The Discovery and Application of Bacteriophage Receptor Binding Proteins

by

David James Simpson

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Microbiology and Biotechnology

Department of Biological Sciences

University of Alberta

Abstract

Bacteriophages are considered to be the most abundant and potentially the most diverse form of life on earth. Phage receptor binding proteins (RBPs), which allow phages to specifically target their host bacteria, consequently represent a massive diversity of bacterial targeting proteins. These RBPs can bind to bacteria with strong affinity and show considerable resistance to proteases and detergents. In recent years, a number of new technologies have been developed for pathogen detection through the attachment of RBPs to surfaces or beads. This thesis describes new techniques that take advantage of these proteins that act as surrogate antibodies.

Recent years have seen a resurgence in bacteriophage research, often in the form of phage therapy, due to the recent rise in antibiotic resistance. However, bacteriophage RBPs remain difficult to identify based on homology alone due to their considerable diversity. While these proteins share a trimeric structure, they can be very dissimilar on the sequence level. In order to further exploit the use of RBPs, I have developed an assay for discovering RBPs using phage genome expression libraries and protein screens to identify binding partners that recognize the host bacterium. Briefly, the phage DNA is sheared and ligated into an expression library, this library is transferred to a nitrocellulose membrane where the colonies are induced to express the inserts, the cells are then lysed and the membrane is probed with the host bacteria. When the *Salmonella enterica* serovar Typhimurium phage P22 was screened using this assay, Gp9 was the only RBP discovered, confirming previous predictions that this is the sole RBP encoded by this phage. I then examined the *Escherichia coli* O157:H7 typing phage 1 using this assay and identified a previously undescribed RBP, Gp145. This general approach has the potential to assist in the identification of RBPs from other relevant bacteriophages.

In previous studies the bacteriophage P22 RBP has also been shown to be able to reduce colonization of *S.* Typhimurium in chickens. In order to exploit this finding further, the protein was expressed in plants with the aim of creating an inexpensively produced selective antimicrobial feed. An elastin like polypeptide (ELP) tag was added to the P22 RBP to increase expression of the protein in *Nicotiana benthamiana*. This thesis demonstrates that the RBP containing plant extract was capable of capturing *S.* Typhimurium on a nitrocellulose membrane, and moderately reducing the ability of *S.* Typhimurium to colonize chickens.

RBPs bind bacteria with high affinity, to make use of this trait in a diagnostics-based platform, the cellulose binding module CBM9, that enables proteins to bind to paper, was added to Gp9 and Gp145. The N-terminally tagged Gp9 and Gp145 constructs were spotted on paper and are able to capture *S.* Typhimurium and *E. coli* O157:H7, respectively on paper. Gp145 was further characterized and shown to bind to the lipopolysaccharide of *E. coli* O157:H7 and interestingly also binds to *S.* Typhimurium, likely through a protein receptor.

Taken together, these results demonstrate that RBPs represent an exciting new technology for microbial detection and treatment which are inexpensive, easy to use and readily scalable.

Preface

Portions of chapter 1 were previously published in “Simpson, David J, Jessica C Sacher, and Christine M Szymanski. 2015. Exploring the Interactions between Bacteriophage-Encoded Glycan Binding Proteins and Carbohydrates. *Current Opinion in Structural Biology* 34 (October): 69–77. doi:10.1016/j.sbi.2015.07.006.” J. Sacher and I were responsible for preparing and editing the manuscript under the supervision of C. Szymanski.

A version of Chapter 2 was published as “Simpson, David, Jessica Sacher, and Christine Szymanski. 2016. Development of an Assay for the Identification of Receptor Binding Proteins from Bacteriophages. *Viruses* 8 (1): 17. doi:10.3390/v8010017." I conceived, designed and performed the experiments, analyzed the data and wrote the paper. Jessica Sacher was involved in designing some of the experiments. Christine Szymanski supervised this project and contributed the materials and reagents. All the authors contributed to the writing of the paper during the editing process.

A version of Chapter 3 was published as “Miletic, S, D J Simpson, C M Szymanski, M K Deyholos, and R Menassa. 2016. A Plant-Produced Bacteriophage Tailspike Protein for the Control of Salmonella. *Frontiers in Plant Science* 6 (January): 1221." S. Miletic and I, designed and performed the experiments and analyzed the data. I was responsible for all the bacterial work and the design of the animal experiments which were performed with the help of Cory Wenzel. S. Miletic wrote the paper. R. Menassa, C. Szymanski and M. Deyholos supervised this project and contributed the materials and reagents. All the authors contributed to the writing of the paper during the editing process. The chicken studies received ethics approval from the University of Alberta animal care committee.

In Chapter 4, I conceived, designed and performed the experiments, analyzed the data and wrote the chapter. Denis Arutyunov designed the CBM-Gp9 construct. Bernadette Beadle performed the CBM-Gp9 capture experiments (Figures 5A and B). Christine Szymanski supervised this project and contributed the materials and reagents.

Acknowledgements

I would like to thank my supervisor Christine Szymanski for the years of support, guidance, mentorship and advice. It was a pleasure working with you and I will always be proud to say that I trained in your lab.

I would also like to thank Michael Deyholos and Stefan Pukatzki for serving on my committee and being very supportive and active committee members. Michael Deyholos was actively involved in the initial development of the plant-based expression work and provided the constructs for the design of the studies described in Chapter 3 and along with Mary de Pauw trained me in all the skills I needed for my work on plants. The suggestions by both my committee members were always helpful and this thesis is much better than it would have been without you.

I am also grateful to all the members of the Szymanski lab particularly: Denis Arytyunov, Jessica Sacher, Harald Nothaft, Bernadette Beadle and Cory Wenzel. Denis trained me in molecular biology and was always ready to discuss ideas and problems. Jessica was always willing to proofread my writing and has been a good friend. Harald and Cory always made themselves available to discuss experimental techniques and resolve problems I was having with my experiments. Bernadette has always been helpful doing anything that needed to be done to get projects moving along.

Thank you to Tom Hantos, Richard Mah and Cecilia Anders who were the lab coordinators for the courses that I was teaching. Their help both in the courses as well as out of them made my program a much better experience.

I would also like to thank my fellow graduate students past and present on the 4th floor of microbiology and the 6th floor of CCIS, your friendship made the years fly by.

Table of contents

List of tables

List of figures

List of abbreviations

APEC: Avian pathogenic *Escherichia coli*

BSA: bovine serum albumin

CBM9: Carbohydrate binding module 9

CFU: Colony-forming units

E. coli: *Escherichia coli*

ECTP1: *Escherichia coli* typing phage 1

ELP: Synthetic elastin-like polypeptide

ER: Endoplasmic reticulum

FLW: Fresh leaf weight

GFP: Green fluorescent protein

GI tract: Gastrointestinal tract

HFBI: Hydrophobin I from *Trichoderma reesei*

Kn: Kanamycin

LB: Lysogeny broth

LPS: Lipopolysaccharide

PBS: Phosphate buffered saline

RBP: Receptor Binding Protein

S. Enteritidis: *Salmonella enterica* serovar Enteriditis

S. Paratyphi: *Salmonella enterica* serovar Paratyphi

S. Typhimurium: *Salmonella enterica* serovar Typhimurium

TSP: Total soluble proteins (Chapter 3)

TSP: Tailspike protein (Chapters 1,2,4,5)

Chapter 1

Introduction

Portions of this chapter were published:

David J Simpson, Jessica C Sacher, and Christine M Szymanski. 2015. "Exploring the Interactions between Bacteriophage-Encoded Glycan Binding Proteins and Carbohydrates." *Current Opinion in Structural Biology* 34 (October): 69–77. doi:10.1016/j.sbi.2015.07.006.

1.1 General information on bacteriophages

Bacteriophages are the most abundant organisms with an estimated 10^{30} - 10^{31} virions on earth (Suttle, 2005) and therefore have considerable diversity, possibly even as great as their hosts (Rohwer, 2003; Casjens, 2008; Ignacio-Espinoza *et al.*, 2013). Phages reproduce by infecting bacterial cells and subverting the genetic machinery of those cells to create progeny phages. In order to accurately identify their hosts prior to infection, phages encode specific receptor binding proteins (RBPs) that allow them to discriminate between different strains of the same species (Casjens and Molineux, 2012). For the last 3 billion years, phages have co-evolved with their hosts and are considered to be one of the main drivers of bacterial evolution, in fact, the development of DNA may have been a product of the arms race between early cellular life and their viruses where the viruses began using DNA as a modified nucleic acid instead or RNA to avoid the host nucleases and eventually the cells also began using it since it is a more stable genetic material (Forterre *et al.*, 2013). Another aspect of the bacteriophage-bacteria coevolution is the constant alteration of bacteria surface glycans which cover the surface of bacteria and are easily altered. These glycans, like lipopolysaccharide (LPS), may have evolved to mask other less changeable epitopes on the surface of the bacteria such as proteins which also act as phage receptors. However, the bacteriophages have counter evolved which has resulted in the generation of phage glycan binding RBPs spanning an immense range of glycan diversity with binding affinities comparable or better than those of most antibodies (Singh *et al.*, 2011). The utilization of these proteins and glycan binding RBPs is the topic of my research.

1.1.1 The discovery of bacteriophages

The discovery of bacteriophages, like many other scientific discoveries throughout history has involved several researchers working independently and is controversial due to independent discoveries.

The story of the discovery of bacteriophage begins with Ernest Hanbury Hankin who in 1896 found that water filtered from the Ganges and Jumna rivers had antimicrobial properties that worked against *Vibrio cholerae* (Sulakvelidze *et al.*, 2001). Gamaleya, a Russian biologist noticed something similar in his studies on *Bacillus subtilis* two years later (Sulakvelidze *et al.*, 2001). However, neither of these researchers followed up on these findings or hypothesized that the cause may be viruses so the full discovery and study of bacteriophage would not begin until the work of Fredrick Twort and Félix d'Herelle.

Frederick Twort discovered that filtrates from cultivations of glycerinated calf vaccinia could create glassy areas on micrococci that would spread to other colonies of micrococci. At the time he believed viruses could be cultured and was looking for a medium to grow them, Twort thought the phenomena he observed may be caused by extremely small bacteria (Twort, 1915). However, he did not fully appreciate what he was looking at and did not understand that viruses were obligate parasites, while he did acknowledge that the phenomenon could be caused by a virus he believed that viruses were likely a much simpler form of life that pre-existed bacteria and so should be able to grow on sterile media if he could find the right media for growth. Unfortunately, Twort had to discontinue his studies because of a lack of funding (Twort, 1915; Sulakvelidze *et al.*, 2001).

While working in North Africa, Félix d'Herelle discovered an anomaly that caused clear spots on coccobacillus, at the time he was using a coccobacillus to infect locusts in order to control them and he first noticed the clear spots on lawns of the bacterium (Duckworth, 1976). Later, when he

began to study dysentery in cavalry soldiers in France, he discovered the same types of spots on lawns of the shiga bacillus (d'Herelle, 1917). He found that if he filtered cultures with the clear spots and spread the filtrate on a pure culture that culture too would produce the spots, moreover the cultures with the filtrate added did not cause disease in guinea pigs or rabbits. He went on to discover the filtrates from a recovering patient were able to kill the bacteria that had infected the patient. He concluded that he had discovered a virus that infected bacteria (d'Herelle, 1917; Duckworth, 1976).

1.2.2 Important discoveries from bacteriophage

Bacteriophages are some of the simplest organisms known, so it is no surprise that several major discoveries in molecular biology were dependent on the use of bacteriophage. Seymore Benzer was one of the first to realize that advantages of using bacteriophage. His rII system used the bacteriophage T4 to make discoveries about the nature of genes. Benzer crossed various mutants in the rII gene of T4 and was able to discover deletions (Nomura and Benzer, 1961) and point mutations in the genes (Benzer, 1961), this system was also used by Francis Crick to determine that the genetic code was in triplicate (Crick *et al.*, 1961).

Hershey and Chase used bacteriophage to show that DNA was the carrier of genetic information. They did experiments by labelling bacteriophage with either ^{35}S or ^{32}P to label either the protein or the DNA, respectively. They discovered the DNA was injected into the bacterial cells and concluded that it was the genetic material since it was being used to create new virions (Hershey and Chase, 1952).

Bacteriophages have also been used to develop important technologies for molecular biology, for instance, the bacteriophage T4 was analyzed in the first major experiment using sodium dodecyl sulfate -polyacrylamide gel electrophoresis SDS-PAGE. Laemmli used SDS-PAGE to separate out

all the proteins in radiolabelled T4 virions. He was also able to observe cleavage of some of the polypeptides in T4 by pulse labelling infected cells with ^{14}C and by taking samples at different times. This enabled him to see that certain peptides would be cleaved during the infections cycle before lysis of the cells (Laemmli, 1970).

Sanger also employed bacteriophages in the development of his sequencing technique, the first sequenced genome was from φX174, a single stranded DNA genome which is 5,375 nucleotides. The same study also showed that two sets of genes were encoded by the same region using different reading frames (Sanger *et al.*, 1977). Furthermore, Sanger sequencing was developed using T4 DNA polymerase on phage φX174 (Sanger and Coulson, 1975).

1.2.3 Phage therapy

Beginning with d'Herelle there has been interest in utilizing bacteriophages as antimicrobials in order to treat infections. In 1919, d'Herelle first began to experiment using bacteriophage as a treatment for dysentery (Summers, 1999). D'Herelle's initial experiments were successful as were experiments using phages to treat staphylococcal infections (Bruynoghe and Maisin, 1921). Several different companies began to manufacture and sell phage cocktails as therapeutics, but the results were inconsistent. With the development of antibiotics, the Western world abandoned phage therapy and almost all phage research focused on molecular biology research where bacteriophage proved extremely useful. However, in Eastern Europe research on phage therapy continued, particularly in Georgia at the Eliava Institute. d'Herelle worked at the institute briefly, but left after his host and founder of the institute, Giorgi Eliava was executed in one of Stalin's purges (Sulakvelidze *et al.*, 2001). Phage therapy continued to be researched and used in Eastern bloc countries until the end of the cold war when it received renewed interest from Western researchers who were looking for ways of treating antibiotic-resistant bacterial pathogens. Today

there is considerable interest in the West as a number of groups and companies are trying to utilize phages as an accepted part of Western medicine (Kutter *et al.*, 2015).

1.1.4 Cheese industry – disadvantages

Bacteriophages also play a role in the cheese production industry. Since lactic acid bacteria (LAB) are necessary to decrease the pH in the fermentation culture, any contamination by phage that infect LAB can have a serious effect on the batch, often causing a downgrading of product quality (Marcó *et al.*, 2012). Since bacteriophages are ubiquitous, it is impossible to completely eliminate phage from the environment, so cheese producers resort to a number of different strategies to reduce the incidence of phage contamination in their fermentation cultures. These strategies include reducing the opportunity for phage contamination through factory design rotation and improvement of starter cultures (Marcó *et al.*, 2012). The fermentation industry is the main area where bacteriophages are viewed as an impediment to human activity.

1.1.5 Bacteriophages and their life cycle - lytic vs lysogenic

Bacteriophages have two main life cycles, the lytic and the lysogenic. The lytic cycle is the simpler of the two, the phage injects its genetic material into the host cell and then takes over the host cell's machinery and often destroys the host's DNA. The virion controlled cell will produce bacteriophage through a genetically regulated sequence that results in progeny phage being completed at the same time the cell is lysed by phage produced holins and endolysins.

The lysogenic cycle is not used by all phage, in fact, its existence was controversial in the early years of phage research and was not proven until 1950 (Lwoff and Gutmann, 1950; Kutter and Sulakvelidze, 2005). In this cycle, the phage genetic material is injected into the host, but instead of taking over the host cell's machinery, the genome of the phage is incorporated into the host genome (Herskowitz and Hagen, 1980; Calendar, 2006). It is also possible for the phage genome

to exist as a plasmid outside the host genome as is the case with phage P1 of *Eschrichia coli* (Lobocka *et al.*, 2004). The phage DNA is then replicated and inherited along with the rest of the host's genome by each of the daughter cells. The genome of a phage within the host genome is called a prophage and this condition can continue indefinitely, though usually the phage genome will become activated and the phage will enter the lytic cycle (Calendar, 2006). Prophages can carry genes that are beneficial to the host and can change the O-antigen of the infected cell (Allison and Verma, 2000; Bondy-Denomy and Davidson, 2014). A prophage can also grant the host cells immunity from infection by other phages that are closely related to the prophage through genetic regulation (Oppenheim *et al.*, 2005; Bondy-Denomy and Davidson, 2014).

Bacteriophages can also exist in a host cell and be secreted, as in the case of the filamentous phages, without ever killing the cell (Marvin, 1998).

1.1.6 Types of phage

Table 1.1 Types of phage

Family	Order	Genetic Material	Morphology	Members	Model
Bacteriophage					
Myoviridae	*Caudoviridae*	Linear dsDNA	Contractile Tail	1312	T4
Siphoviridae	*Caudoviridae*	Linear dsDNA	Long noncontractile tail	3262	λ
Podoviridae	*Caudoviridae*	Linear dsDNA	Short noncontractile tail	771	T7
Plasmaviridae		Circular dsDNA	Lipid envelope, no capsid	5	L2
Microviridae		Circular ssDNA	Cubic symmetry	38	φX174
Corticoviridae		Circular dsDNA	Two protein shell with a lipid bilayer	3	PM2
Tectiviridae		Linear	Outer protein	19	PRD1

		dsDNA	capsid surrounding lipoprotein vesicle		
Leviviridae		Linear ssRNA	Cubic symmetry	38	MS2
Cystoriridae		Linear dsRNA	Lipid envelope	3	Φ6
Inoviridae		Circular ssDNA	Helical symmetry	66	fd
Archeal Viruses					
Sphaerolipoviridae		dsDNA	Lipoprotein envelope, rodlike shape		
Fuselloviridae		dsDNA	Lemon shape, short spikes		
Lipothrixviridae	*Ligamenvirales*	Linear dsDNA	Helical rodlike shape		
Rudiviridae	*Ligamenvirales*	Linear dsDNA	Helical rigid rod		

Bacteriophages are among the most diverse organism on earth. The majority of bacteriophages discovered so far are double-stranded DNA tailed phages belonging to the order *Caudovirales*. However, there are several other phage families and it is believed that all bacteria have multiple phage predators.

The *Caudoviridae*

The most prevalent group of bacteriophage is the *Caudovirdae* which consists of three families: the *Myoviridae*, the *Siphoviridae* and the *Podoviridae*. These are known as the tailed phages and all have a head-like isosahedral capsid which contains the double stranded DNA and a tail. The tail of these phages consists of a tube, a baseplate, the injection machinery and the RBPs. These phages typically have a primary and secondary receptor (eg. T4), but some only utilize a single receptor (eg. P22, Lambda). Phages will often bind with lower affinity to the surface of the

bacterial cell until they identify the secondary receptor at which point they will begin the DNA ejection process (Kutter and Sulakvelidze, 2005).

Myoviridae

The *Myoviridae* are distinguished by their contractile tail which acts like a syringe to penetrate the bacteria membranes/cell wall and eject their DNA. The *Myoviridae* have some of the largest genomes of any viruses, with some as large as 500 kb (Hatfull and Hendrix, 2011). The tail has two tubes, an inner tube which does not contract, but is surrounded by an outer tube which does contract. The contraction of the outer tube or sheath pushes the inner tube through the bacterial through the cell envelope. Contraction of the outer tube is initiated by conformational changes to the baseplate caused by the binding of the RBP to the host cell (Leiman and Shneider, 2012). The myoviridae phages are among the most studied phages and they include T4.

