous host may have been rare codon usage, mRNA and/or protein instability, or stress induced metabolic burden from expressing the recombinant protein as well as incorrect protein conformation (Le Loir *et al.*, 2005; Sørensen and Mortensen, 2005).

The third attempt at creating a *C. difficile* lysin delivery system to the GIT with improved tolerance to bile involved the use of the *E. coli*, which includes both commensal as well as pathogenic strains. A small number of the commensal strains have been used in the development of probiotics and these include the *E. coli* Nissile 1917 (Mutaflor) and *E. coli* DSM 17252 (Symbioflor 2) and it was on this basis that *E. coli* was chosen. Most recombinant proteins secreted in *E. coli* are usually translocated to the periplasmic space (Choi and Lee, 2004). In order to release these proteins to the extracellular environment, strategies such as treating host cells with certain agent like glycine/lysozyme (Yang *et al.*, 1998; Jang *et al.*, 1999) or co-expression with *Kil* genes (Kliest *et al.*, 2003) or bacteriocin-release-protein (BRP) genes (Hsiung *et al.*, 1989) have been adopted. We used co-expression of the BRP with *C. difficile* endolysin with the goal of releasing the endolysin to the extracellular environment. *C. difficile* endolysin expression was successful as a result of the BRP activating the dormant phospholipase A present in *E. coli* leading to formation of trans-

envelope pores releasing the *C. difficile* lysin. This secretion also led to the release of the many native cytoplasmic and periplasmic proteins from the *E. coli* host. Controlled expression of BRP by modulating the IPTG concentrations was necessary to prevent complete lysis of the recombinant cell from the membrane pore formation, thereby maintaining the viability of the *E. coli* secretion host. The *C. difficile* lysin secreted by this system was shown to be enzymatically active. Given the earlier technical difficulties, and the inability of other workers to set up such a system, this was considered to be a big achievement representing a first step towards the delivery of *C. difficile* lysin to the lower intestinal tract using a probiotic *E. coli* host.

6.6 References

1. Begley, M., Gahan, C. G. M. and Hill, C. (2005) 'The interaction between bacteria and bile', *FEMS Microbiology Reviews*, 29(4), pp. 625–651. doi: 10.1016/j.femsre.2004.09.003.
2. Bentley, W. E., Mirjalili, N., Andersen, D. C., Davis, R. H. and Kompala, D. S. (1990) 'Plasmid-encoded protein: The principal factor in the "metabolic burden" associated with recombinant bacteria', *Biotechnology and Bioengineering*, 35(7), pp. 668–681. doi: 10.1002/bit.260350704.
3. Berlec, A., Ravnikar, M. and Strukelj, B. (2012) 'Lactic acid bacteria as oral delivery systems for biomolecules.', *Die Pharmazie*, 67(11), pp. 891–898. doi: 10.1126/science.289.5483.1352.
4. Bermúdez-Humarán, L. G., Kharrat, P., Chatel, J. M. and Langella, P. (2011) 'Lactococci and lactobacilli as mucosal delivery vectors for therapeutic proteins and DNA vaccines', *Microbial Cell Factories*. BioMed Central Ltd, 10(suppl. 1), p. S4. doi: 10.1186/1475-2859-10-S1-S4.
5. Braat, H., Rottiers, P., Hommes, D. W., Huyghebaert, N., Remaut, E., Remon, J. P., van Deventer, S. J. H., Neirynck, S., Peppelenbosch, M. P. and Steidler, L. (2006) 'A phase I trial with transgenic bacteria expressing Interleukin-10 in crohn's disease', *Clinical Gastroenterology and Hepatology*, 4(6), pp. 754–759. doi: 10.1016/j.cgh.2006.03.028.
6. Burke, K. E. and Lamont, J. T. (2014) '*Clostridium difficile* infection: A worldwide disease', *Gut and Liver*, 8(1), pp. 1–6. doi: 10.5009/gnl.2014.8.1.1.
7. Choi, J. H. and Lee, S. Y. (2004) 'Secretory and extracellular production of recombinant proteins using *Escherichia coli*', *Applied Microbiology and Biotechnology*, 64(5), pp. 625–635. doi: 10.1007/s00253-004-1559-9.
8. Grangette, C., Müller-Alouf, H., Geoffroy, M. C., Goudercourt, D., Turneer, M. and Mercenier, A. (2002) 'Protection against tetanus toxin after intragastric administration of two

recombinant lactic acid bacteria: Impact of strain viability and *in vivo* persistence', *Vaccine*, 20(27–28), pp. 3304–3309. doi: 10.1016/S0264-410X(02)00301-8.