Siphoviridae

The family *Siphoviridae* have a long noncontractile tail and represent 60% of the known phages on earth (Davidson *et al.*, 2012). *Siphoviridae* tails are made of the tail tube protein which is bound to the capsid by the tail terminator protein. The tail tube is often a hexamer of the tail tube protein and its length is regulated by the tape measure protein. The tails of *Siphoviridae* have a tail tip complex which can be a very narrow tip or a wider baseplate (Davidson *et al.*, 2012). Upon binding of the RBPs to the host cell, a channel is created between the two membranes to allow the DNA to pass through. The phages T5 and Lambda are the two major phage representatives that are investigated in the *Siphoviridae* family.

Podoviridae

The *Podoviridae* are the short-tailed phages, and like the *Siphoviridae* they also have noncontractile tails. *Podoviridae* phage often have receptor binding proteins that have an enzymatic activity to degrade either the LPS O-antigen repeat or the capsule in order to bring the phages closer to the host surface where they can eject their DNA into the host. Unlike the tails of *Siphoviridae* and *Myoviridae*, the tails of *Podoviridae* are assembled directly on the capsid instead of being assembled separately and then attached (Casjens and Molineux, 2012).

Non-tailed phages

Plasmaviridae

The only well-studied phage in this family is the *Acholeplasma* phage L2. The virions are 50-125 nm in diameter and have a lipid-protein membrane. The genome, composed of dsDNA, is 12 kb and encodes 15 ORFs including 4 structural proteins. The lipid component comes from the host cell membrane. This virus exists as a prophage in the genome and reproduces by budding (Maniloff *et al.*, 1994).

Microviridae

The *Microviridae* are a family of ssDNA phage which have icosahedral capsids. The model organism for this family of phages is the *E. coli* phage φX174. The capsids of these phages are composed of a coat protein F which from 12 pentamers and each of these pentamers is decorated with a pentamer of protein G which is the major spike protein responsible for host recognition and attachment (Doore and Fane, 2016).

Corticoviridae

The only known corticoviridae phage currently known is PM2 which was isolated from coastal waters in the 1960s (Espejo and Canelo, 1968). It has a circular 10 079 bp dsDNA genome contained within a protein rich membrane which is itself contained within a protein capsid. The capsid is decorated with the P1 spike protein which is responsible for host recognition and attachment. The virion is 57 nm in diameter and the genome contains 17 ORFs. The virions are released from the *Pseudoaltermonas* host by lysis (Poranen *et al.*, 2015). There is evidence from metagenomics analysis that more of these phages are present in the oceans.

Tectiviridae

Tectiviridae are similar to *Corticoviridae* in structure, both have outer protein capsids which surround a protein rich lipid membrane. Also, both have a protein capsid that is decorated with spike proteins responsible for host recognition and attachment. However, *Tectiviridae* have linear genomes instead of the circular genomes of the *Corticoviridae*. *Tectiviridae* have been found to infect both Gram-positive and Gram-negative hosts and can be lytic or lysogenic (Poranen *et al.*, 2015). The model organism for this family is PRD1.

Leviviridae

Leviviridae have ssRNA genomes and resemble polioviruses (Ackermann, 2009). *Leviviridae* are composed of two genera: *Levivirus* and *Allolevivirus* and these two groups differ in genome length, 3.5 vs. 4.2 kb and in the proteins that comprise the capsid (Calendar, 2006). The virions have no envelope and have an icosahedral shell. They infect the host by utilizing the F-pilus to get their RNA into the host cell (Calendar, 2006).

Cystoviridae

Cystoviridae phages have a dsRNA genome contained within two protein capsid shells which are contained within a lipid envelope. The lipid envelope contains spike proteins which recognize and attach to the host. These phages use their envelopes to fuse with the bacterial membrane as a mechanism of infection. The model organism for this family is φ6, and it has three linear RNA fragments for its genome which encode for 14 proteins (Poranen *et al.*, 2015).

Inoviridae

Inoviridae are ssDNA filamentous phages, they have circular genomes. They are very long, up to 3700 nm with their length corresponding to their genome size. The capsids are made of thousands of copies of the major capsid protein. These viruses do not lyse their hosts but instead establish a chronic infection where the virions are secreted from the host cell (Székely *et al.*, 2016). These phages are represented by the fd phage which is almost identical to the M13 and f1 phages, best known for their use in phage display. These phages use a minor capsid protein PIII found on the end of the phage to bind to their hosts (Calendar, 2006). The CTX phage, which carry the cholera toxin, are also a part of this group.

1.1.7 Phage-like machines

In addition to phages themselves, several types of molecular machines with structural and functional relatedness to phages exist, such as the type VI secretion system (T6SS) (Pukatzki *et al.*, 2007), pyocins (Michel-Briand and Baysse, 2002) and gene transfer agents (GTA) (Lang *et al.*, 2012).

T6SSs are bacterial secretion systems with homology to T4 phage tails, the apparatus assembles inside the bacterial cell and ejects proteins out of the cell. The T6SS has been found in 25% of sequenced Gram-negative bacteria (Salomon and Orth, 2015) with the best-studied examples of

these coming from *Vibrio cholerae* and *Pseudomonas* species. This system is sometimes used as a virulence mechanism, but seems to be more often used as an antimicrobial weapon (Unterweger *et al.*, 2012; Sana *et al.*, 2016).

R-type pyocins are proteinaceous bacterial killing entities naturally expressed by *Pseudomonas* species and that share ancestry with phages (Michel-Briand and Baysse, 2002). These "headless phages" kill bacteria by puncturing the cell membrane and dissipating membrane potential (Michel-Briand and Baysse, 2002). Recent cryo-EM-generated atomic models of the *P. aeruginosa* R2 pyocin have provided unparalleled insights into the membrane-puncturing mechanism (Ge *et al.*, 2015). Pyocin host specificity is imparted by phage-like RBPs, which can be swapped with phage RBPs to change the host killing range of the particles (Williams *et al.*, 2008; Scholl *et al.*, 2009; Gebhart *et al.*, 2015). For example, a *Clostridium difficile* RBP-pyocin fusion was recently shown to specifically prevent intestinal colonization of mice by *C. difficile* spores, which are typically impervious to intervention (Gebhart *et al.*, 2015). Overall, pyocin-RBP fusions represent a highly effective means of utilizing phage RBPs as antibacterial homing devices.

GTAs are phage-like particles that package portions of the bacterial chromosome and inject that DNA into a subsequent target cell. Each of these machines is comprised of proteins with homology to bacteriophages, which in some cases also includes their RBPs. Interestingly, GTAs have been found to be encoded in the genome of diverse species: *Rhodobacter capsulatus*, *Desulfovibrio desulfuricans*, *Brachyspira hydysenteriae* and *Methanococcus voltae* (Lang *et al.*, 2012). Strangely however, the capsids of these phage-like particles are too small to contain all of the genes which encode the GTA, meaning that GTAs cannot transfer their own genes to organisms, in fact, the DNA that is packaged into the GTA appears to be random.

1.1.8 Methods of phage resistance by bacteria

Bacteriophages are the primary predators of bacteria and because of this, bacteria have evolved a number of ways to counteract infection by phages.

The first and most obvious method of stopping phage infection is to simply avoid being infected by the phage in the first place. This can be achieved by either blocking or altering the phage receptor. Altering the receptor to avoid phage binding is an established tactic which is used for both protein and carbohydrate receptors (Drexler *et al.*, 1989; Zaleski *et al.*, 2005). Cells have also been found to produce proteins that block the receptors of their phage predators (Labrie *et al.*, 2010). Bacteria will also use extracellular polysaccharides to restrict access of the phage to the cell surface and some bacteriophages have evolved depolymerases to counter this (Born *et al.*, 2014). Numerous phages will also alter their own receptor when they enter the lysogenic phase in order to stop another related phage from entering the cell (Bondy-Denomy and Davidson, 2014).

Another method of countering bacteriophage infection is having restriction-modification (R-M) systems digest the phage DNA after it enters the host cell. Bacteria will methylate their DNA which is not cleaved by the R-M system, however, the phage DNA which is not modified will be cleaved upon entry into the cell (Labrie *et al.*, 2010).

The clustered regularly interspaced short palindromic repeats (CRISPR) are a newly discovered mechanism of bacteriophage resistance. Briefly, CRISPRs consist of short regions of bacteriophage or plasmid DNA, these sections of DNA are used by the CRISPR-associated proteins to target any DNA that matches that sequence. Bacterial cells acquire new sections of DNA from incoming extracellular DNA giving the bacterial or archaeal cell a form of an

adaptive immune system (Garneau *et al.*, 2010). For a review on this topic see Sorek *et al.* (Sorek *et al.*, 2013).

1.1.9 Receptor binding proteins

The phage must first bind to the target cell through either a protein or carbohydrate receptor (Lindberg, 1973) and often recognizes both a primary and secondary receptor on the host cell. RBPs are found at the ends of phage tails and are responsible for the specific binding to these receptors (Garcia-Doval and van Raaij, 2013). Receptor binding proteins are the host recognition factors of phages. They are also responsible for properly orienting the phage on the host cell, which must occur prior to a successful infection. These proteins can be called tailspikes, tail fibres or spike proteins, depending on the phage. Due to their diversity in host-binding specificity, RBP genes often cannot be recognized in a sequenced genome based solely on homology to already characterized RBPs (Kropinski *et al.*, 2011). Known phage RBPs have been shown to be very stable proteins, displaying high resistance to proteases and detergents, which are often found in their native environments such as the animal gut. This stability is attributed to the fact that RBPs are often rich in β-structures. The intertwined beta-helix architecture of phage P22 trimeric tailspike protein serves as the classic example of RBP spatial arrangement (Singh *et al.*, 2012). Thus, the high stability, specificity, and ease of overexpression make RBPs excellent alternatives to antibodies and ideal tools for the development of new diagnostic technologies.

A modular nature is common among phage RBPs. Typically, the N-terminus serves to connect an RBP to the phage head while the C-terminus is responsible for protein multimerization, receptor binding, and occasionally glycosidase activity (Danner *et al.*, 1993; Stummeyer *et al.*, 2005; Casjens and Molineux, 2012; Born *et al.*, 2014; Javed *et al.*, 2014; Schwarzer *et al.*, 2015). The N- and C-terminal domains of the *Salmonella enterica* serovar Typhimurium phage P22

TSP (Gp9) were crystallized separately in the 1990s (Steinbacher *et al.*, 1994; Steinbacher *et al.*, 1997). However, in 2014, Seul *et al.* introduced a single amino acid change (Y108W) to create a more rigid linker region between the N- and C-termini, enabling a full-length TSP to be crystallized for the first time (Seul *et al.*, 2014). This modular nature of phage RBPs adds to their exploitation potential, as working with smaller proteins can simplify recombinant expression and characterization. For example, crystallization of the *E. coli* phage phi92 RBP (EndoN92) was facilitated by generating the protein without the N-terminus once it was established that this domain was not required for glycan binding or hydrolysis (Stummeyer *et al.*, 2005; Schwarzer *et al.*, 2015).

Of added benefit, phage RBPs have evolved in the presence of the bacteria they recognize, which commonly live in harsh environments such as the mammalian intestine, soil, and water, so these RBPs tend to be resistant to conditions frequently responsible for antibody degradation. This robustness makes phage RBPs ideal for use as "surrogate antibodies" in diagnostics and therapeutics. As will be shown in chapter 3, phage RBPs are readily amenable to recombinant expression in both *E. coli* and plants (Miletic *et al.*, 2016), pointing toward a role for RBPs as accessible tools for glycan analysis. Many RBPs are also capable of cleaving the oligosaccharides they bind, further contributing to their value (Schulz *et al.*, 2010; Born *et al.*, 2014).

Phage host range is dictated by RBPs, which frequently recognize bacterial polysaccharides such as LPS, e.g. *S. Typhimurium* phage P22 (Figure 1.1) (Steinbacher *et al.*, 1996); capsular polysaccharides (CPS), e.g. *Campylobacter* phage F336 (Sørensen *et al.*, 2012) and teichoic acids, e.g. *Listeria* phages A511 (Habann *et al.*, 2014), A118 and P35 (Bielmann *et al.*, 2015). Additionally, phage RBPs can be protein-specific (Mutoh *et al.*, 1978; Drexler *et al.*, 1991), but

protein-specific RBPs will not be discussed. RBPs can also bind much smaller epitopes, such as the *O*-methyl phosphoramidate moiety decorating most *Campylobacter jejuni* CPS (Sørensen *et al.*, 2012).

1.2 Proteins and organisms used in these studies

1.2.1 RBPs from *Caudoviridae*

RBPs come in two forms, tailspikes and tail fibres. Tail fibres are long and thin while tailspikes shorter and thicker (Figure 1.1).

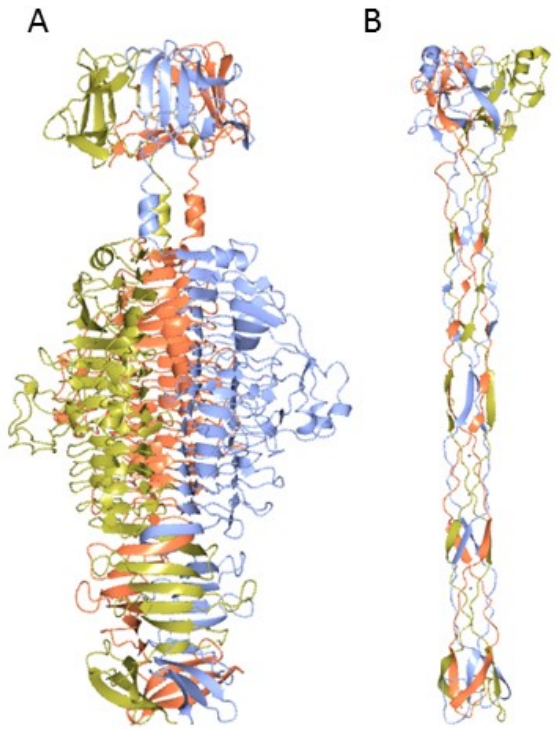


Figure 1.1: Tailspikes and tail fibres. A) Full length P22 Gp9 from 2XC1 (Seul *et al.*, 2014). B) The T4 Gp37 C-terminal domain (785-1026) from PDB 2XGF (Sergio G Bartual *et al.*, 2010).

Among both tailspikes and tail fibres the head binding domains, towards the N-termini, tend to share homology, however, the C-termini which are responsible for host attachment are very diverse.

1.2.2 P22 TSP

The *S.* Typhimurium P22 phage belongs to the *Podoviridae* family of dsDNA phages bearing short non-contractile tails. Its trimeric RBP (Gp9) is called a tailspike protein (TSP) and 18 of these TSP molecules are located at the tip of the phage tail. Each 69-kDa P22 TSP monomer binds to the repeating α-D-mannose-(1,4)- α-L-rhamnose-(1,3)- α-D-galactose O-antigen of the bacterial LPS, and also has endorhamnosidase activity capable of cleaving the polysaccharide to assist in phage binding to the bacterial surface (Gemski and Stocker, 1967; Iwashita and Kanegasaki, 1976; Steinbacher *et al.*, 1996). The first 106 amino acids make up the N-terminal domain of the P22 RBP and are responsible for connecting the RBP to the phage head. Consequently, these amino acids can be removed without influencing the binding or enzymatic activity of the RBP (Miller *et al.*, 1998). The C-terminal domain is responsible for protein trimerization (Steinbacher *et al.*, 1994) as well as for binding and degrading the LPS (Steinbacher *et al.*, 1994). This protein has been studied in great detail in order to understand the nature of P22-LPS interactions, thus making the P22 RBP an ideal choice for proof-of-principle studies used in this thesis.

The P22 phage TSP binds to three common LPS O-antigen serotypes of *S.* Typhimurium, accommodating variations in glucosylation and deoxysugar modification (abequose, tyvelose or paratose) (Figure 1.2). In contrast, the deoxysugar branch is the main target for antibody recognition, thus defining the serotype (Steinbacher *et al.*, 1996). In fact, the first crystal structure of an antibody-carbohydrate complex was reported for a Fab fragment (Se155-4) bound

to the *S.* Typhimurium O-antigen (Cygler *et al.*, 1991). A novel proxy protein electrospray ionization mass spectrometry technique compared Se155-4 versus P22 TSP binding to two O-antigen repeats of *S. enterica* Typhimurium and demonstrated similar association constants: (4.3 ± 1.0) x 10^5 M^{-1} and (2.9 ± 0.9) x 10^5 M^{-1}, respectively at 25ºC (Santoso *et al.*, 2010; El-Hawiet *et al.*, 2012). The crystal structures of the P22 TSP bound to O-antigen repeats of *S. enterica* serovars Typhimurium and Enteritidis (serotypes B and D1 respectively) have long been available (Steinbacher *et al.*, 1996), and the crystal structure of the P22 TSP bound to the third O-antigen type from *S. enterica* serovar Paratyphi A (serotype A) was recently described (Andres *et al.*, 2013).

P22 related tailspikes can also be found on Myoviruses; the Det7 tailspike is an example from this family This tailspike shares 50% homology to the P22 TSP and also binds to LPS of its *S.* Typhimurium host (Walter *et al.*, 2008).

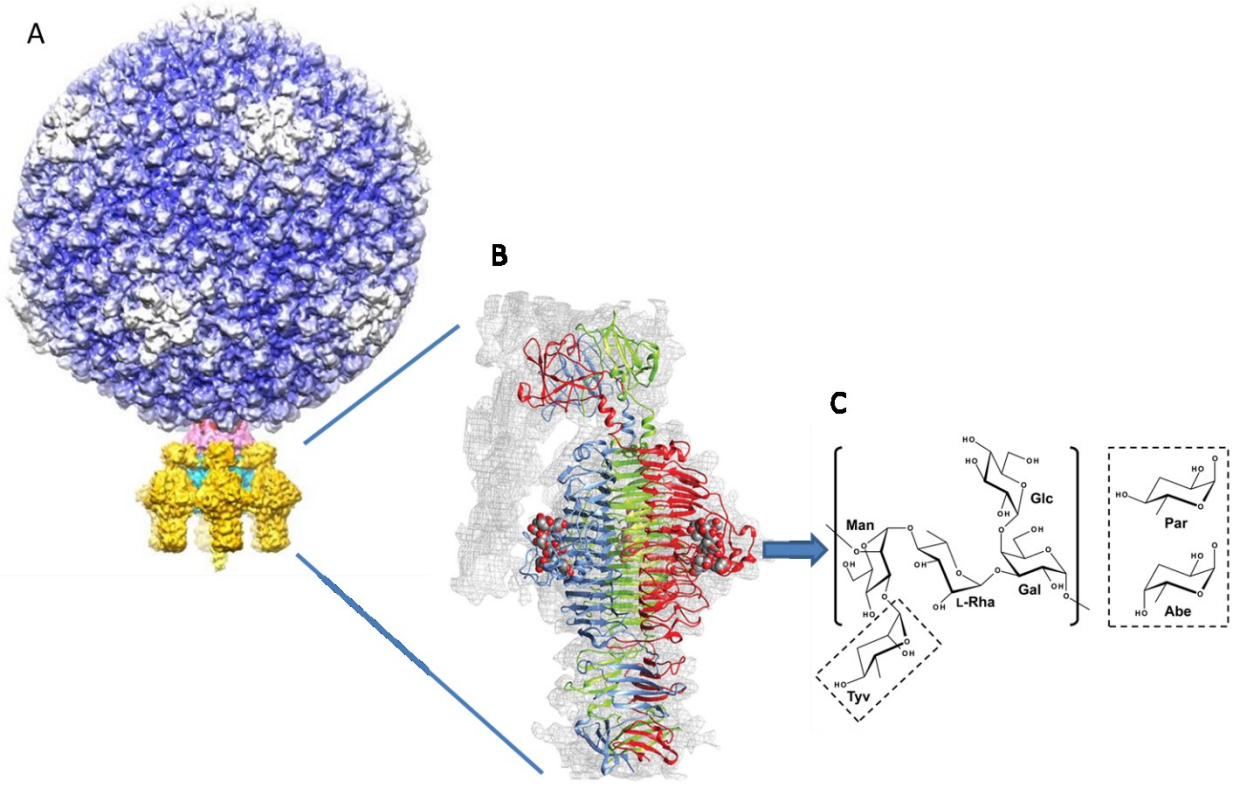


Figure 1.2: Salmonella bacteriophage P22 – a model system for understanding phage-glycan binding protein interactions. (A) Cryo-EM of phage P22 from Tang *et al.* (2011) (Tang *et al.*, 2011). Phage P22 exhibits 6 tail spikes, each comprised of a homotrimer of tailspike proteins (TSPs). The intertwined β-helix architecture of the trimeric P22 TSP serves as a classic example of RBP spatial arrangement (Casjens and Molineux, 2012). On P22 virions, 6 of these TSP trimers are located on the phage tail tip, and each monomer (69 kDa) binds to two repeats of the *Salmonella* O-antigen, which is composed of repeating α-D-mannose-(1,4)-α-L-rhamnose-(1,3)-α-D-galactose, with variations in the α-(1,4)-linked D-glucose at the D-galactose and the 3,6-dideoxyhexose that is linked α-(1,3) to mannose. (B) The full-length P22 TSP was fit into the 7.8 Å resolution cryo-EM map of the P22 virion. The receptor binding portion of the TSP-Y108W trimer fit into the map (Tang *et al.*, 2011), while the linker region and particle-binding domain were distorted for best fit into the map. The coordinates for the O-antigen phage receptors were taken from PDB entry 1TYU (shown as spheres). The P22 TSP structure originally appeared in Seul *et al*. (2014) (Seul *et al.*, 2014), and is reproduced with permission. (C) Structure of the *Salmonella* O-antigen repeat with the possible dideoxy sugars (Tyv: tyvelose, Par: paratose, and Abe: abequose). Figure taken from Simpson et al. 2015 (Simpson *et al.*, 2015).

1.2.3 Well studied tail fibres

There are also several well-studied tail fibres, the T4's Gp37 is a particularly, well studied tail fibre. The T4 Gp37 has homology to the lambda tail fibre (George *et al.*, 1983; Michel *et al.*, 1986) and the lambda tail fibre can be switched into the T4 tail to alter the host range (Montag *et al.*, 1989). Gp37 attaches to the outer membrane porin protein C and LPS while lambda attaches to LamB, a maltose transport protein. Both of these proteins have highly conserved histidines that in Gp37 coordinate iron ions (Sergio G Bartual *et al.*, 2010). The T4 short tail fibre, Gp12,

also has structural homology to Gp37 however it binds to LPS inner core and has a zinc ion coordinated by six histidines (Thomassen *et al.*, 2003). All of these proteins have a homotrimeric beta-helical like structure. Expression of Gp37 requires two chaperones, Gp57 which acts as a general chaperone for tail fibres in T4 (Hashemolhosseini *et al.*, 1996) and Gp38 which is a specific chaperone for Gp37 (Bartual *et al.*, 2010). The crystal structure of the T4 short tail fibre was partially solved in 2003 (van Raaij *et al.*, 2001a; Thomassen *et al.*, 2003) and a fragment of the long tail fibre was solved in 2010 (Sergio G Bartual *et al.*, 2010), whereas the full structure of the P22 TSP has been solved (Seul *et al.*, 2014) and the structure of the binding domain bound to the ligand has also been solved (Steinbacher *et al.*, 1996). As a result, there is considerably more known about tailspikes than about tail fibres.

T7 gp17 is an LPS binding tail fibre with a pyramid like structure (Garcia-Doval and van Raaij, 2012). While this protein formed a beta-helical structure and SDS-resistant trimers it does not bear close homology to T4 tail fibres, demonstrating the diversity in this family of proteins.

1.2.4 RBPs can be altered to change host range

Phages are also capable of altering their host range over successive generations (Kiem *et al.*, 2004), as demonstrated by the *Bordetella* phage BPP-1, which alters its RBP specificity at a high frequency by a reverse transcriptase-based mechanism analogous to that of antibodies (Liu *et al.*, 2002). Alternatively, some phages fine-tune their host range by simultaneous expression of multiple different RBPs. For example, Schwarzer *et al.* (2012) demonstrated that phage phi92 encodes a multivalent tail apparatus made of at least 4 different RBPs (Schwarzer *et al.*, 2012). This “Swiss army knife” binding apparatus allows the phage to infect polysialic acid-encapsulated, poly(α-1,4-glucuronic acid-α-1,4-*N*-acetylglucosamine)-encapsulated and non-encapsulated strains of *E. coli*. This phage can also infect 19 different *Salmonella* serovars,

containing distinct O-antigens, with similar infection efficiencies (Schwarzer *et al.*, 2012). Exploiting host range-expanding strategies of phages could be useful in antimicrobial therapeutics, where resistance development is a concern.

The cytoviridae phage phi6 has also been shown to alter its host range mostly through mutations in it RBP (Ferris *et al.*, 2007). The high mutation rate associated with its RNA-based genome allow it to exploit new niches by infecting closely related pseudomonas species.

1.2.5 Glycosidase activity

In addition to glycan binding activity, many phage RBPs degrade cell surface glycans in order to access the cell membrane following host recognition (Drulis-Kawa *et al.*, 2012). This includes the P22 TSP, which has endorhamnosidase activity that allows it to cleave between the rhamnose-galactose backbone of *S.* Typhimurium LPS (Iwashita and Kanegasaki, 1976; Steinbacher *et al.*, 1996). Additionally, exopolysaccharides (EPS) produced by some bacteria as a means of protection, particularly in biofilms, can be degraded by phage-encoded EPS depolymerases. For instance, the phage L1 RBP (DpoL1) cleaves the galactose backbone of amylovoran, the EPS structure produced by the potato blight pathogen *Erwinia amylovora* (Born *et al.*, 2014). A recent example of a unique RBP-attributed glycosidase involves sialic acid polymer (polySia) degradation. *E. coli* phage K1F encodes EndoNF, the first characterized phage endosialidase that cleaves α-2,8 linked polySia (Stummeyer *et al.*, 2005). Recently, Schwarzer *et al.* (2014) discovered another endosialidase, EndoN92 from phage phi92, which is capable of cleaving the alternating α–2,8 and α-2,9 linked polySia in *E. coli* K92 CPS (Schwarzer *et al.*, 2015). Moreover, when incubated at a 25-fold higher concentration, EndoN92 cleaved the *Neisseria meningitidis* serogroup C α–2,9 linked polySia CPS homopolymer, making it the first example of an α-2,9-specific endosialidase.

1.2.5 Endolysins as phage recognition proteins

Over the past decade, several studies have described the antimicrobial use of endolysins, which are peptidoglycan hydrolases that enable phages to lyse host cells after infection (for a comprehensive review, see (Schmelcher *et al.*, 2012). Endolysins from phages infecting Gram-positive bacteria typically contain a C-terminal cell wall binding domain with specificity for surface glycans such as *N*-acetylglucosamine- or choline-modified teichoic acids, and an N-terminal peptidoglycan-cleaving domain which possesses endopeptidase, *N*-acetylmuramoyl-L-alanine amidase, *N*-acetyl-β-D-muramidase, or *N*-acetyl-β-D-glucosaminidase activity (Schmelcher *et al.*, 2012). While there has been a great deal of research into Gram-positive-specific endolysin treatments, few endolysins can pass through the Gram-negative outer membrane (Morita *et al.*, 2001). This was recently addressed by attaching a cationic peptide to the endolysins to facilitate their transport through the LPS layer (Briers *et al.*, 2014). These constructs, called Artilysins, reduced *Pseudomonas aeruginosa* and *Acinetobacter baumannii* cell numbers by 4-5 orders of magnitude *in vitro* and were effective in a *Caenorhabditis elegans* model *in vivo* (Briers *et al.*, 2014).

The Cell wall Binding Domains (CBDs) of endolysins bind strongly to the Gram-positive cell wall and have been tagged with GFP to label their target cells (Loessner *et al.*, 2002). The CBDs can bind with higher affinity than antibodies (Loessner *et al.*, 2002) and they also can possess greater specificity that antibodies (Schmelcher *et al.*, 2010). CBDs have been attached to magnetic beads to selectively capture Listeria strains (Kretzer *et al.*, 2007), even out of raw milk (Walcher *et al.*, 2010). They have also been used as components of electrochemical sensors (Tolba *et al.*, 2012).

1.2.6 P22 bacteriophage

P22 is a member of the *Podoviridae* family of phages, it is a lysogenic phage with a lambda like DNA arrangement, which is worth noting since lambda is a member of the *Siphoviridae* family. P22 was initially discovered as a lysogen of *S.* Typhimurium LT22 and was initially named PLT-22, but this was later shortened to P22 (Zinder and Lederberg, 1952). The phage attracted interest because it was the first phage capable of transducing DNA between bacteria (Zinder and Lederberg, 1952). P22 binds to the LPS of *S.* Typhimurium and injects its 42 kb genome. It must then enter either the lytic or the lysogenic cycle. If P22 enters the lytic cycle, it will take over the host cell and produce between 300-500 phage particles in approximately one hour (Calendar, 2006). If the phage enters the lysogenic cycle, the DNA will become incorporated into the genome of the host and passed onto all the daughter cells. The decision whether to enter into the lytic or lysogenic cycle seems to be dependent on the multiplicity of infection (MOI) with higher MOI favoring lysogeny (Levine, 1957).

P22 only requires LPS for DNA injection, however, the while the RBP binds to the O-antigen DNA injection will not occur unless the LPS is attached to lipid A (Andres *et al.*, 2010). The phage ejects the proteins gp7, gp16 and gp20 along with the DNA during the ejection process (Israel, 1977) These proteins are thought to form a conduit through the cell envelope to the cytoplasm. Gp26 is the tail needle responsible for penetrating the cell envelope and also serves as a plug, containing the phage DNA inside the capsid and is a homotrimeric fibre (Olia *et al.*, 2007).

The P22 tail machine is a three megadalton complex consisting of 12 copies of Gp1, the portal protein, which form a ring at the base of the capsid (figure 1.2). Twelve copies of Gp4 are attached to the bottom of the Gp1 ring and 6 copies of Gp10 attach to the base of the Gp4 ring.

Three copies of Gp26 block the portal by inserting into the center of the ring complex and 6 trimers of Gp9 are attached to the copies of Gp10 on the outside of the ring (Lander *et al.*, 2009).

1.2.7 *E. coli* O157 typing phage 1

The *E. coli* typing phage 1 is a phage that was isolated in 1987 and first used as a typing phage for *E. coli* O157:H7 (Ahmed *et al.*, 1987). The phage is T4 like and has an 88.5 kb genome (Cowley *et al.*, 2015), for comparison, the T4 phage genome size is 169 kb. This phage not been studied in great detail, which was one of the reasons it was selected for the studies described in this thesis. It is known that this phage is a myoviridae and thus has a long contractile tail. This phage also has tail fibres instead of tail spikes like T4.

1.2.8 *Salmonella enterica* serovar Typhimurium

Salmonella is a Gram-negative nonspore-forming gamma proteobacterium. It belongs to the *Enterobactericae* family along with *E. coli*. Salmonella was first discovered in 1885 by Theobald Smith and Daniel Elmer Salmon, while Smith was the actual discoverer of the bacterium, Salmon took credit and Salmonella was named after him (Fabrega and Vila, 2013). The genus is divided into two species, *S. enterica* and *S. bongori*. *S. enterica* is divided into 6 sub-species with Typhimurium being one. Non-typhoidal Salmonella is estimated to cause one million foodborne illnesses in the United States, with 19,000 hospitalizations and 380 deaths every year (Painter *et al.*, 2013). Salmonella has been linked to 50% of produce related bacterial foodborne outbreaks in Canada (Kozak *et al.*, 2013). *S.* Typhimurium has five pathogenicity islands two of which, SPI1 and SPI2, encode type three secretion systems (Marcus *et al.*, 2000). Salmonella species can also be classified based on the O and H antigens (LPS and flagella respectively). *S.* Typhimurium is an invasive pathogen which uses its first type three secretion system T3SS-1 to cause cytoskeletal rearrangements in the host cell and internalization of the bacterium (Fabrega

and Vila, 2013). The bacterium is then contained within a Salmonella Containing Vacuole (SCV) where the second T3SS is required for the replication of *S.* Typhimurium inside the SCV (Hensel *et al.*, 1995; Ochman *et al.*, 1996).

S. Typhimurium is usually a commensal in avian species and is commonly found in chicken meat (Revolledo *et al.*, 2006; Chambers and Gong, 2011). It has also been shown in numerous studies to be able to contaminate produce by internalizing in plant tissues (Deering *et al.*, 2012). The bacteria can invade the plant tissues through natural openings like stomata or damaged tissues, or they can be brought into the tissues along with contaminated water (Deering *et al.*, 2012).

1.2.9 *Escherichia coli* O157:H7

E. coli was first isolated by Theodor Escherich in 1885. It is a Gram-negative gamma protobacterium. Most *E. coli* are commensals in humans and because they are easy to grow and work with, as a result *E. coli* has become the best-understood organism in the world. However, there are strains of *E. coli* that are pathogenic including Enterohemorrhagic *E. coli* (EHEC). EHEC produces two different Shiga toxins, which are bacteriophage derived AB5 toxins that act to halt protein synthesis in the target cell (Saxena *et al.*, 1989). Both of these shiga toxins bind to globotriosyl ceramide (galactose-α-1-4-galactose-β-1-4-glucosyl ceramide) (Waddell *et al.*, 1988). EHEC causes hemorrhagic colitis and hemolytic uremic syndrome in humans and while there are other serotypes associated with EHEC, O157:H7 is the most important. *E. coli* O157:H7 causes 63,000 foodborne illnesses in the United States, 2,100 hospitalizations and 20 deaths every year (Painter *et al.*, 2013).

Commonly found in cattle guts since they lack the receptor for the shiga toxin, cows are asymptomatic carriers and are a common source of *E. coli* O157:H7 contamination (Lim *et al.*, 2010). Like *S.* Typhimurium, *E. coli* O157:H7 is commonly transmitted to humans through

produce and is able to invade into plant tissues using the same mechanisms (Deering *et al.*, 2012).

1.3 Current state of research

1.3.1 Past uses of receptor binding proteins

It has been appreciated for some time that phage RBPs can be viewed as surrogates to antibodies in traditional assays such as ELISA. However, recently our lab has demonstrated that in addition to being attractive reagents for diagnostics, RBPs can be used as therapeutics against bacterial pathogens since the oral administration of recombinant P22 RBP reduces *S.* Typhimurium colonization of the chicken ceca and bacterial spread to the liver and spleen (Waseh *et al.*, 2010). As expected, P22 RBP is resistant to chicken cecal proteases and is also capable of agglutinating *S.* Typhimurium cells and reducing their motility *in vitro* (Waseh *et al.*, 2010), showing similar agglutination levels as antibodies recognizing the *S.* Typhimurium LPS (Sigurskjold *et al.*, 1991; Cygler *et al.*, 1991; Waseh *et al.*, 2010). P22 RBPs and anti-LPS antibodies also have comparable binding affinities as determined by a recently developed proxy protein electrospray ionization mass spectrometry-based assay, which has shown binding affinities that agree with the results of a previously published fluorescence-quenching assay (El-Hawiet *et al.*, 2012). Our recent developments in RBP immobilization onto chips used in surface plasmon resonance (SPR) spectroscopy analysis of binding interactions (Naidoo *et al.*, 2012) and magnetic microbeads further demonstrate the utility of phage RBPs (Singh *et al.*, 2011; Poshtiban *et al.*, 2013). These techniques have been used to develop rapid pathogen detection tools based on the *S.* Typhimurium phage P22 RBP and on the newly identified *C. jejuni* phage NCTC 12673 RBP-like protein, Gp047 (Kropinski *et al.*, 2011; Tay *et al.*, 2012; Singh *et al.*, 2012; Javed *et al.*, 2013).

1.3.2 Current methods of bacterial detection

Conventional methods of testing for bacteria in foodstuffs involves the growing of bacterial colonies using selective media (Mandal *et al.*, 2011). These techniques can take several days due to the need to grow the cultures. Because of this delay, the contamination of food products may not be discovered until it has already arrived in consumers' homes. There is considerable interest in developing faster techniques (Singh *et al.*, 2013; Poshtiban *et al.*, 2013) including using immunomagnetic beads (Chapman and Ashton, 2003), SPR (Leonard *et al.*, 2004) and quanitative polymerase chain reaction (Wolffs *et al.*, 2006), but so far none have attained the same specificity and sensitivity of the more traditional methods.

1.3.3 Production of therapeutic proteins in plants

Protein production in plants has become an important technology for the overexpression of proteins (Hellwig *et al.*, 2004; Ahmad *et al.*, 2010). The benefits of using plants for production generally revolve around cost since it is simple to create large biomass with plants, production is easily scalable and the infrastructure to grow plants is abundant. Plant protein production also offers a ready delivery method in the case of vaccines and therapeutic proteins/peptides; this technology has been termed molecular pharming (Stoger *et al.*, 2014). Antibodies have been produced in plants since the late 1980s (Hiatt *et al.*, 1989) and insulin was first produced in a transgenic plant in 1998 (Arakawa *et al.*, 1998; Nykiforuk *et al.*, 2006). Bacteriophage proteins have also been produced in plants in the past, T7 RNA polymerase has been transformed into tobacco and rice creating an inducible expression system in those organisms (Nguyen *et al.*, 2004). The site-specific recombinase Cre from the P1 phage has been cloned into and expressed in *Nicotiana benthamiana* creating a genetic tool for use in that organism (Kopertekh *et al.*, 2004). Phage lysins have also been produced in plants; the lysin PlyGBS from the Streptococcus

phage NCTC 11261 was expressed in the chloroplasts of *N. tabacum* (Oey *et al.*, 2009). Not only was this lysin shown to accumulate in great abundance in the chloroplasts, it was also very stable. A bactericidal assay also showed that the plant extract's total soluble protein was effective in killing *Streptococcus pyogenes* (Oey *et al.*, 2009).

1.4 Objectives of these studies

1.4.1 Objective 1: Develop an assay to rapidly identify receptor binding proteins from bacteriophages without available genome sequences

Bacteriophage RBPs are not always identifiable by looking for homology to known RBPs in the genome. In order to overcome this difficulty, I developed an assay for discovering RBPs using phage genome expression libraries and protein screens to identify binding partners that recognize the host bacterium. When phage P22 was screened using this assay, Gp9 was the only RBP discovered, confirming previous predictions that this is the sole RBP encoded by this phage. I then examined the *E. coli* O157:H7 typing phage 1 (ECTP1) in our assay and identified a previously undescribed RBP, Gp145. This general approach has the potential to assist in the identification of RBPs from other bacteriophages (Simpson *et al.*, 2016).

1.4.2 Objective 2: Utilize plant produced RBPs as a therapeutic to reduce *S.* Typhimurium colonization in chickens.

Previous studies have shown that the P22 TSP Gp9 is able to reduce the colonization of *S.* Typhimurium in chicken ceca and reduce the spread of this bacterium to the liver and spleen (Waseh *et al.*, 2010). However, purifying this protein and administering it to chickens is too costly to be used in agriculture. To overcome this difficulty, we cloned *gp9* into *Nicotiana benthamiana* to utilize the benefits of plant protein expression. This also had the advantage of

creating a convenient oral administration route for the protein since the plants could be feed to chickens without any further purification or processing. An elastin-like polypeptide tag was added to the P22 RBP in order to increase expression. The RBP containing plant extract was found to bind to *S.* Typhimurium and demonstrated minor reduction in *S.* Typhimurium colonization of chickens (Miletic *et al.*, 2016).

1.4.3 Objective 3: Attach a cellulose binding module to RBPs to demonstrate that this is an inexpensive and effective detection technology.