9. Hsiung, H. M., Cantrell, A., Luirink, J., Oudega, B., Veros, A. J. and Becker, G. W. (1989) 'Use of bacteriocin release protein in *E. coli* for excretion of human growth hormone into the culture medium', *Bio/Technology*, 7(3), pp. 267–271. doi: 10.1038/nbt0389-267.
10. Islam, K. B. M. S., Fukiya, S., Hagio, M., Fujii, N., Ishizuka, S., Ooka, T., Ogura, Y., Hayashi, T. and Yokota, A. (2011) 'Bile acid is a host factor that regulates the composition of the cecal microbiota in rats', *Gastroenterology*. Elsevier Inc., 141(5), pp. 1773–1781. doi: 10.1053/j.gastro.2011.07.046.
11. Jang, K. H., Seo, J. W., Song, K. B., Kim, C. H. and Rhee, S. K. (1999) 'Extracellular secretion of levansucrase from *Zymomonas mobilis* in *Escherichia coil*', *Bioprocess Engineering*, 21(5), pp. 453–458. doi: 10.1007/s004490050701.
12. Kimoto, H., Nomura, M., Kobayashi, M., Mizumachi, K. and Okamoto, T. (2003) 'Survival of lactococci during passage through mouse digestive tract', *Canadian Journal of Microbiology*, 49(11), pp. 707–711. doi: 10.1139/w03-092.
13. Kliest, S., Miksch, G., Hitzmann, B., Arndt, M., Friehs, K. and Flaschel, E. (2003) 'Optimization of the extracellular production of a bacterial phytase with *Escherichia coli* by using different fed-batch fermentation strategies', *Applied Microbiology and Biotechnology*, 61(5–6), pp. 456–462. doi: 10.1007/s00253-003-1229-3.
14. Kuipers, O. P., De Ruyter, P. G. G. A., Kleerebezem, M. and De Vos, W. M. (1998) 'Quorum sensing-controlled gene expression in lactic acid bacteria', *Journal of Biotechnology*, 64(1), pp. 15–21. doi: 10.1016/S0168-1656(98)00100-X.
15. Li, Y., Kan, Z., You, Y., Gao, X., Wang, Z. and Fu, R. (2015) 'Exogenous transglutaminase improves multiple-stress tolerance in *Lactococcus lactis* and other lactic acid bacteria with glutamine and lysine in the cell wall', *Biotechnology Letters*, 37(12), pp. 2467–2474. doi:

10.1007/s10529-015-1942-x.