Detection of pathogenic bacteria remains an important area of research since the best way to deal with infectious diseases is to stop the organisms from coming into contact with people in the first place. In order to create an inexpensive technology to detect bacteria, the RBPs Gp9 from P22 and Gp145 from ECTP1 were tagged with the cellulose binding module (CBM). CBM9, is a protein tag that binds to β-D-Glcp-(1-4)-β-D-Glcp and is isolated from *Thermotoga maritima* (Notenboom *et al.*, 2001). This module is commonly used as a tag that will allow proteins to bind to paper. This tag, when attached to the N-terminus of the Gp9 or Gp145, is able to capture *S.* Typhimurium or *E. coli* O157:H7 respectively, on paper much better than the RBP alone. Moreover, the CBM-Gp9 construct is able to capture *S.* Typhimurium from natural sources like ceca.

1.4.4 Final statement

Taken together the studies shown in this thesis demonstrate that bacteriophage RBPs are a useful and versatile tool with tremendous potential for exploitation as both therapeutics and diagnostics.

1.5 References

Ackermann, H. (2009) *Bacteriophages*. Humana Press, Totowa, NJ. http://link.springer.com/10.1007/978-1-60327-164-6.

Ahmad, A., O. Pereira, E., J. Conley, A., S. Richman, A., and Menassa, R. (2010) Green Biofactories: Recombinant Protein Production in Plants. *Recent Pat Biotechnol* **4**: 242–259

Ahmed, R., Bopp, C., Borczyk, A., and Kasatiya, S. (1987) Phage-typing scheme for *Escherichia coli* O157:H7. *J Infect Dis* **155**: 806–809 http://www.jstor.org/stable/30105855.

Allison, G.E., and Verma, N.K. (2000) Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol* **8**: 17–23.

Andres, D., Gohlke, U., Broeker, N.K., Schulze, S., Rabsch, W., Heinemann, U., *et al.* (2013) An essential serotype recognition pocket on phage P22 tailspike protein forces *Salmonella enterica* serovar Paratyphi A O-antigen fragments to bind as nonsolution conformers. *Glycobiology* **23**: 486–94 http://www.ncbi.nlm.nih.gov/pubmed/23292517.

Andres, D., Hanke, C., Baxa, U., Seul, A., Barbirz, S., and Seckler, R. (2010) Tailspike interactions with lipopolysaccharide effect DNA ejection from phage P22 particles in vitro. *J Biol Chem* **285**: 36768–36775.

Arakawa, T., Yu, J., Chong, D.K.X., Hough, J., Engen, P.C., and Langridge, W.H.R. (1998) A plant-based cholera toxin B subunit–insulin fusion protein protects against the development of autoimmune diabetes. *Nat Biotechnol* **16**: 934–938 http://www.nature.com/doifinder/10.1038/nbt1098-934.

Bartual, S.G., Garcia-Doval, C., Alonso, J., Schoehn, G., and Raaij, M.J. van (2010) Two-chaperone assisted soluble expression and purification of the bacteriophage T4 long tail fibre protein gp37. *Protein Expr Purif* **70**: 116–121 http://dx.doi.org/10.1016/j.pep.2009.11.005.

Bartual, S.G., Otero, J.M., Garcia-Doval, C., Llamas-Saiz, A.L., Kahn, R., Fox, G.C., and Raaij, M.J. van (2010) Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proc Natl Acad Sci U S A* **107**: 20287–92

Benzer, S. (1961) On the Topography of the Genetic Fine Structure. *PNAS* **47**: 62–509 http://www.pnas.org/content/47/3/403.short.

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

Bondy-Denomy, J., and Davidson, A.R. (2014) When a virus is not a parasite: The beneficial effects of prophages on bacterial fitness. *J Microbiol* **52**: 235–242 http://link.springer.com/10.1007/s12275-014-4083-3.

Born, Y., Fieseler, L., Klumpp, J., Eugster, M.R., Zurfluh, K., Duffy, B., and Loessner, M.J. (2014) The tail-associated depolymerase of *Erwinia amylovora* phage L1 mediates host cell adsorption and enzymatic capsule removal, which can enhance infection by other phage. *Environ Microbiol* **16**: 2168–2180 http://doi.wiley.com/10.1111/1462-2920.12212.

Briers, Y., Walmagh, M., Puyenbroeck, V. Van, Cornelissen, A., Cenens, W., Aertsen, A., *et al.* (2014) Engineered Endolysin-Based "Artilysins" To Combat Multidrug-Resistant Gram-Negative Pathogens. *MBio* **5**: e01379-14-e01379-14

Bruynoghe, R., and Maisin, J. (1921) Essais de thérapeutique au moyen du bacteriophage. *CR Soc Biol* **85**: 120–121.

Calendar, R. (2006) *The Bacteriophages*. 2nd ed., Oxford University Press, Oxford.

Casjens, S.R. (2008) Diversity among the tailed-bacteriophages that infect the Enterobacteriaceae. *Res Microbiol* **159**: 340–348.

Casjens, S.R., and Molineux, I.J. (2012) Short Noncontractile Tail Machines: Adsorption and DNA Delivery by Podoviruses. In *Viral Molecular Machines*. Springer US, pp. 143–179

Chambers, J.R., and Gong, J. (2011) The intestinal microbiota and its modulation for Salmonella control in chickens. *Food Res Int* **44**: 3149–3159.

Chapman, P.A., and Ashton, R. (2003) An evaluation of rapid methods for detecting *Escherichia coli* O157 on beef carcasses. *Int J Food Microbiol* **87**: 279–285.

Cowley, L.A., Beckett, S.J., Chase-Topping, M., Perry, N., Dallman, T.J., Gally, D.L., and Jenkins, C. (2015) Analysis of whole genome sequencing for the *Escherichia coli* O157:H7 typing phages. *BMC Genomics* **16**: 271

Crick, F.H.C., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961) General Nature of the Genetic Code for Proteins. *Nature* **192**: 1227–1232

Cygler, M., Rose, D.R., and Bundle, D.R. (1991) Recognition of a cell-surface oligosaccharide of pathogenic Salmonella by an antibody Fab fragment. *Science* **253**: 442–445.

d'Herelle, F. (1917) Sur un microbe invisible antagoniste des bacilles dysentériques. *CR Acad Sci Paris* **165**: 373–375.

Danner, M., Fuchs, A., Miller, S., and Seckler, R. (1993) Folding and assembly of phage P22 tailspike endorhamnosidase lacking the N-terminal, head-binding domain. *Eur J Biochem* **215**: 653–661

Davidson, A.R., Cardarelli, L., Pell, L.G., Radford, D.R., and Maxwell, K.L. (2012) Long Noncontractile Tail Machines of Bacteriophages. In *Viral Molecular Machines*. pp. 115–142

Deering, A.J., Mauer, L.J., and Pruitt, R.E. (2012) Internalization of *E. coli* O157:H7 and Salmonella spp. in plants: A review. *Food Res Int* **45**: 567–575.

Doore, S.M., and Fane, B.A. (2016) The microviridae: Diversity, assembly, and experimental evolution. *Virology* **491**: 45–55.

Drexler, K., Dannull, J., Hindennach, I., Mutschler, B., and Henning, U. (1991) Single mutations in a gene for a tail fiber component of an *Escherichia coli* phage can cause an extension from a protein to a carbohydrate as a receptor. *J Mol Biol* **219**: 655–663.

Drexler, K., Riede, I., Montag, D., Eschbach, M.-L., and Henning, U. (1989) Receptor specificity of the *Escherichia coli* T-even type phage Ox2: Mutational alterations in host range mutants. *J*

Mol Biol **207**: 797–803.

Drulis-Kawa, Z., Majkowska-Skrobek, G., Maciejewska, B., Delattre, A.-S., and Lavigne, R. (2012) Learning from Bacteriophages - Advantages and Limitations of Phage and Phage-Encoded Protein Applications. *Curr Protein Pept Sci* **13**: 699–722.

Duckworth, D.H. (1976) "Who discovered bacteriophage?". *Bacteriol Rev* **40**: 793–802

El-Hawiet, A., Kitova, E.N., Arutyunov, D., Simpson, D.J., Szymanski, C.M., and Klassen, J.S. (2012) Quantifying ligand binding to large protein complexes using electrospray ionization mass spectrometry. *Anal Chem* **84**: 3867–3870

Espejo, R.T., and Canelo, E.S. (1968) Properties of bacteriophage PM2: A lipid-containing bacterial virus. *Virology* **34**: 738–747

Fabrega, A., and Vila, J. (2013) *Salmonella enterica* Serovar Typhimurium Skills To Succeed in the Host: Virulence and Regulation. *Clin Microbiol Rev* **26**: 308–341

Ferris, M.T., Joyce, P., and Burch, C.L. (2007) High frequency of mutations that expand the host range of an RNA virus. *Genetics* **176**: 1013–1022.

Forterre, P., Filée, J., and Myllykallio, H. (2013) *Origin and Evolution of DNA and DNA Replication Machineries*. Landes Bioscience, Austin, Texas.

Garcia-Doval, C., and Raaij, M.J. van (2012) Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Proc Natl Acad Sci* **109**: 9390–9395

Garcia-Doval, C., and Raaij, M.J. van (2013) Bacteriophage receptor recognition and nucleic Acid transfer. *Subcell Biochem* **68**: 489–518.

Garneau, J.E., Dupuis, M.-V., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., *et al.* (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**: 67–71.

Ge, P., Scholl, D., Leiman, P.G., Yu, X., Miller, J.F., and Zhou, Z.H. (2015) Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states. *Nat Struct Mol Biol* **22**: 377–382 http://www.nature.com/doifinder/10.1038/nsmb.2995.

Gebhart, D., Lok, S., Clare, S., Tomas, M., Stares, M., Scholl, D., *et al.* (2015) A modified R-type bacteriocin specifically targeting *Clostridium difficile* prevents colonization of mice without affecting gut microbiota diversity. *MBio* **6**: 10.1128/mBio.02368-14.

Gemski, P., and Stocker, B.A. (1967) Transduction by bacteriophage P22 in nonsmooth mutants of *Salmonella typhimurium*. *J Bacteriol* **93**: 1588–1597.

George, D.G., Yeh, L.-S.L., and Barker, W.C. (1983) Unexpected relationships between bacteriophage lambda hypothetical proteins and bacteriophage T4 tail-fiber proteins. *Biochem Biophys Res Commun* **115**: 1061–1068

Habann, M., Leiman, P.G., Vandersteegen, K., Bossche, A. Van den, Lavigne, R., Shneider, M.M., *et al.* (2014) Listeria phage A511, a model for the contractile tail machineries of SPO1-related bacteriophages. *Mol Microbiol* **92**: 84–99.

Hashemolhosseini, S., Stierhof, Y.-D., Hindennach, I., and Henning, U. (1996) Characterization of the Helper Proteins for the Assembly of Tail Fibers of Coliphages T4 and lambda. *J Bacteriol* **178**: 6258–6265.

Hatfull, G.F., and Hendrix, R.W. (2011) Bacteriophages and their genomes. *Curr Opin Virol* **1**: 298–303.

Hellwig, S., Drossard, J., Twyman, R.M., and Fischer, R. (2004) Plant cell cultures for the production of recombinant proteins. *Nat Biotechnol* **22**: 1415–1422 http://www.nature.com/doifinder/10.1038/nbt1027.

Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E., and Holden, D.W. (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science* **269**: 400–3 http://www.ncbi.nlm.nih.gov/pubmed/7618105.

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

Herskowitz, I., and Hagen, D. (1980) The Lysis-Lysongeny Decision of Phage lambda: Explict Programming and Responsiveness. *Ann Rev Genet* **14**: 399–445.

Hiatt, A., Caffferkey, R., and Bowdish, K. (1989) Production of antibodies in transgenic plants. *Nature* **342**: 76–78 http://www.nature.com/doifinder/10.1038/342076a0.

Ignacio-Espinoza, C.J., Solonenko, S.A., Sullivan, M.B., Ignacio-Espinoza, J.C., Solonenko, S.A., and Sullivan, M.B. (2013) The global virome: Not as big as we thought? *Curr Opin Virol* **3**: 566–571

Israel, V. (1977) E proteins of bacteriophage P22. I. Identification and ejection from wild-type and defective particles. *J Virol* **23**: 91–7

Iwashita, S., and Kanegasaki, S. (1976) Enzymic and molecular properties of base-plate parts of bacteriophage P22. *Eur J Biochem* **65**: 87–94.

Javed, M.A., Ackermann, H.W., Azeredo, J., Carvalho, C.M., Connerton, I., Evoy, S., *et al.* (2014) A suggested classification for two groups of Campylobacter myoviruses. *Arch Virol* **159**: 181–190.

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**.

Kiem, S., Oh, W.S., Peck, K.R., Lee, N.Y., Lee, J.Y., Song, J.H., *et al.* (2004) Phase variation of biofilm formation in *Staphylococcus aureus* by IS 256 insertion and its impact on the capacity adhering to polyurethane surface. *J Korean Med Sci* **19**: 779–782.

Kopertekh, L., Jüttner, G., and Schiemann, J. (2004) Site-specific recombination induced in transgenic plants by PVX virus vector expressing bacteriophage P1 recombinase. *Plant Sci* **166**: 485–492.

Kozak, G.K., MacDonald, D., Landry, L., and Farber, J.M. (2013) Foodborne Outbreaks in Canada Linked to Produce: 2001 through 2009. *J Food Prot* **76**: 173–183

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

Kropinski, A.M., Arutyunov, D., Foss, M., Cunningham, A., Ding, W., Singh, A., *et al.* (2011) Genome and proteome of *Campylobacter jejuni* bacteriophage NCTC 12673. *Appl Environ Microbiol* **77**: 8265–8271.

Kutter, E., and Sulakvelidze, A. (2005) *Bacteriophages : biology and applications*. CRC Press, Boca Raton.

Kutter, E.M., Kuhl, S.J., and Abedon, S.T. (2015) Re-establishing a place for phage therapy in western medicine. **10**: 685–688

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

Laemmli, U.K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nat Publ Gr* **227**: 680–685.

Lander, G.C., Khayat, R., Li, R., Prevelige, P.E., Potter, C.S., Carragher, B., and Johnson, J.E. (2009) The P22 Tail Machine at Subnanometer Resolution Reveals the Architecture of an Infection Conduit. *Structure* **17**: 789–799

Lang, A.S., Zhaxybayeva, O., and Beatty, J.T. (2012) Gene transfer agents: phage-like elements of genetic exchange. *Nat Rev* **10**: 472–482

Leiman, P.G., and Shneider, M.M. (2012) Contractile Tail Machines of Bacteriophages. Springer US, pp. 93–114 http://link.springer.com/10.1007/978-1-4614-0980-9_5. Accessed May 31, 2016.

Leonard, P., Hearty, S., Quinn, J., and O'Kennedy, R. (2004) A generic approach for the detection of whole *Listeria monocytogenes* cells in contaminated samples using surface plasmon resonance. .

Levine, M. (1957) Mutations in the temperate phage P22 and lysogeny in Salmonella. *Virology* **3**: 22–41

Lim, J.Y., Yoon, J., and Hovde, C.J. (2010) A brief overview of *Escherichia coli* O157:H7 and its plasmid O157. *J Microbiol Biotechnol* **20**: 5–14

Lindberg, A.A. (1973) Bacteriophage receptors. *Annu Rev Microbiol* **27**: 205–241.

Liu, M., Deora, R., Doulatov, S.R., Gingery, M., Eiserling, F.A., Preston, A., *et al.* (2002) Reverse Transcriptase-Mediated Tropism Switching in Bordetella Bacteriophage. *Science (80-)* **295**: 2091–2094

Lobocka, M.B., Rose, D.J., Plunkett, G., Rusin, M., Samojedny, A., Lehnherr, H., *et al.* (2004) Genome of Bacteriophage P1. *J Bacteriol* **186**: 7032–7068

Loessner, M.J., Kramer, K., Ebel, F., and Scherer, S. (2002) C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. *Mol Microbiol* **44**: 335–349

Lwoff, A., and Gutmann, A. (1950) Recherches sur un Bacillus-Megatherium Lysogene. *Ann L Inst PASTEUR* **78**: 711–739.

Mandal, P., Biswas, A., Choi, K., and Pal, U. (2011) Methods for Rapid Detection of Foodborne Pathogens: An Overview. *Am J Food Technol* **6**: 87–102

Maniloff, J., Kampo, G.J., and Dascher, C.C. (1994) Sequence analysis of a unique temperate phage: mycoplasma virus L2. *Gene* **141**: 1–8.

Marcó, M.B., Moineau, S., and Quiberoni, A. (2012) Bacteriophages and dairy fermentations. *Bacteriophage* **2**: 149–158 http://www.ncbi.nlm.nih.gov/pubmed/23275866.

Marcus, S.L., Brumell, J.H., Pfeifer, C.G., and Finlay, B.B. (2000) Salmonella pathogenicity islands: big virulence in small packages. *Microbes Infect* **2**: 145–156.

Marvin, D. (1998) Filamentous phage structure, infection and assembly. *Curr Opin Struct Biol* **8**: 150–158

Michel, C.J., Jacq, B., Arquès, D.G., and Bickle, T.A. (1986) A remarkable amino acid sequence homology between a phage T4 tail fibre protein and ORF314 of phage λ located in the tail operon. .

Michel-Briand, Y., and Baysse, C. (2002) The pyocins of *Pseudomonas aeruginosa*. *Biochimie* **84**: 499–510.

Miletic, S., Simpson, D.J., Szymanski, C.M., Deyholos, M.K., and Menassa, R. (2016) A Plant-Produced Bacteriophage Tailspike Protein for the Control of Salmonella. *Front Plant Sci* **6**: 1221.

Miller, S., Schuler, B., and Seckler, R. (1998) Phage P22 tailspike protein: removal of head-binding domain unmasks effects of folding mutations on native-state thermal stability. *Protein Sci* **7**: 2223–2232.

Montag, D., Schwarz, H., and Henning, U. (1989) A component of the side tail fiber of *Escherichia coli* bacteriophage lambda can functionally replace the receptor-recognizing part of a long tail fiber protein of the unrelated bacteriophage T4. *J Bacteriol* **171**: 4378–84

Morita, M., Tanji, Y., Orito, Y., Mizoguchi, K., Soejima, A., and Unno, H. (2001) Functional analysis of antibacterial activity of *Bacillus amyloliquefaciens* phage endolysin against Gram-negative bacteria. *FEBS Lett* **500**: 56–59

Mutoh, N., Furukawa, H., and Mizushima, S. (1978) Role of lipopolysaccharide and outer membrane protein of *Escherichia coli* K-12 in the receptor activity for bacteriophage T4. *J Bacteriol* **136**: 693–699.

Naidoo, R., Singh, A., Arya, S.K., Beadle, B., Glass, N., Tanha, J., *et al.* (2012) Surface-immobilization of chromatographically purified bacteriophages for the optimized capture of bacteria. *Bacteriophage* **2**: 15–24

Nguyen, H.T., Leelavathi, S., and Reddy, V.S. (2004) Bacteriophage T7 RNA polymerase-directed, inducible and tissue-specific over-expression of foreign genes in transgenic plants. *Plant Biotechnol J* **2**: 301–310

Nomura, M., and Benzer, S. (1961) The nature of the “deletion” mutants in the rII region of phage T4. *J Mol Biol* **3**: 684–692

Notenboom, V., Boraston, A.B., Kilburn, D.G., and Rose, D.R. (2001) Crystal Structures of the Family 9 Carbohydrate-Binding Module from *Thermotoga maritima* Xylanase 10A in Native and Ligand-Bound Forms. *Biochemistry* **40**: 6248–6256

Nykiforuk, C.L., Boothe, J.G., Murray, E.W., Keon, R.G., Goren, H.J., Markley, N.A., and Moloney, M.M. (2006) Transgenic expression and recovery of biologically active recombinant human insulin from *Arabidopsis thaliana* seeds. *Plant Biotechnol J* **4**: 77–85

Ochman, H., Soncini, F.C., Solomon, F., and Groisman, E.A. (1996) Identification of a pathogenicity island required for Salmonella survival in host cells. *Proc Natl Acad Sci U S A* **93**: 7800–4

Oey, M., Lohse, M., Kreikemeyer, B., and Bock, R. (2009) Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *Plant J* **57**: 436–445

Olia, A.S., Casjens, S., and Cingolani, G. (2007) Structure of phage P22 cell envelope–penetrating needle. *Nat Struct Mol Biol* **14**: 1221–1226

Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L., and Adhya, S. (2005) Switches in Bacteriophage Lambda Development. *Annu Rev Genet*. **39**:409-29

Painter, J.A., Hoekstra, R.M., Ayers, T., Tauxe, R. V., Braden, C.R., Angulo, F.J., and Griffin, P.M. (2013) Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 1998–2008. *Emerg Infect Dis* **19**: 407–415

Poranen, M.M., Bamford, D.H., Oksanen, H.M., Poranen, M.M., Bamford, D.H., and Oksanen, H.M. (2015) Membrane-Containing Bacteriophages. In *eLS*. John Wiley & Sons, Ltd, Chichester, UK. pp. 1–11

Poshtiban, S., Javed, M.A., Arutyunov, D., Singh, A., Banting, G., Szymanski, C.M., and Evoy, S. (2013) Phage receptor binding protein-based magnetic enrichment method as an aid for real time PCR detection of foodborne bacteria. *Analyst* **138**: 5619–5626

Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D., and Mekalanos, J.J. (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc Natl Acad Sci U S A* **104**: 15508–15513.

Raaij, M.J. van, Schoehn, G., Burda, M.R., and Miller, S. (2001) Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. *J Mol Biol* **314**: 1137–1146.

Revolledo, L., Ferreira, A.J.P., and Mead, G.C. (2006) Prospects in Salmonella Control: Competitive Exclusion, Probiotics, and Enhancement of Avian Intestinal Immunity. *J Appl Poult Res* **15**: 341–351

Rohwer, F. (2003) Global phage diversity. *Cell* **113**: 141.

Salomon, D., and Orth, K. (2015) Type VI secretion system. *Curr Biol* **25**: R265–R266.

Sana, T.G., Berni, B., and Bleves, S. (2016) The T6SSs of *Pseudomonas aeruginosa* Strain

PAO1 and Their Effectors: Beyond Bacterial-Cell Targeting. *Front Cell Infect Microbiol* **6**: 61

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., *et al.* (1977) Nucleotide sequence of bacteriophage φX174 DNA. *Nature* **265**: 687–695

Sanger, F., and Coulson, A.R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* **94**: 441–448

Santoso, Y., Joyce, C.M., Potapova, O., Reste, L. Le, Hohlbein, J., Torella, J.P., *et al.* (2010) Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc Natl Acad Sci U S A* **107**: 715–720.

Saxena, S.K., O'Brien, A.D., and Ackerman, E.J. (1989) Shiga toxin, Shiga-like toxin II variant, and ricin are all single-site RNA N-glycosidases of 28 S RNA when microinjected into Xenopus oocytes. *J Biol Chem* **264**: 596–601

Schmelcher, M., Donovan, D.M., and Loessner, M.J. (2012) Bacteriophage endolysins as novel antimicrobials. *Future Microbiol* **7**: 1147–71

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. *Appl Environ Microbiol* **76**: 5745–5756

Scholl, D., Cooley, M., Williams, S.R., Gebhart, D., Martin, D., Bates, A., and Mandrell, R. (2009) An Engineered R-Type Pyocin Is a Highly Specific and Sensitive Bactericidal Agent for the Food-Borne Pathogen *Escherichia coli* O157:H7. *Antimicrob Agents Chemother* **53**: 3074–3080

Schulz, E.C., Schwarzer, D., Frank, M., Stummeyer, K., Mühlenhoff, M., Dickmanns, A., *et al.* (2010) Structural Basis for the Recognition and Cleavage of Polysialic Acid by the Bacteriophage K1F Tailspike Protein EndoNF. *J Mol Biol* **397**: 341–351.

Schwarzer, D., Browning, C., Stummeyer, K., Oberbeck, A., Mühlenhoff, M., Gerardy-Schahn, R., and Leiman, P.G. (2015) Structure and biochemical characterization of bacteriophage phi92 endosialidase. *Virology* **477**: 133–143

Schwarzer, D., Buettner, F.F.R., Browning, C., Nazarov, S., Rabsch, W., Bethe, A., *et al.* (2012) A Multivalent Adsorption Apparatus Explains the Broad Host Range of Phage phi92: a Comprehensive Genomic and Structural Analysis. *J Virol* **86**: 10384–10398

Seul, A., Müller, J.J., Andres, D., Stettner, E., Heinemann, U., and Seckler, R. (2014) Bacteriophage P22 tailspike: structure of the complete protein and function of the interdomain linker. *Acta Crystallogr Sect D Biol Crystallogr* **70**: 1336–1345

Sigurskjold, B.W., Altman, E., and Bundle, D.R. (1991) Sensitive titration microcalorimetric study of the binding of Salmonella O-antigenic oligosaccharides by a monoclonal antibody. *Eur J Biochem* **197**: 239–246

Simpson, D., Sacher, J., and Szymanski, C. (2016) Development of an Assay for the Identification of Receptor Binding Proteins from Bacteriophages. *Viruses* **8**: 17

Simpson, D.J., Sacher, J.C., and Szymanski, C.M. (2015) Exploring the interactions between bacteriophage-encoded glycan binding proteins and carbohydrates. *Curr Opin Struct Biol* **34**: 69–77

Singh, A., Arutyunov, D., McDermott, M.T., Szymanski, C.M., and Evoy, S. (2011) Specific detection of *Campylobacter jejuni* using the bacteriophage NCTC 12673 receptor binding protein as a probe. *Analyst* **136**: 4780–4786.

Singh, A., Arutyunov, D., Szymanski, C.M., and Evoy, S. (2012) Bacteriophage based probes for pathogen detection. *Analyst* **137**: 3405–3421.

Singh, A., Poshtiban, S., and Evoy, S. (2013) Recent Advances in Bacteriophage Based Biosensors for Food-Borne Pathogen Detection. *Sensors* **13**: 1763–1786

Sorek, R., Lawrence, C.M., and Wiedenheft, B. (2013) CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea. *Annu Rev Biochem* **82**: 237–266

Sørensen, M.C.H., Alphen, L.B. van, Fodor, C., Crowley, S.M., Christensen, B.B., Szymanski, C.M., *et al.* (2012) Phase Variable Expression of Capsular Polysaccharide Modifications Allows *Campylobacter jejuni* to Avoid Bacteriophage Infection in Chickens. *Front Cell Infect Microbiol* **2**: 11

Steinbacher, S., Baxa, U., Miller, S., Weintraub, A., Seckler, R., and Huber, R. (1996) Crystal structure of phage P22 tailspike protein complexed with Salmonella sp. O-antigen receptors. *Proc Natl Acad Sci U S A* **93**: 10584–10588

Steinbacher, S., Miller, S., Baxa, U., Budisa, N., Weintraub, A., Seckler, R., and Huber, R. (1997) Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 A, fully refined structure of the endorhamnosidase at 1.56 A resolution, and the molecular basis of O-antigen recognition and cleavage. *J Mol Biol* **267**: 865–80

Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. *Science* **265**: 383–386.

Stoger, E., Fischer, R., Moloney, M., and Ma, J.K.-C. (2014) Plant Molecular Pharming for the Treatment of Chronic and Infectious Diseases. *Annu Rev Plant Biol* **65**: 743–768

Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R., and Ficner, R. (2005) Crystal structure of the polysialic acid–degrading endosialidase of bacteriophage K1F. *Nat Struct Mol Biol* **12**: 90–96

Sulakvelidze, A., Alavidze, Z., and Morris, J.G. (2001) Bacteriophage Therapy. *Antimicrob Agents Chemother* **45**: 649–659

Summers, W.C. (1999) *Félix d'Herelle and the origins of molecular biology*. Yale University Press.

Suttle, C. a (2005) Viruses in the sea. *Nature* **437**: 356–361.

Székely, A.J., Breitbart, M., Ackermann, H.-W., Adams, M., Lefkowitz, E., King, A., *et al.* (2016) Single-stranded DNA phages: from early molecular biology tools to recent revolutions in

environmental microbiology. *FEMS Microbiol Lett* **363**: 1–7

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

Tay, L.-L.L., Huang, P.-J.J., Tanha, J., Ryan, S., Wu, X., Hulse, J., and Chau, L.-K.K. (2012) Silica encapsulated SERS nanoprobe conjugated to the bacteriophage tailspike protein for targeted detection of Salmonella. *Chem Commun (Camb)* **48**: 1024–6

Thomassen, E., Gielen, G., Schütz, M., Schoehn, G., Abrahams, J.P., Miller, S., and Raaij, M.J. van (2003) The Structure of the Receptor-binding Domain of the Bacteriophage T4 Short Tail Fibre Reveals a Knitted Trimeric Metal-binding Fold. *J Mol Biol* **331**: 361–373.

Tolba, M., Ahmed, M.U., Tlili, C., Eichenseher, F., Loessner, M.J., Zourob, M., *et al.* (2012) A bacteriophage endolysin-based electrochemical impedance biosensor for the rapid detection of Listeria cells. *Analyst* **137**: 5749

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

Unterweger, D., Kitaoka, M., Miyata, S.T., Bachmann, V., Brooks, T.M., Moloney, J., *et al.* (2012) Constitutive Type VI Secretion System Expression Gives *Vibrio cholerae* Intra- and Interspecific Competitive Advantages. *PLoS One* **7**: e48320

Waddell, T., Head, S., Petric, M., Cohen, A., and Lingwood, C. (1988) Globotriosyl ceramide is specifically recognized by the *Escherichia coli* verocytotoxin 2. *Biochem Biophys Res Commun* **152**: 674–679

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–Ba. *Foodborne Pathog Dis* **7**: 1019–1024

Walter, M., Fiedler, C., Grassl, R., Biebl, M., Rachel, R., Hermo-Parrado, X.L., *et al.* (2008) Structure of the Receptor-Binding Protein of Bacteriophage Det7: a Podoviral Tail Spike in a Myovirus. *J Virol* **82**: 2265–2273

Waseh, S., Hanifi-Moghaddam, P., Coleman, R., Masotti, M., Ryan, S., Foss, M., *et al.* (2010) Orally Administered P22 Phage Tailspike Protein Reduces Salmonella Colonization in Chickens: Prospects of a Novel Therapy against Bacterial Infections. *PLoS One* **5**: e13904.

Williams, S.R., Gebhart, D., Martin, D.W., and Scholl, D. (2008) Retargeting R-type pyocins to generate novel bactericidal protein complexes. *Appl Environ Microbiol* **74**: 3868–3876.

Wolffs, P.F.G., Glencross, K., Thibaudeau, R., and Griffiths, M.W. (2006) Direct Quantitation and Detection of Salmonellae in Biological Samples without Enrichment, Using Two-Step Filtration and Real-Time PCR. *Appl Environ Microbiol* **72**: 3896–3900

Zaleski, P., Wojciechowski, M., and Piekarowicz, A. (2005) The role of Dam methylation in phase variation of *Haemophilus influenzae* genes involved in defence against phage infection.

Microbiology **151**: 3361–3369

Zinder, N.D., and Lederberg, J. (1952) Genetic exchange in Salmonella. *J Bacteriol* **64**: 679–699

Chapter 2

Development of an assay for the identification of receptor binding proteins from bacteriophages

A version of this chapter was published:

Simpson, D., Sacher, J., and Szymanski, C. (2016) Development of an Assay for the Identification of Receptor Binding Proteins from Bacteriophages. *Viruses* **8**: 17

Author Contributions: I conceived, designed and performed the experiments, analyzed the data and wrote the paper. Jessica Sacher was involved in designing some of the experiments. Christine Szymanski supervised this project and contributed the materials and reagents. All the authors contributed to the writing of the paper during the editing process.

2.1 Introduction

Bacteriophage receptor binding proteins (RBPs) have recently been developed into a number of tools that make use of their high specificity and robustness (Singh *et al.*, 2012). These technologies include diagnostics involving RBPs bound to surfaces (Singh *et al.*, 2009) or to beads (Tay *et al.*, 2012; Poshtiban *et al.*, 2013) for the selective capture of bacteria. Moreover, RBPs have also been employed as therapeutics that reduce specific bacterial colonization *in vivo*, either alone (Waseh *et al.*, 2010) or as part of a larger molecular machine (Ritchie *et al.*, 2011). As diagnostics, RBPs offer several advantages over other technologies, such as antibodies, including greater stability, ligand specificity, and affinity even against carbohydrate epitopes which are not typically recognized effectively by most antibodies (Novotny *et al.*, 2013; Simpson *et al.*, 2015). The importance of developing technologies for the rapid identification of pathogens cannot be understated, as food contaminated with bacterial pathogens presents both a public health concern and a serious financial liability for producers. Also, rapid methods of bacterial detection can help surmount challenges in diagnosing bacterial agents of disease prior to progression and/or spread of the infection. RBPs therefore represent a highly versatile, effective and much-needed technology for bacterial pathogen detection.

Phages recognize their hosts through RBP binding to a specific receptor on the host cell surface. RBPs are also responsible for properly orienting the phage onto the host cell, which must occur prior to a successful infection (Casjens and Molineux, 2012). These proteins can be called tailspikes, tail fibres or spike proteins. Due to their diversity in host-binding specificity, RBP genes often cannot be recognized in a sequenced genome based solely on homology with already characterized RBPs (Kropinski *et al.*, 2011), since even if the phage RBPs share structural homology, they do not tend to share sequence homology (Freiberg *et al.*, 2003; Barbirz *et al.*,

2008). Known phage RBPs have been shown to be very stable proteins, displaying high resistance to proteases and detergents. These properties are inherent to phage RBPs, presumably since these proteins have evolved to be functional in harsh native environments such as the intestinal tract. On a molecular level, this stability is generally attributed to the fact that RBPs are often trimers rich in β-structures that are intertwined in a β-helical architecture. The phage P22 trimeric tailspike, Gp9, serves as a classic example of RBP spatial arrangement (please refer to Figure 2.4B) (Singh *et al.*, 2012). Overall, the high stability, specificity, and ease of recombinant overexpression make RBPs excellent alternatives to antibodies and ideal tools for the development of new diagnostic technologies.

In addition, RBPs are not limited to phages, but can also be found as components of phage-like molecular machines such as pyocins (Michel-Briand and Baysse, 2002) and gene transfer agents (GTAs) (Lang *et al.*, 2012). Pyocins, also known as headless phages, bind to a target cell and create a channel across the membrane(s), disrupting the transmembrane potential and killing the cell (Ge *et al.*, 2015). These machines are currently being exploited as antibacterial agents and have had their target range altered by substituting different RBPs with their host recognition moieties, giving known pyocins novel target ranges (Gebhart *et al.*, 2015). GTAs are non-replicative phage-like particles that package portions of the bacterial chromosome and inject that DNA into a neighboring target cell (Lang *et al.*, 2012). Both of these machines rely on their RBPs for binding to and identification of their host, and thus present other natural reservoirs for proteins of this class.

Considerable research has been done on RBPs, especially the P22 RBP, which has been used as a model for understanding protein folding and phage infection (Casjens and Molineux, 2012). However, to date there is no rapid method for identifying phage RBPs, even within sequenced

genomes, although, lambda has been shown to be a useful tool for annotating phage genomes. This is a result of the fact that RBPs are difficult to identify based on homology, since RBPs each bind to distinct receptors. As well, phage DNA has been shown to be challenging to sequence, likely as a result of hyper-modification of the phage DNA bases (Warren, 1980), further contributing to challenges associated with RBP identification. In order to overcome these difficulties, we developed a method to rapidly identify RBPs in the absence of sequence information. We generated *E. coli* gene expression libraries expressing randomly sheared fragments of phage DNA and then screened these libraries for colonies producing a gene product able to bind to the host organism. We tested this technique with two phages: *Salmonella enterica* serovar Typhimurium phage P22, which belongs to the *Podoviridae* family, and *E. coli* O157:H7 typing phage 1, belonging to the *Myoviridae* family (Ahmed *et al.*, 1987; Cowley *et al.*, 2015). In both cases, the tailspike or putative tail fibre was successfully identified using the screen and all identified genome fragments contained coding regions for putative RBPs. Together these results demonstrate the utility of an assay for accurate RBP identification from phages belonging to two of the three major families of tailed phages.

2.2 Materials and methods

2.2.1 Strains and phages

The bacteriophage P22 and its propagating strain *S.* Typhimurium ATCC 19585 were obtained from the American Type Culture Collection. The *E. coli* O157:H7 typing phage 1 and its propagating strain C-8299-83 were obtained from the Félix d'Herelle Centre.

2.2.2 Construction of gene expression libraries

Phage DNA was extracted using the Phage DNA Isolation Kit (Norgen Biotek Corp) according to the manufacturer's instructions. The phage DNA was then nebulized using a nebulizer kit (Invitrogen). Briefly, 10 µg DNA was diluted in 1 ml of shearing buffer (10% glycerol, 10 mM Tris, 1mM EDTA, pH 8) and nebulized under 10 psi of nitrogen gas for 30 seconds. The sheared DNA was precipitated overnight at -20°C in isopropanol, 40 µg/ml glycogen and 0.3 M sodium acetate (pH 4.8) and then blunt-end repaired using T4 DNA polymerase (Thermo Scientific) according to the manufacturer's instructions. The DNA fragments were precipitated again in isopropanol and then ligated into the *Eco*RV site of pET30a using T4 DNA ligase (NEB). Next, 20 ng of pET30a was ligated with 50 ng, 100 ng or 200 ng of sheared phage DNA in a 20 µl reaction volume overnight at 16 °C, and the ligation mixture was transformed into chemically competent Top10 cells (Invitrogen). The resulting colonies were then screened for the presence of inserted DNA into pET30a. If at least 2/10 verified plasmids contained inserts between 1 kB and 3 kB, the entire library was plated out for single colonies, pooled together and plasmid DNA was extracted using a plasmid mini prep kit (Thermo Scientific). The library was then transformed into *E. coli* BL21 (DE3) cells.

2.2.3 Screen for identification of receptor binding protein-encoding genes

E. coli BL21 (DE3) cells containing the phage genomic expression library described above were plated for isolated colonies (150 to 200 colonies per plate) on LB containing 25 µg/ml kanamycin (Figure 2.1). A nitrocellulose membrane cut to fit each plate was placed on top of the colonies, gently pressed down and removed after 5-10 sec. Membranes were then placed colony side up on LB (25 µg/ml kanamycin + 0.4 mM IPTG) agar plates and incubated overnight at 30°C for induction of protein expression. The colony lifts were then incubated colony side up for 1 h at room temperature on top of a similarly sized Whatman filter paper that was pre-saturated

with bacterial protein extraction reagent (B-PER, Thermo Scientific) containing 1x protease inhibitor cocktail (Roche), DNAse I (1 U/ml) (Thermo Scientific) and lysozyme (500 µg/ml) (Sigma). The colony lifts were blocked for 1 h in 5% skim milk (Difco), washed in PBST, and remnants of lysed colonies were then gently wiped off using a Kimwipe. The membranes were then exposed to UV light for 15 min to kill any remaining intact *E. coli* cells. The lifts were incubated overnight in 500 mM NaCl at 4°C and then incubated in a suspension of the host organism at 10^8 CFU/ml in blocking solution for 30 min at room temperature with gentle shaking. The membranes were then washed 3 X 10 min in PBST, dried by dabbing with a paper towel and then placed colony side up on LB agar plates. Plates were incubated overnight at room temperature to allow bound bacterial cells to grow into visible colonies on the membrane. The colonies that appeared on the membranes were matched to their corresponding colonies on the master plates containing the original *E. coli* expression library and the corresponding colonies were patched to a new plate. Patched colonies were used in a subsequent colony lift and carried through a second round of the assay to confirm the results. The genomic expression library colonies that demonstrated a positive result in the second round were grown up, frozen down and the inserts in the plasmids were sequenced.