16. Le Loir, Y., Azevedo, V., Oliveira, S. C., Freitas, D. A, Miyoshi, A., Bermúdez-Humarán, L. G., Nouaille, S., Ribeiro, L. A, Leclercq, S., Gabriel, J. E., Guimaraes, V. D., Oliveira, M. N., Charlier, C., Gautier, M. and Langella, P. (2005) 'Protein secretion in *Lactococcus lactis*: an efficient way to increase the overall heterologous protein production', *Microbial Cell Factories*, 4(1), p. 2. doi: 10.1186/1475-2859-4-2.
17. Messaoudi, S., Manai, M., Kergourlay, G., Prévost, H., Connil, N., Chobert, J. M. and Dousset, X. (2013) '*Lactobacillus salivarius*: Bacteriocin and probiotic activity', *Food Microbiology*. Elsevier Ltd, 36(2), pp. 296–304. doi: 10.1016/j.fm.2013.05.010.
18. Mierau, I. and Kleerebezem, M. (2005) '10 Years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*', *Applied Microbiology and Biotechnology*, 68, pp. 705–717. doi: 10.1007/s00253-005-0107-6.
19. Mohamadzadeh, M., Duong, T., Sandwick, S. J., Hoover, T. and Klaenhammer, T. R. (2009) 'Dendritic cell targeting of *Bacillus anthracis* protective antigen expressed by *Lactobacillus acidophilus* protects mice from lethal challenge', *Proc Natl Acad Sci U S A*, 106(11), pp. 4331–4336. doi: 10.1073/pnas.0900029106.
20. Park, S. F. and Stewart, G. S. A. B. (1990) 'High-efficiency transformation of *Listeria monocytogenes* by electroporation of penicillin-treated cells (Recombinant DNA; bacterial transformation; shuttle plasmid; field strength)', 94(1), pp. 129–132.
21. Robinson, K., Chamberlain, L. M., Schofield, K. M., Wells, J. M. and Le Page, R. W. F. (1997) 'Oral vaccination of mice against tetanus with recombinant *Lactococcus lactis*', *Nature Biotechnology*, 15(7), pp. 653–657. doi: 10.1038/nbt0797-653.
22. Rosano, G. L. and Ceccarelli, E. A. (2014) 'Recombinant protein expression in *Escherichia coli*: Advances and challenges', *Frontiers in Microbiology*, 5(172), pp. 1–17. doi: 10.3389/fmicb.2014.00172.

23. Ruiz, L., Margolles, A. and Sánchez, B. (2013) ‘Bile resistance mechanisms in *Lactobacillus* and *Bifidobacterium*.’, *Frontiers in Microbiology*, 4(396), pp. 1-8. doi: 10.3389/fmicb.2013.00396.
24. Sleator, R. D., Wemekamp-Kamphuis, H. H., Gahan, C. G. M., Abee, T. and Hill, C. (2005) ‘A PrfA-regulated bile exclusion system (*Bil*E) is a novel virulence factor in *Listeria monocytogenes*’, *Molecular Microbiology*, 55(4), pp. 1183–1195. doi: 10.1111/j.1365-2958.2004.04454.x.
25. Sørensen, H. P. and Mortensen, K. K. (2005) ‘Advanced genetic strategies for recombinant protein expression in *Escherichia coli*’, *Journal of Biotechnology*, 115(2), pp. 113–128. doi: 10.1016/j.jbiotec.2004.08.004.
26. Steidler, L., Hans, W., Schotte, L., Neirynck, S., Obermeier, F., Falk, W., Fiers, W. and Remaut, E. (2000) ‘Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10.’, *Science (New York, N.Y.)*, 289(5483), pp. 1352–1355. doi: 10.1126/science.289.5483.1352.
27. Stern, N. J., Svetoch, E. A., Eruslanov, B. V., Perelygin, V. V., Mitsevich, E. V., Mitsevich, I. P., Pokhilenko, V. D., Levchuk, V. P., Svetoch, O. E. and Seal, B. S. (2006) ‘Isolation of a *Lactobacillus salivarius* strain and purification of its bacteriocin, which is inhibitory to *Campylobacter jejuni* in the chicken gastrointestinal system’, *Antimicrobial Agents and Chemotherapy*, 50(9), pp. 3111–3116. doi: 10.1128/AAC.00259-06.
28. van der Wal, F. J., Luirink, J. and Oudega, B. (1995) ‘Bacteriocin release proteins: mode of action, structure, and biotechnological application’, *FEMS Microbiology Reviews*, 17(4), pp. 381–399. doi: 10.1111/j.1574-6976.1995.tb00221.x.
29. Watson, D., Sleator, R. D., Hill, C. and Gahan, C. G. M. (2008) ‘Enhancing bile tolerance improves survival and persistence of *Bifidobacterium* and *Lactococcus* in the murine gastrointestinal tract.’, *BMC Microbiology*, 8(176), pp. 1-10. doi: 10.1186/1471-2180-8-176.