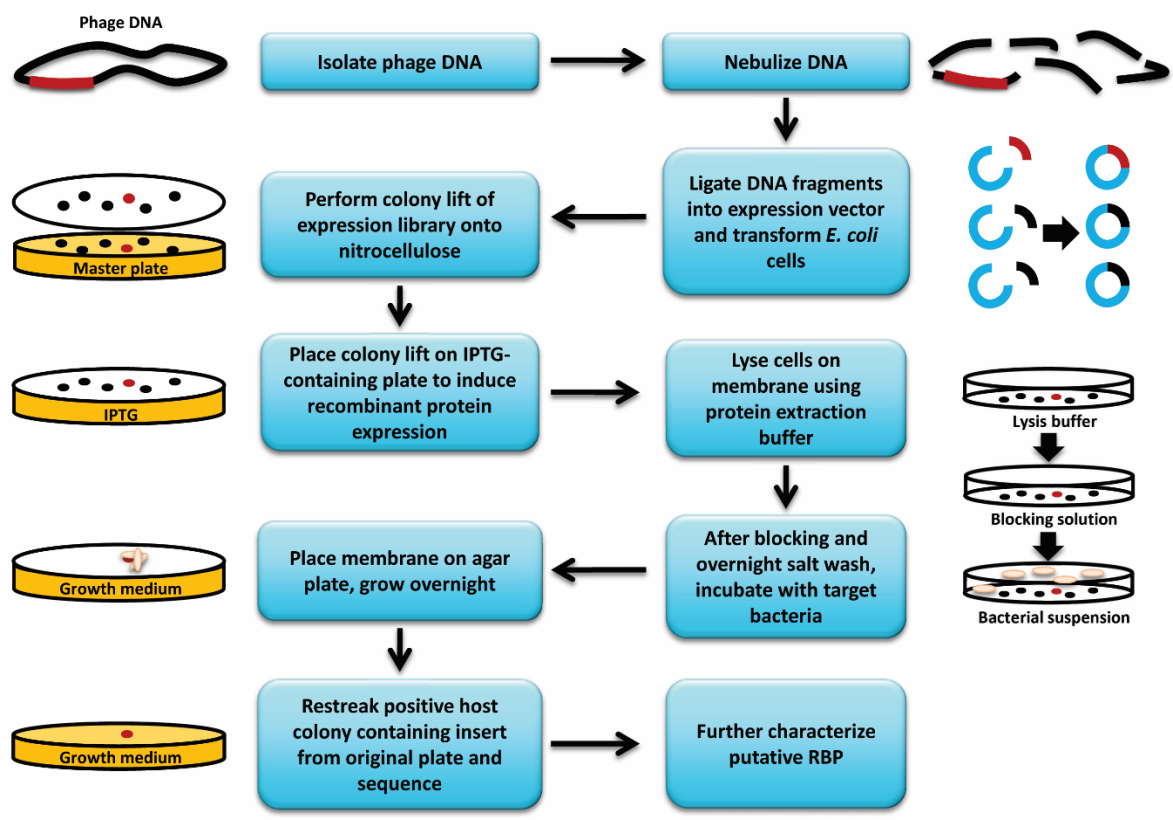


Figure 2.1 Overview of RBP discovery assay. A flowchart and graphical representation summarizing the methodology used to discover new phage receptor binding proteins.

2.2.4 Imaging of colony lifts with green fluorescence protein (GFP) expressing *S.* Typhimurium

Since it is difficult to photograph bacterial colonies growing on nitrocellulose membranes, colonies of *E. coli* BL21 (DE3) containing either the empty vector (pET30a) or plasmid expressing the Gp9 insert 16155-19050 that was identified in this study (pET30a-Gp9) were subjected to the assay conditions and then probed with *S.* Typhimurium expressing GFP (pWM 1007 (Miller *et al.*, 2000)) to generate a visual image of the assay. After overnight growth, the membranes were imaged with a FujiFilm FLA-5000 system using the 473 nm laser at 400 V for excitation and LPB (Y510) filter for emission.

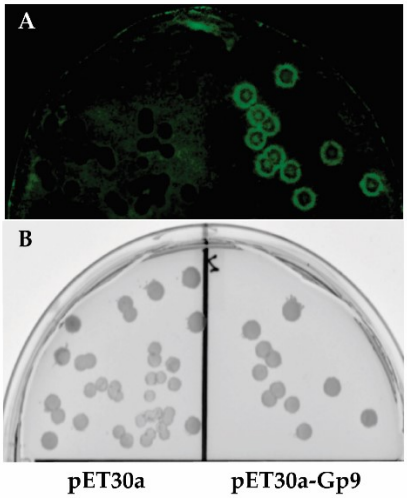


Figure 2.2 Image of colony lift. A) Fluorescent image of GFP expressing *S.* Typhimurium growing on a nitrocellulose membrane. B) The master plate of the colony lift showing *E. coli* colonies from the empty vector control (pET30a) on the left and *E. coli* colonies expressing the positive insert (16155-19050) from the assay on the right..

2.2.5 Protein modelling

All putative RBPs were sent to the Phyre2 server for protein modeling (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) (Kelley *et al.*, 2015)**.** Table 1 shows the top result for each of the genes discovered in the assay and all results with a confidence greater than 90%.

2.3 Results

The *S.* Typhimurium phage P22 library produced seven unique gene fragments that bound to *S.* Typhimurium. All fragments contained an intact copy of the known tailspike gene, *gp9*. Approximately 5700 colonies were screened, with 70% containing inserts. Eleven positive colonies were obtained, all of which contained the known tailspike gene *gp9*, and 7 of these

contained unique inserts. All inserts were in the correct orientation for expression from the vector promoter (Figure 2.3A).

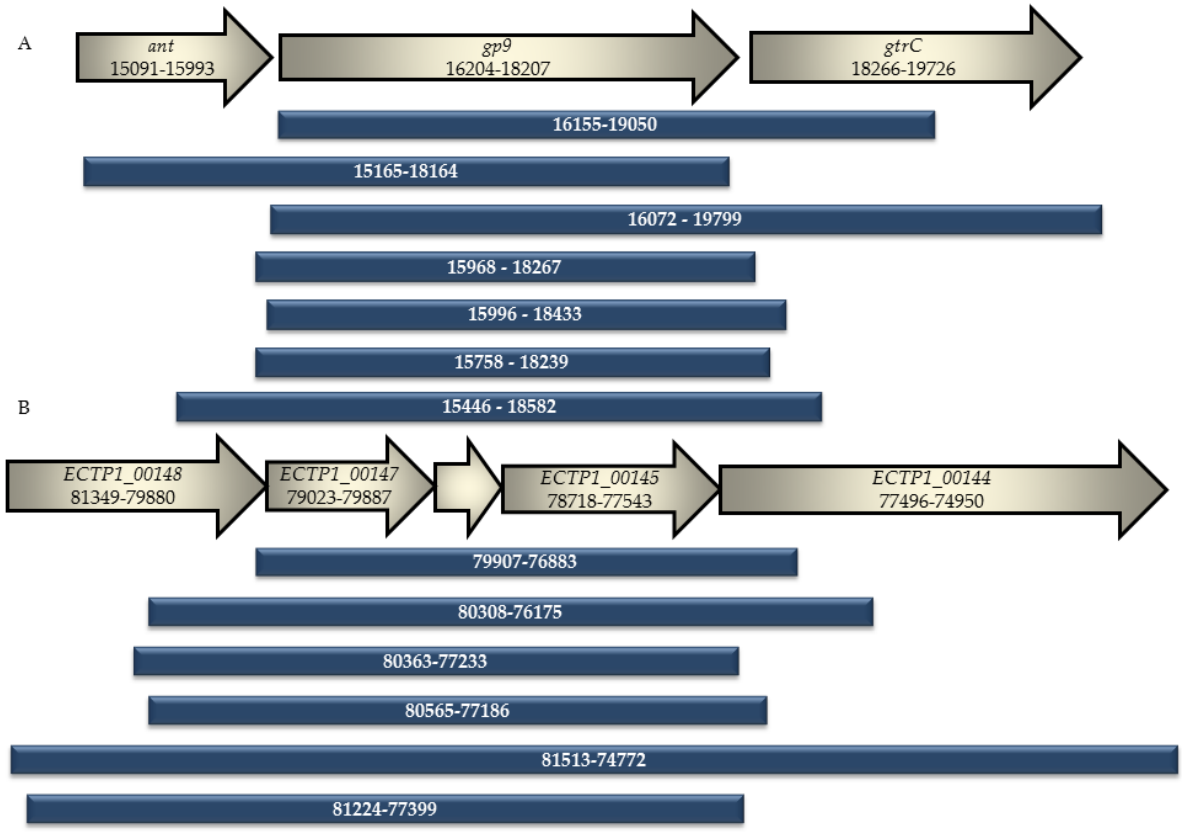


Figure 2.3 Alignment of sequenced fragments from positive colonies compared to GenBank annotated sequences. A) The inserts identified in this study following screening of the P22 phage genomic library are shown alongside the corresponding portion of the phage P22 genome. *Ant* encodes the antirepressor associated with the prophage switch from lysogenic to lytic mode, *gp9* encodes the known P22 phage tailspike protein and *gtrC* encodes the O-antigen conversion glucosyltransferase. All inserts discovered using the assay contained *gp9.* The numbers correspond to the location in the P22 phage genome BK000583.1. **B)** The inserts identified in this study following screening of the *E. coli* O157 typing phage 1 library are shown alongside the

corresponding portion of the phage 1 genome. All inserts contained the unknown genes *ECTP1_00145*, *ECTP1_00146* and *ECTP1_00147*. *ECTP_00145* was predicted to be a tail fibre by protein modelling. *ECTP1_00146* is in between *ECTP1_00147* and *ECTP1_00145* and is unmarked in the figure. The numbers correspond to the location in the phage ECTP1 genome KP869100.1

The *E. coli* O157:H7 typing phage 1 library produced six unique gene fragments that bound to *E. coli* O157. Approximately 4000 colonies were screened, with 30% containing inserts. Eight colonies showed positive binding, with 6 of these containing unique inserts. All inserts were in the correct orientation for expression from the vector promoter (Figure 2.3B). The observed difference in the percentage of vectors containing inserts, as compared to the P22 phage library, is likely due to using cleaner DNA when the P22 library was constructed.

Table 2.1 Results of Phyre analysis of the putative ECTP1 RBPs isolated from the assay.

Protein	Phyre Prediction	Identity	Confidence
ECTP1_00144	Baseplate structural protein Gp10 (T4 phage)	19%	99.2%
	Receptor binding tip Gp37 (T4)	27%	99.1%
ECTP1_00145	Baseplate structural protein Gp10 (T4)	43%	100%
	Receptor binding tip Gp37 (T4)	35%	99.9%
	Short tail fibre Gp12 (T4)	16%	98.9%

ECTP1_00146	Restriction endonuclease	100%	28%
ECTP1_00147	Restriction endonuclease	20%	56.3%
ECTP1_00148	Baseplate structural protein Gp6	12%	94.2%

All inserts contained the genes ECTP1_00145-00147. ECTP1_00146 and ECTP1_00147_were both predicted with low confidence to be restriction endonucleases. *ECTP1_00145,* which when translated and submitted to Phyre2 analysis was modeled to have homology to the Gp37 tail fibre from the T4 phage and to the baseplate protein Gp10 from T4. ECTP1_00144 and ECTP1_00148 were also submitted to the Phyre server, but while ECTP1_00144 was also predicted to be a tail fibre, it was only isolated in one insert. Even then, this protein was isolated together with ECTP1_00145, so it is unlikely that ECTP1_00144 is the RBP in this case (Table 2.1). Since Gp37 is the known RBP for T4, we further compared Gp37 and ECTP1_00145 (Figure 2.4) and found limited homology between the two. BlastP analysis showed 36% sequence identity between Gp37 and ECTP1_00145 over 23% sequence coverage, with an E-value of $6e^{-29}$ (Figure 2.4A). Interestingly, most of the homology was at the C-terminal end of the proteins, the region known to be involved in receptor binding in Gp37. However, the Gp37 domain responsible for binding, between amino acids 931 and 966 (Bartual *et al.*, 2010), was less conserved with ECTP1_00145. The difference in this particular region may be indicating that the two proteins have different receptors. Interestingly, the 14 histidine residues found in this region, which form pairs needed for iron coordination, were all conserved but one in ECTP1_00145 (indicated by asterisks in Figure 2.4A). Although only the 24.6 kDa binding tip of the full 109.2 kDa Gp37 protein has been crystallized, other putative Gp37-like tail fibres have

been described that more closely resemble ECTP1_00145 in size and sequence (Table 2.2). As expected, there was even less homology between T4 Gp37 and P22 Gp9, with 35% identity over 19% sequence coverage and an E-value of 0.12. These differences are further reflected in the crystal structures of the proteins (Figures 2.4B and 2.4C). This lack of homology between the two proteins used in this assay demonstrates the capacity of the assay to identify very different RBPs, in this case a tailspike and a tail fibre which are encoded by two separate families of tailed phages.

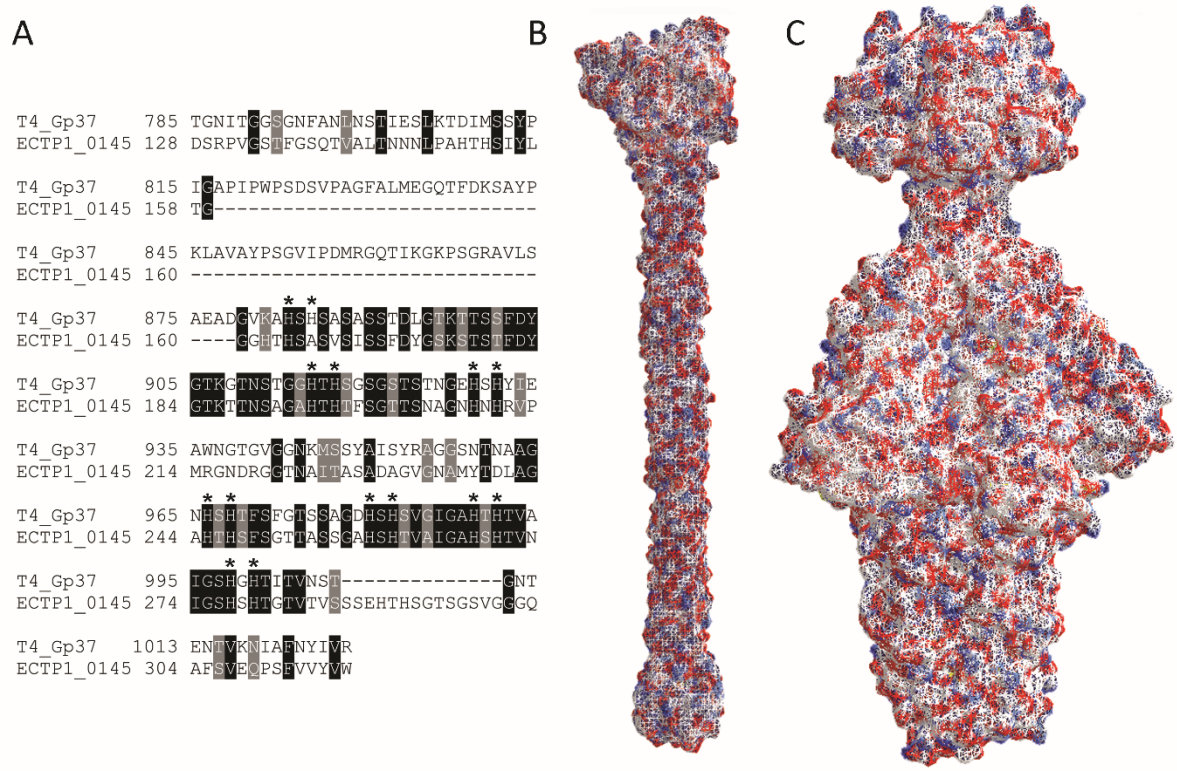


Figure 2.4 Alignment and structure of RBPs detected in this study. A) BlastP alignment of phage T4 Gp37 receptor binding tip (785-1026 from 1026 amino acids in the full length protein), and the corresponding homologous region of ECTP1_00145 (128-317 from 391 amino acids),

showing conserved amino acids in light grey and identical amino acids in dark grey. Histidines involved in iron coordination are marked with an asterisk. **B)** Protein structure of the T4 Gp37 receptor binding tip trimer (785-1026, PDB ID: 2XGF) (Bartual *et al.*, 2010). **C)** Protein structure of the P22 Gp9 trimer (PDB ID: 2XC1) (Seul *et al.*, 2014).

Table 2.2 Blastx analysis of ECTP1_00145

Description	Max score	Total score	Query cover	E value	Ident	Accession
putative tail fibre protein GP37 [*Escherichia* phage wV8]	731	731	99%	0	100%	YP_002922865.1
putative tail fibre protein [*Escherichia* phage HY02]	721	721	99%	0	98%	AIK67887.1
capsid and scaffold protein [*Escherichia* phage EC6]	711	711	99%	0	97%	YP_009151296.1
tail fibre protein [*Escherichia* phage vB_EcoM-VpaE1]	575	575	99%	0	78%	YP_009147343.1
tail fibre protein [*Escherichia* phage vB_EcoM_AYO145A]	574	574	99%	0	78%	AKC04905.1
putative tail fibre protein [*Escherichia* phage JH2]	555	555	99%	0	76%	AGR48475.1
long tail fibre protein [*Salmonella* phage HB-2014]	551	551	99%	0	75%	YP_009146321.1

putative tail fibre protein [*Staphylococcus* phage SA1]	550	550	99%	0	75%	YP_009168663.1
tail fibre protein [*Salmonella* phage Mushroom]	547	547	99%	0	76%	AJF40589.1
tail fibre protein [*Salmonella* phage SBA-1781]	546	546	99%	0	76%	AFU63437.1
putative tail fibre protein GP37 [*Salmonella* phage FelixO1]	545	545	99%	0	76%	NP_944921.1
putative tail fibre protein [Enterobacteria phage UAB_Phi87]	534	534	99%	0	77%	YP_009150171.1
putative tail fibre protein [*Escherichia* phage phiSUSP2]	505	505	99%	9.00E-175	70%	ALH47185.1
baseplate wedge subunit and tail pin [*Citrobacter* phage Moogle]	498	498	99%	8.00E-172	68%	YP_009145700.1
tail fibre protein [*Citrobacter* phage Michonne]	493	493	99%	7.00E-170	68%	AKU44009.1
putative tail fibre protein [*Escherichia* phage phiSUSP1]	486	486	99%	2.00E-167	68%	ALH47055.1
tail fibre protein [*Erwinia* phage phiEa104]	295	295	99%	6.00E-93	47%	YP_004327045.1

putative tail fibre protein [*Erwinia* phage phiEa21-4]	294	294	99%	2.00E-92	47%	YP_002456092.1
tail fibre [*Erwinia* phage vB_EamM-M7]	279	279	99%	9.00E-87	45%	AEJ81298.1
Putative tail fibre protein [*Erwinia* phage phiEa116]	278	278	99%	3.00E-86	44%	CCA66288.1
putative tail fibre protein [*Salmonella* phage FSL SP-010]	478	478	68%	6.00E-166	86%	AGF88729.1
putative tail fibre protein [*Salmonella* phage FSL SP-107]	478	478	67%	7.00E-166	86%	AGF89440.1
phage tail fibre protein [*Escherichia coli*]	110	110	55%	9.00E-23	39%	WP_001613294.1
hypothetical protein WR17_12490 [*Escherichia coli*]	104	104	55%	4.00E-22	36%	KLH27863.1
phage tail protein [*Escherichia coli*]	106	106	55%	1.00E-21	37%	WP_032250185.1
phage tail protein [*Escherichia coli*]	103	103	55%	1.00E-20	39%	WP_024216547.1
phage tail fibre protein [*Escherichia coli*]	103	103	55%	2.00E-20	39%	WP_001559528.1
phage tail protein [*Escherichia coli*]	103	103	55%	2.00E-20	39%	WP_053881948.1
phage tail protein [*Escherichia coli*]	103	103	55%	2.00E-20	39%	WP_047088576.1

phage tail protein [*Escherichia coli*]	103	103	55%	2.00E-20	39%	WP_040090438.1
hypothetical protein [*Escherichia coli*]	102	102	55%	2.00E-20	39%	WP_038820201.1
gp37, tail fibre [Enterobacteria phage SV14]	100	100	50%	7.00E-21	35%	CAA91919.1
gp37 long tail fibre distal subunit [Enterobacteria phage vB_EcoM-VR7]	104	104	50%	1.00E-20	37%	YP_004063946.1
putative tail fibre protein [*Escherichia* phage phiSUSP2]	117	117	48%	1.00E-24	39%	ALH47186.1
putative tail fibre protein [*Escherichia* phage phiSUSP1]	113	113	48%	2.00E-23	38%	ALH47056.1
hypothetical protein [*Escherichia coli*]	109	109	47%	4.00E-24	41%	WP_038428222.1
putative tail fibre protein [*Staphylococcus* phage SA1]	119	119	47%	2.00E-25	43%	YP_009168664.1
putative tail fibre protein [*Salmonella* phage HB-2014]	117	117	47%	8.00E-25	43%	YP_009146322.1
putative tail fibre protein [Enterobacteriaphage UAB_Phi87]	113	113	47%	2.00E-23	42%	YP_009150170.1
tail fibre protein [*Citrobacter* phage	112	112	47%	2.00E-23	41%	AKU44008.1

Michonne]						
tail fibre protein [*Citrobacter* phage Mordin]	112	112	47%	2.00E-23	41%	ALA06874.1
tail fibres protein [*Salmonella* phage SBA-1781]	111	111	47%	6.00E-23	41%	AFU63436.1
tail fibres protein [*Salmonella* phage SPT-1]	111	111	47%	6.00E-23	41%	AFU63552.1
Putative tail fibre protein [*Salmonella* phage FelixO1]	110	110	47%	2.00E-22	41%	NP_944923.1
tail fibre protein [*Salmonella* phage Mushroom]	109	109	47%	2.00E-22	41%	AJF40588.1
hypothetical protein AN659_02340 [*Enterobacter cloacae* subsp. cloacae]	107	107	29%	2.00E-23	56%	KPQ85085.1
short-chain fatty acid transporter [*Escherichia coli*]	106	106	29%	3.00E-23	57%	AHM43535.1
hypothetical protein UH30_23970 [*Escherichia coli*]	104	104	29%	7.00E-23	56%	KJA01516.1
phage tail fibre repeat family protein [*Escherichia coli* P0304816.12]	103	103	29%	2.00E-22	53%	ENF33961.1
short-chain fatty acid transporter [*Escherichia coli*]	104	104	29%	2.00E-22	53%	WP_024208059.1

hypothetical protein ABF63_18835 [*Enterobacter cloacae*]	104	104	29%	2.00E-22	52%	KLQ77673.1
phage tail fibre repeat family protein [*Escherichia coli*]	104	104	29%	3.00E-22	53%	WP_001620747.1
hypothetical protein SS01_22115 [*Enterobacter cloacae*]	104	104	29%	3.00E-22	52%	KJO25483.1
short-chain fatty acid transporter [*Escherichia coli*]	101	101	29%	3.00E-22	53%	WP_024246402.1
hypothetical protein [*Escherichia coli*]	108	108	29%	3.00E-22	56%	WP_044081857.1
phage tail fibre repeat family protein [*Escherichia coli* 2730350]	105	105	29%	7.00E-22	53%	ENA86895.1
phage tail protein [*Escherichia coli*]	107	107	29%	7.00E-22	53%	WP_047090223.1
phage tail fibre repeat family protein [*Escherichia coli* P0304816.7]	103	103	29%	9.00E-22	53%	ENF61429.1
phage tail protein [*Escherichia coli*]	106	106	29%	1.00E-21	53%	WP_042969324.1
short-chain fatty acid transporter [*Escherichia coli*]	103	103	29%	1.00E-21	53%	KLH91524.1
hypothetical protein SS19_06215 [*Enterobacter cloacae*]	102	102	29%	1.00E-21	52%	KJL56021.1

hypothetical protein SR76_12125 [*Enterobacter cloacae*]	101	101	29%	2.00E-21	51%	KJP72894.1
phage tail fibre repeat family protein [*Escherichia coli* 1-392-07_S4_C3]	99.4	99.4	29%	3.00E-21	54%	KEN81314.1
hypothetical protein LI66_16940 [*Enterobacter cloacae*]	100	100	29%	4.00E-21	51%	AJB84181.1
hypothetical protein SG79_17960 [*Enterobacter cloacae*]	100	100	29%	4.00E-21	51%	KJX44172.1
phage tail fibre repeat family protein [*Escherichia coli* 2-005-03_S1_C3]	100	100	29%	4.00E-21	54%	EZK13318.1
hypothetical protein ABF71_22225 [*Enterobacter cloacae*]	100	100	29%	5.00E-21	51%	KLQ26844.1
phage tail fibre repeat family protein [*Escherichia coli* 2-156-04_S3_C1]	100	100	29%	6.00E-21	53%	KDW02205.1
phage tail protein [*Escherichia coli]*	105	105	26%	5.00E-21	57%	WP_032191999.1
phage tail protein [*Escherichia coli*]	105	105	26%	6.00E-21	57%	WP_000108489.1
phage tail protein [*Escherichia coli*]	105	105	26%	6.00E-21	57%	WP_052923054.1
phage tail protein *[Escherichia coli*]	104	104	26%	6.00E-21	57%	WP_054626495.1

phage tail fibre repeat family protein [*Escherichia coli* 1-182-04_S4_C2]	100	100	26%	6.00E-21	56%	EZJ42148.1
side tail fibre protein from bacteriophage origin [*Escherichia coli* HM605]	101	101	26%	7.00E-21	56%	EIL71065.1
phage tail fibre repeat family protein [*Escherichia coli* 2846750]	97.8	97.8	26%	9.00E-21	56%	EMZ59450.1
phage tail fibre repeat family protein [*Escherichia coli* BCE032_MS-12]	101	101	26%	9.00E-21	56%	ENB38924.1
long tail fibre protein p37 [*Shigella flexneri* K-315]	102	102	26%	9.00E-21	56%	EIQ21292.1
tail protein *[Escherichia coli*]	104	104	26%	1.00E-20	56%	WP_033550501.1
hypothetical protein [*Escherichia coli*]	104	104	26%	1.00E-20	56%	WP_001534332.1
hypothetical protein [*Escherichia coli*]	104	104	26%	1.00E-20	56%	WP_021563149.1
short-chain fatty acid transporter [*Escherichia coli*]	104	104	26%	1.00E-20	56%	WP_000104716.1
phage tail fibre repeat family protein [*Escherichia coli* 2872000]	101	101	26%	2.00E-20	55%	EMV62153.1

phage tail fibre repeat family protein [*Escherichia coli* BCE030_MS-09]	102	102	26%	4.00E-20	56%	ENB28302.1
phage tail protein [*Escherichia coli*]	102	102	26%	4.00E-20	56%	WP_053889027.1
phage tail fibre protein [*Escherichia coli*]	102	102	26%	4.00E-20	56%	WP_032287570.1
phage tail fibre repeat family protein [*Escherichia coli* 3-073-06_S4_C3]	102	102	26%	4.00E-20	56%	KDZ79081.1
phage tail protein [*Escherichia coli*]	102	102	26%	4.00E-20	56%	WP_053902476.1
phage tail fibre repeat family protein [*Escherichia coli* 3-073-06_S4_C1]	102	102	26%	4.00E-20	56%	KDU34276.1
hypothetical protein [*Escherichia coli*]	97.1	97.1	26%	4.00E-20	52%	WP_001624807.1
short-chain fatty acid transporter [*Salmonella enterica subsp. enterica* serovar Infantis]	99	99	26%	8.00E-20	55%	KNB28030.1
hypothetical protein KV39_06595 [*Escherichia coli*]	95.1	95.1	26%	8.00E-20	52%	KGA87659.1
hypothetical protein AE17_01327 [*Escherichia coli* UCI 58]	95.9	95.9	26%	8.00E-20	52%	KDG89942.1

short-chain fatty acid transporter [*Escherichia coli*]	96.7	96.7	26%	9.00E-20	50%	KFH78014.1
hypothetical protein [*Escherichia coli*]	99	99	26%	1.00E-19	53%	WP_016233435.1
prophage tail fibre domain-containing protein [*Escherichia coli* K71]	99	99	26%	1.00E-19	53%	KRR57681.1
hypothetical protein Q457_17870 [*Escherichia coli* ATCC BAA-2196]	96.3	96.3	26%	1.00E-19	50%	ETI74864.1
short-chain fatty acid transporter [*Escherichia coli* O26:H11 str. 2010C-4819]	95.5	95.5	26%	1.00E-19	50%	EZG56210.1
phage tail protein [*Escherichia coli*]	100	100	26%	1.00E-19	50%	WP_045894452.1
short-chain fatty acid transporter [*Escherichia coli* O123:H11 str. 2009C-3307]	95.5	95.5	26%	1.00E-19	50%	EZE17057.1
phage tail protein [*Escherichia coli*]	100	100	26%	1.00E-19	54%	WP_040234776.1

2.4 Discussion

Receptor binding proteins are the components used by phages for host recognition and attachment. Previously, we along with other groups have shown that there are several applications for RBPs, both as therapeutics and as diagnostics (Waseh *et al.*, 2010; Singh *et al.*,

2012). However, in order to advance the usage of these technologies, a method for the timely identification of RBPs is necessary. This research describes a method for the discovery of RBPs from phages that does not depend on sequencing or bioinformatic comparison. Briefly, phage DNA is extracted, fragmented, ligated into an expression vector, transformed into *E. coli* and then transferred to a nitrocellulose membrane where recombinant proteins are overexpressed. The membrane is then probed for binding to the phage host bacterium.

This assay allows for the discovery of RBPs based on their ability to bind to a host bacterium, and thus represents an improvement over previous methods for identifying RBPs based on homology to already known RBPs or on genome synteny, since the latter methods prohibit the discovery of truly novel and unrelated RBPs. Phenotypic screens have been used in the past to discover phage proteins, including discovery of phage lysins from both isolated phages and from metagenomic samples (Schmitz *et al.*, 2008). It may eventually be possible to utilize DNA from an environmental sample to search for RBPs that originate from phages that cannot be propagated under laboratory conditions, which would help elucidate the characteristics of these phages. However, we acknowledge that the current form of this assay is limited by the fact that the target bacteria must be culturable on agar plates and if a chaperone is necessary for RBP folding, that chaperone must be located in close proximity to the RBP in the phage genome. To circumvent the need to culture the target organism, future studies could use this method in combination with other detection methods, such as antibodies, to distinguish host bacteria bound to the nitrocellulose membranes. Interestingly, we noticed that all of the inserts discovered in the assay had the same orientation in the vector, suggesting that they were utilizing the vector promoter. However, protein expression most likely involved the native ribosome binding sites of

the individual genes, as opposed to that of the vector, since the inserts were found to be translated in different frames.

To date, we have not utilized this assay for species beyond the *Enterobactericae*, but due to the wealth of proteins from other species that have been expressed in *E. coli*, which includes RBPs (Arutyunov *et al.*, 2014), it is likely that this assay will be amenable to RBP discovery from phages targeting other species. For example, *E. coli* BL21(DE3) has been used in the past for over-expression of Gram-positive phage tailspikes (Smith *et al.*, 2005). It is possible that a *Bacillus subtilis* expression system might facilitate the identification of RBPs from Gram-positive phages, but this has yet to be assessed. An *in vitro* expression system is also an alternative technology that may produce superior results with some phages, such as those whose gene products are not readily expressed in *E. coli.*

This assay may also have utility beyond identifying phage RBPs. For instance, both GTAs and pyocins have RBPs, which could be detected by this assay if a library of GTA or pyocin DNA were generated. This would provide both a novel source of bacterial binding proteins and a means of better characterizing these molecular machines. Relatedly, RBPs identified in this assay could be used to alter the target of pyocins or whole phages.

RBPs are some of the most biotechnologically relevant proteins encoded in phage genomes, as they have evolved to specifically and effectively bind bacterial hosts. As well, RBPs represent the most dynamic proteins encoded by most phages, as constant evolutionary arms race dynamics between phages and bacteria over time lead to a strong selection for new variants (Ferris *et al.*, 2007). RBPs encompass a considerable diversity of binding targets, making them attractive in commercial exploitation and important for understanding the lifecycle of any phage.

Over the last decade, phage research has exhibited a resurged interest, especially in the area of phage therapy. However, the fact still remains that phages, and particularly the bacterial receptors they recognize, need to be well characterized and carefully chosen in order to ensure the success of phage therapy strategies. Therefore, rapid RBP identification and careful characterization of RBP-receptor interactions is a logical and promising avenue to develop in order to help move this field forward.

Acknowledgments: I we would like to thank Harald Nothaft for his help with Figure 2.4. JCS is a recipient of an Alexander Graham Bell Canada Graduate Scholarship. CMS is an Alberta Innovates Technology Futures iCORE Strategic Chair in Bacterial Glycomics.

2.5 References:

Ahmed, R., Bopp, C., Borczyk, A., and Kasatiya, S. (1987) Phage-typing scheme for *Escherichia coli* O157:H7. *J Infect Dis* **155**: 806–809

Arutyunov, D., Singh, U., El-Hawiet, A., Seckler, H. dos S., Nikjah, S., Joe, M., *et al.* (2014) Mycobacteriophage cell binding proteins for the capture of mycobacteria. *Bacteriophage* **4**: e960346 http://dx.doi.org/10.4161/21597073.2014.960346.

Barbirz, S., Müller, J.J., Uetrecht, C., Clark, A.J., Heinemann, U., Seckler, R., *et al.* (2008) Crystal structure of *Escherichia coli* phage HK620 tailspike: podoviral tailspike endoglycosidase modules are evolutionarily related. *Mol Microbiol* **69**: 303–316.

Bartual, S.G., Otero, J.M., Garcia-Doval, C., Llamas-Saiz, A.L., Kahn, R., Fox, G.C., and Raaij, M.J. van (2010) Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proc Natl Acad Sci U S A* **107**: 20287–92

Casjens, S.R., and Molineux, I.J. (2012) Short Noncontractile Tail Machines: Adsorption and DNA Delivery by Podoviruses. In *Viral Molecular Machines*. Springer US, pp. 143–179 http://link.springer.com/10.1007/978-1-4614-0980-9_7.

Cowley, L.A., Beckett, S.J., Chase-Topping, M., Perry, N., Dallman, T.J., Gally, D.L., and Jenkins, C. (2015) Analysis of whole genome sequencing for the *Escherichia coli* O157:H7 typing phages. *BMC Genomics* **16**: 271

Ferris, M.T., Joyce, P., and Burch, C.L. (2007) High frequency of mutations that expand the host range of an RNA virus. *Genetics* **176**: 1013–1022.

Freiberg, A., Morona, R., Bosch, L. Van den, Jung, C., Behlke, J., Carlin, N., *et al.* (2003) The tailspike protein of Shigella phage Sf6. A structural homolog of Salmonella phage P22 tailspike protein without sequence similarity in the beta-helix domain. *J Biol Chem* **278**: 1542–1548.

Ge, P., Scholl, D., Leiman, P.G., Yu, X., Miller, J.F., and Zhou, Z.H. (2015) Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states. *Nat Struct Mol Biol* **22**: 377–382

Gebhart, D., Lok, S., Clare, S., Tomas, M., Stares, M., Scholl, D., *et al.* (2015) A modified R-type bacteriocin specifically targeting *Clostridium difficile* prevents colonization of mice without affecting gut microbiota diversity. *MBio* **6**: 10.1128/mBio.02368-14.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845–858

Kropinski, A.M., Arutyunov, D., Foss, M., Cunningham, A., Ding, W., Singh, A., *et al.* (2011) Genome and proteome of *Campylobacter jejuni* bacteriophage NCTC 12673. *Appl Environ Microbiol* **77**: 8265–8271.

Lang, A.S., Zhaxybayeva, O., and Beatty, J.T. (2012) Gene transfer agents: phage-like elements of genetic exchange. *Nat Rev* **10**: 472–482

Michel-Briand, Y., and Baysse, C. (2002) The pyocins of *Pseudomonas aeruginosa*. *Biochimie* **84**: 499–510.

Miller, W.G., Bates, A.H., Horn, S.T., Brandl, M.T., Wachtel, M.R., and Mandrell, R.E. (2000) Detection on surfaces and in Caco-2 cells of *Campylobacter jejuni* cells transformed with new gfp, yfp, and cfp marker plasmids. *Appl Environ Microbiol* **66**: 5426–5436.

Novotny, M. V, Alley, W.R., and Mann, B.F. (2013) Analytical glycobiology at high sensitivity: current approaches and directions. *Glycoconj J* **30**: 89–117

Poshtiban, S., Javed, M.A., Arutyunov, D., Singh, A., Banting, G., Szymanski, C.M., and Evoy, S. (2013) Phage receptor binding protein-based magnetic enrichment method as an aid for real time PCR detection of foodborne bacteria. *Analyst* **138**: 5619–5626

Ritchie, J.M., Greenwich, J.L., Davis, B.M., Bronson, R.T., Gebhart, D., Williams, S.R., *et al.* (2011) An *Escherichia coli* O157-Specific Engineered Pyocin Prevents and Ameliorates Infection by *E. coli* O157:H7 in an Animal Model of Diarrheal Disease. *Antimicrob Agents Chemother* **55**: 5469–5474

Schmitz, J.E., Daniel, A., Collin, M., Schuch, R., and Fischetti, V.A. (2008) Rapid DNA library construction for functional genomic and metagenomic screening. *Appl Environ Microbiol* **74**: 1649–1652.

Seul, A., Müller, J.J., Andres, D., Stettner, E., Heinemann, U., and Seckler, R. (2014) Bacteriophage P22 tailspike: structure of the complete protein and function of the interdomain linker. *Acta Crystallogr Sect D Biol Crystallogr* **70**: 1336–1345

Simpson, D.J., Sacher, J.C., and Szymanski, C.M. (2015) Exploring the interactions between bacteriophage-encoded glycan binding proteins and carbohydrates. *Curr Opin Struct Biol* **34**: 69–77

Singh, A., Arutyunov, D., Szymanski, C.M., and Evoy, S. (2012) Bacteriophage based probes for pathogen detection. *Analyst* **137**: 3405–3421.

Singh, A., Glass, N., Tolba, M., Brovko, L., Griffiths, M., and Evoy, S. (2009) Immobilization of bacteriophages on gold surfaces for the specific capture of pathogens. *Biosens Bioelectron* **24**: 3645–3651.