30. Yang, J., Moyana, T., Mackenzie, S., Xia, Q. and Xiang, J. (1998) 'One hundred seventy-fold increase in excretion of an FV fragment-tumor necrosis factor alpha fusion protein (sFV/TNF-α) from *Escherichia coli* caused by the synergistic effects of glycine and Triton X-100', *Applied and Environmental Microbiology*, 64(8), pp. 2869–2874. doi: PMC106785.
31. Yoon, S. W., Lee, C. H., Kim, J. Y., Kim, J. Y., Sung, M. H. and Poo, H. (2008) '*Lactobacillus casei* secreting α-MSH induces the therapeutic effect on DSS-induced acute colitis in Balb/c mice', *Journal of Microbiology and Biotechnology*, 18(12), pp. 1975–1983. doi: 10.4014/jmb.0800.445.

Thesis conclusions

This thesis focussed on phage therapy against two Gram-positive bacteria, which are very significant in hospital acquired infections, namely *Staphylococcus aureus* and *Clostridium difficile.* Both pathogens are known for their antibiotic resistance, rendering infections difficult to treat. The experimentation in the thesis focussed initially on whole phages and then in more depth on the peptidoglycan hydrolases encoded by the phages. The results presented in the thesis are divided into five chapters, each with its own distinct discussion. The first results chapter (Chapter II**)** focused on the isolation of two new *S. aureus* phages designated B1 and JA1, both of which were members of commercial phage mixtures used in therapeutics at the George Eliava Institute of Bacteriophage, Microbiology and Virology (the Tbilisi Institute). Following their isolation, the host range of these phages on significant *S. aureus* MLST sequence types were compared to phage K. Significantly, both phages B1 and JA1 possessed a much wider host range in comparison to phage K, thus confirming the significance of these phages as therapeutics in the elimination of infectious *S. aureus* and their superiority to the well-known phage K. Between the two new phages, all but two of the twenty-one Irish MLST MRSA isolates were eliminated, an observation which indicates a big improvement compared with the lytic ability of phage K. The next results chapter (Chapter III) details the genome characterization of these two new phages. They had genome sizes of 140,808bp (B1) and 139,484bp (JA1) and the same G+C content of 30.3%. Both genomes were completely annotated and were observed to be organised into modules including the DNA replication/transcription module, structural/morphogenesis module, DNA packaging module and lysis module. Both phages also lack the restriction sites for the common staphylococcal host-encoded endonuclease, *Sau*3a1 and are members of the genus *Kayvirus*. Although both B1 and JA1 were 99% similar to each other, they differed by the presence of four open reading frames with no known function. The next results chapter (Chapter IV) focusses in depth on the endolysin common to all three phages, designated $CHAP_k$. This

enzyme was used for the development of a secretion system, using *L. lactis* as a host bacterium. The successful secretion of $CHAP_k$ by the recombinant *L. lactis* strain was detected by streaking this recombinant *L. lactis* strain on media containing heat-inactivated *S. aureus* cells. It was shown that $CHAP_k$ could successfully reduce the titre of *S. aureus* cells in milk. The next results (Chapter V) stayed on the topic of endolysins, focussing on the *C. difficile* phage encoded amidase enzyme, and its evaluation for the elimination of *C. difficile*, the causitive agent of human pseudomembranous colitis. Similar to $CHAP_k$, this *C. difficile* endolysin was also used in the development of a model secretion system in lactic acid bacteria. Interestingly, a truncated version of this endolysin was generated and was found to be the only form of the endolysin that could be secreted from the bacterium. This deleted derivative of the amidase is 175 amino acids in length and it was assumed that the smaller size would facilitate further applications than the native endolysin. The ideal site of action of the *C. difficile* endolysin is the gastrointestinal tract, and accordingly the last chapter (Chapter VI) focussed on attempting to express and secrete the endolysin in bacteria that survive in the intestine. Despite (a) exploring the development of bile tolerance in *L. lactis* by incorporating a *bilE* gene and (b) expressing the endolysin in the naturally bile resistant *L. salivarius*, an intestinal *E. coli* was found to be the best mode of secretion of this endolysin in a bile-containing environment exhibiting clear elimination of *C. difficile*, thus setting up a possibility for *in-vivo* animal trials.

Appendix