Smith, N.L., Taylor, E.J., Lindsay, A.-M., Charnock, S.J., Turkenburg, J.P., Dodson, E.J., *et al.* (2005) Structure of a group A streptococcal phage-encoded virulence factor reveals a catalytically active triple-stranded -helix. *Proc Natl Acad Sci* **102**: 17652–17657

Tay, L.-L.L., Huang, P.-J.J., Tanha, J., Ryan, S., Wu, X., Hulse, J., and Chau, L.-K.K. (2012) Silica encapsulated SERS nanoprobe conjugated to the bacteriophage tailspike protein for targeted detection of Salmonella. *Chem Commun (Camb)* **48**: 1024–6

Warren, R.A. (1980) Modified bases in bacteriophage DNAs. *Annu Rev Microbiol* **34**: 137–158.

Waseh, S., Hanifi-Moghaddam, P., Coleman, R., Masotti, M., Ryan, S., Foss, M., *et al.* (2010) Orally Administered P22 Phage Tailspike Protein Reduces Salmonella Colonization in Chickens: Prospects of a Novel Therapy against Bacterial Infections. *PLoS One* **5**: e13904.

Chapter 3

A plant-produced bacteriophage tailspike protein for the control of Salmonella

A version of this chapter was published:

Miletic, S., Simpson, D.J., Szymanski, C.M., Deyholos, M.K., and Menassa, R. (2016) A Plant-Produced Bacteriophage Tailspike Protein for the Control of Salmonella. *Front Plant Sci* **6**: 1221.

Author Contributions: Sean Miletic and I, designed and performed the experiments and analyzed the data. I was responsible for all the bacterial work and the design of the animal experiments which were performed with the help of Cory Wenzel. Sean Miletic wrote the paper. Rima Menassa, Christine Szymanski and Michael Deyholos supervised this project and contributed the materials and reagents. All the authors contributed to the writing of the paper during the editing process.

3.1 Introduction

The human intestine is a complex microbial ecosystem where hundreds of species of bacteria have adapted to live and grow. A mutualistic relationship has evolved benefiting the health of the host while providing an optimal habitat for microflora to thrive (Guarner and Malagelada, 2003) However, some of these bacteria are pathogenic, causing a wide array of intestinal pathologies. *Salmonella enterica* is a Gram-negative enteropathogenic bacterium that is widely prevalent and is one of the primary causes of foodborne illness in humans. There are roughly 1.4 million nontyphoidal salmonellosis cases each year in North America, causing approximately 25% of all hospitalizations due to foodborne illness (Mead, 1999). *S. enterica* serotype Typhimurium also referred to as *S.* Typhimurium, causes gastroenteritis characterized by diarrhea, vomiting and abdominal pain and is showing the emergence of multidrug resistant strains (Su *et al.*, 2004; Chen *et al.*, 2013). Poultry and eggs are a major source of infection, but other sources such as vegetables, fruits, nuts, sprouts, leafy greens, roots, and beans have been reported (Rodrigue *et al.*, 1990; Hammack, 2011). In chickens, *Salmonella* is found throughout the digestive tract (Fanelli *et al.*, 1971) and the rupturing of intestinal contents during evisceration can readily contaminate poultry meat. For instance, *Salmonella* has been isolated from 33% of raw chicken breasts sampled from retail grocery stores in Ontario, Canada (Cook *et al.*, 2012).

Antibiotic use has led to the emergence of antibiotic-resistant *Salmonella* strains. In 2013, 17% of typhoidal *Salmonella* isolates from Canadians were resistant to ciprofloxacin and 41% of *Salmonella* Heidelberg infections were resistant to at least one antibiotic [1]. This growing concern has provoked research into alternative methods for controlling bacterial outbreaks. Considerable

[1] http://healthycanadians.gc.ca/alt/pdf/publications/drugs-products-medicamentsproduits/antibiotic-resistance-antibiotique/antimicrobial-surveillance-antimicrobioresistanceeng.pdf

research into using bacteriophage therapy to treat or prevent bacterial infections progressed in Eastern Europe and the former Soviet Union during the latter part of the 20th century and could potentially be reconsidered as a viable alternative to antibiotics (Sulakvelidze *et al.*, 2001). Lytic bacteriophages are host-specific, self-replicating, and virtually nontoxic making them attractive alternatives to control bacteria such as *Salmonella* and bacteriophages have been shown to reduce *Salmonella* colonization in chickens (Goode *et al.*, 2003; Atterbury *et al.*, 2007). Despite these successes, this therapy is not without drawbacks. Bacteriophages are host-specific requiring diagnosis of the pathogen before the phage is administered (Waseh *et al.*, 2010). Phages can also carry harmful genes and can potentially transfer these genes to the bacteria, increasing virulence (Skurnik and Strauch, 2006). As a result, there has been interest in the use of phage proteins such as endolysins (Roach and Donovan, 2015) as tools for the specific targeting of bacteria and the exploitation of phage receptor binding proteins for use in diagnostics and engineered phage-derived killing machines (Singh *et al.*, 2012; Simpson *et al.*, 2015). Unexpectedly, Waseh et al. have demonstrated that the P22 phage tailspike protein alone is effective in controlling *Salmonella* colonization and spread in chickens, presumably through its binding capability (Waseh *et al.*, 2010). These tailspike proteins are highly stable homotrimers that form the short tail of the bacteriophage and bind to the O-antigenic repeating units on the outer membrane lipopolysaccharide (Baxa *et al.*, 1996). The tailspike protein Gp9 from the P22 bacteriophage can recognize several serovars of *Salmonella* including *S.* Typhimurium, *S*. Paratyphi A, and *S*. Enteritidis. A shortened version of Gp9 has been shown to agglutinate *S.* Typhimurium, inhibit bacterial motility and reduce colonization in the chicken gut (Waseh *et al.*, 2010). Therefore, this protein has the potential to act as an effective pre-slaughter feed additive to reduce *Salmonella* contamination in chickens.

Plant bioreactors have been growing in acceptance as feasible production platforms for therapeutic proteins, as they are highly scalable and can be established with little upfront cost (Fischer *et al.*, 2012). Protein drugs expressed in plant tissue are thought to be protected from digestive enzymes by the plant cell wall (Kwon and Daniell, 2015), and are especially useful for veterinary applications where regulations allow administration of unpurified or partially purified extracts (MacDonald *et al.*, 2015). For example, leaf tissue can be harvested, lyophilized, and orally administered in capsules or suspended in a slurry removing costs associated with protein purification, administration, and cold-storage (Kolotilin *et al.*, 2014). As higher eukaryotic organisms, plants can introduce post-translational modifications required for complex recombinant proteins. Despite these benefits, protein yield remains a major factor limiting the widespread adoption of plant bioreactors for commercial protein production. Consequently, several approaches are currently being used to increase protein accumulation in plants. Proteins can be targeted to different subcellular compartments such as the endoplasmic reticulum (ER), the chloroplasts, and the apoplast using signal and transit peptides (Conley *et al.*, 2009). This is because each subcellular compartment has a unique biochemical environment, protease content and physical size which influence protein accumulation levels (Streatfield, 2007; Pillay *et al.*, 2014). Additionally, peptide tags can be fused to recombinant protein to increase accumulation. For example, fusion tags such as elastin-like polypeptides (ELPs) and hydrophobin I (HFBI) can increase recombinant protein accumulation levels, and have also been used to purify proteins from plant extracts (Conley *et al.*, 2011).

The goal of this project was to transiently produce the truncated version of Gp9 in *N. benthamiana* by targeting the protein to the chloroplasts, to the ER, or to the ER fused with an ELP or HFBI tag. The activity of plant-produced Gp9 was then tested by examining its ability to

bind to *S.* Typhimurium. Lastly, plant tissue containing Gp9 was orally administered to chicks inoculated with *S.* Typhimurium to determine if this plant produced therapeutic has the potential to limit *Salmonella* colonization.

3.2 Materials and methods

3.2.1 Gene cloning

The truncated version (encoding amino acids 109-666) of the endorhamnosidase mutant of *gp9* (as described in Waseh et al., 2010) was codon optimized for plant expression and synthesized by Biobasic Inc. (Markham, Ontario). *Gp9* was recombined into the previously constructed pCaMGate expression vectors using the LR reaction of Gateway® technology (Invitrogen, Thermo Fischer Scientific, Waltham, U.S.A.), courtesy of Dr. Andrew Conley of Agriculture and Agri-Food Canada, London Ontario. Recombinant pCAMGate vectors were transformed in *E. coli* XL1-Blue using the Gene Pulser II system (Bio-Rad Laboratories Inc., Hercules, U.S.A,) and PCR screening using gene-specific primers (Forward primer: CGTTAGGTGTAGGTTTTGGTATGGATGGT, Reverse primer: CCGGCAACAGGATTCAATCTTAA) was conducted to screen for positive transformants containing the correct insert. Plasmid DNA was isolated from positive colonies and transformed into electro-competent *A. tumefaciens* EHA105 cells. Electroporated *A. tumefaciens* cells were spread on yeast extract broth (YEB) plates containing 50 µg/ml kanamycin and 10 µg/ml rifampicin and incubated for two days at 28°C.

3.2.2 Transient expression in *N. benthamiana* plants

Suspensions of *A. tumefaciens* carrying *Gp9* or *A. tumefaciens* carrying the post-transcriptional gene silencing suppressor *p19* from Cymbidium ringspot virus (Silhavy *et al.*, 2002), were incubated overnight at 28°C with shaking at 250 rpm until an optical density at 600 nm (OD600)

of 0.5- 1.0 was reached. Cultures were then centrifuged at 6000 x g for 30 min and resuspended to an OD600 of 1.0 in Gamborg's solution containing 3.2 g/l Gamborg's B5 with vitamins, 20 g/l sucrose, 10 mM MES (pH 5.6) and 200 µM acetosyringone. Cultures were then incubated at room temperature with gentle agitation for one hour. An equal volume of *A. tumefaciens* culture containing *Gp9* was combined with *A. tumefaciens* culture carrying *p19* and Gamborg's solution to give a total *A. tumefaciens* OD600 of approximately 0.67. These suspensions were used to infiltrate 7-8 week old *N. benthamiana* plants grown in a growth room under 16 hour light/8 hour dark conditions at 21-22°C with 55% humidity, and receiving roughly 100 µmol/photons $m^{-2}s^{-1}$ of light. A 3 ml syringe was used to infiltrate the *A. tumefaciens* suspensions through the stomata of the abaxial leaf epidermis of *N. benthamiana*. After infiltration, plants were returned to the growth chamber for up to six days.

3.2.3 Tissue collection and protein extraction

Four biological replicates were used in all experiments and consisted of four plants sampled as follows: two leaf disks/leaf (7 mm diameter) were collected from three infiltrated leaves of each plant and pooled. Tissue was flash frozen in liquid nitrogen and stored at -80°C until use. For protein extraction, tissue was homogenized twice in 30 second pulses using a TissueLyser (Qiagen, Venlo, Netherlands) and total soluble proteins were extracted in 200 µl of plant extraction buffer (PEB) containing 1X phosphate-buffered saline (PBS), 0.1% (v/v) Tween-20, 2% (w/v) polyvinylpolypyrrolidone (PVPP), 100 mM ascorbic acid, 1 mM ethylenediaminetetraccetic acid (EDTA), 1mM of phenylmethanesulfonylfluoride (PMSF) and 1 µg/ml leupeptin. Total soluble protein (TSP) concentration for each sample was determined using the Bradford assay (Bradford, 1976). For electrophoresis under non-reducing conditions, protein samples were stored in a 5% SDS, 250 mg/ml glycerol, 0.1 mg/ml bromophenol blue,

0.16 M Tris/HCl sample buffer (Seckler *et al.*, 1989) to better visualize protein trimerization. Samples were frozen at -80°C until use.

3.2.4 Western blotting and gel staining

Pooled sample extracts and individual replicates were immunodetected against a standard curve of known amounts of cellulose binding domain (CBD) synthetic protein standard (synthesized by Genscript, Piscataway, USA) or purified Gp9-ELP to accurately quantify protein accumulation levels. Samples were either boiled for ten minutes or not boiled and loaded onto a Bio-Rad MiniProtean® TGX™ Precast 4-20% (w/v) polyacrylamide gradient gel. Separated proteins were transferred to polyvinylidene difluoride (PVDF) membranes and blocked overnight in a 5% (w/v) skim milk powder in TBS-T (Tris-buffered saline-Tween 20) blocking solution.

Membranes were incubated with a 1:5000 dilution of mouse anti-c-Myc antibody (Genscript, A00864, Piscataway, USA) or a 1:8000 dilution of rabbit polyclonal anti-Gp9 antibody (Kropinski *et al.*, 2011). Membranes were washed and incubated with a 1:5000 dilution of goat anti-mouse or anti-rabbit secondary antibody conjugated with horseradish peroxidase (HRP), and visualized with the GE Healthcare Life Sciences (Little Chalfont, UK) ECL Prime Western Blotting Detection Reagent. Recombinant protein was quantified by image densitometry using Totallab TL100 software (Nonlinear Dynamics, Durham, USA). For staining of separated proteins, gels were washed for 5 minutes in water and stained for 1 hour with GelCode™ Blue stain reagent (Thermo Fischer Scientific, Waltham, U.S.A.) at room temperature. Gels were then destained using three 5-minute washes with water and imaged.

3.2.5 Protein purification

Gp9-ELP and Gp9-HFBI proteins were purified using a c-Myc tag purification kit from MBL International Corporation MBL (3305, Woburn, USA) according to the manufacturer's

instructions. Purified protein was stored at -80°C until use. The *E. coli* produced His_6-Gp9 was purified as described by Waseh et al., 2010.

3.2.6 Gp9 and Gp9-ELP adherence to *S.* Typhimurium

S. enterica serovar Typhimurium (ATCC19585) was purchased from the American Type Culture Collection (Manassas, VA) and grown under aerobic conditions at 37°C on Lysogeny Broth (LB) agar plates. The strain used in the adherence assay was transformed with the pWM1007 plasmid (Miller *et al.*, 2000) which expresses the green fluorescence protein (GFP) and grown on LB supplemented with 25 µg/ml kanamycin.

Five hundred nanograms of the *E. coli* produced His_6-Gp9, BSA or plant-produced Gp9-ELP were spotted onto hole punch sized pieces of nitrocellulose membranes and then blocked for 1 hour in 5% skim milk in PBS with 0.05% Tween (PBS-T). The membranes were then probed with 10^8 cfu/ml of GFP-expressing *S.* Typhimurium in 5% skim milk PBS-T. The disks were washed 3 times for 5 minutes in PBS-T and were then placed onto LB agar plates with 25 µg/ml kanamycin and allowed to grow at room temperature overnight, followed by growth at 37°C for 8 hours.The disks were imaged with a FujiFilm FLA-5000 system using the 473nm laser at 400V for excitation and LPB (Y510) filter for emission. Fluorescence intensity was measured using the MultiGauge version 3.0 software.

3.2.7 Animal studies

Animal studies were carried out in accordance with the protocol approved by the Animal Care and Use Committee at the University of Alberta following the procedure described by Waseh et al., (2010). Each group contained 5-8 SPF leghorn birds (Poultry Research Facility, University of Alberta) that were provided with feed and water *ad libitum* and were randomly tested for the presence of *Salmonella* on the day of hatch by plating cloacal swabs onto selective Oxoid

Brilliance *Salmonella* agar (Oxoid, Ontario, Canada). In all cases no *Salmonella* colonies were observed after 24 h of incubation at 37°C. Chickens were orally gavaged with 300 µL PBS containing 10^7 colony forming units (CFU) of *S.* Typhimurium the next day and then gavaged with 35 mg lyophilized and powdered leaves resuspended in 300µl of PBS at 1 h, 18 h and 42 h post-infection. The chickens were culled at 47 h post-infection and the collected cecal contents were serially diluted and plated onto Oxoid Brilliance *Salmonella* agar. *Salmonella* CFU were counted after the agar plates were incubated for 24 h at 37°C.

3.2.8 Statistics

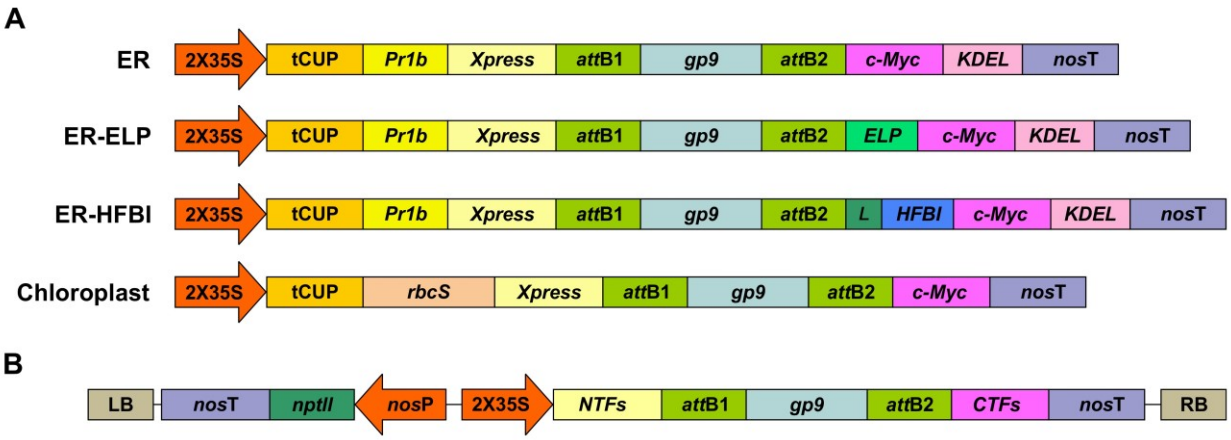
Minitab® 17 statistical software (Minitab Ltd., Coventry, UK) was used to perform statistical analysis on the Gp9 accumulation data. A one-way analysis of variance (ANOVA) was performed with a Tukey test on the mean Gp9 accumulation levels for each day of the time course. $P < 0.05$ was considered significant. A two-tailed student T-test was performed on the data from the chicken experiment. $P < 0.05$ was once again considered significant.

3.3 Results

3.3.1 Gp9 transient expression in *N. benthamiana*

The *Gp9* tailspike gene was cloned into pCaMGate expression vectors (Pereira et al., 2014) targeting the ER, the ER fused with an ELP tag (ER-ELP), the ER fused with a HFBI tag (ERHFBI), or the chloroplasts (Figure 3.1). Constructs were then agroinfiltrated along with the gene silencing suppressor, p19 (Silhavy et al., 2002), into the leaves of *N. benthamiana* plants. Plants were monitored over the course of 6 day(s) post-infiltration (dpi). Young, upper leaves infiltrated with the ER-targeted constructs turned a slight yellow green color but otherwise the phenotypes remained unchanged. Leaf tissue samples were collected from 3 to 6 dpi and

analyzed for Gp9 accumulation via Western blot using a c-Myc antibody specific for the C-terminal Myc peptide found in all constructs (Figure 3.2A). Gp9 targeted to the ER was faintly visible, running higher than predicted (Figure 3.1C), and no bands were visible for the chloroplast-targeted Gp9. Conversely, Gp9-ELP and Gp9-HFBI were readily detected, and faint bands were observed for both the ELP and HFBI fused constructs migrating above 150 kDa which could represent potential trimers. Smaller bands ranging in size between 25-50 kDa were also observed in the Gp9-ELP and Gp9HFBI samples which most likely correspond to Gp9 degradation products since they are absent from the p19 negative control lane (Figure 3.2A). Interestingly, when immunoblots were probed with an anti-Gp9- antibody, a strong band slightly under 75 kDa was observed for ER-targeted Gp9, and a somewhat fainter band was seen in the chloroplast-targeted Gp9 (Figure 3.2B). Bands potentially representing dimers and trimers were also observed for all four proteins, as well as smaller bands that may represent Gp9 degradation products. It is interesting that the smaller bands observed in the Gp9-ELP and Gp9-HFBI are of a different size when detected with the two antibodies. The protein band migrating between 37.5 and 50 kDa may be a plant protein as it appears faintly in the p19 lane, while the band running between 50 and 75 kDa may represent a degradation product of Gp9. This band is present in all samples, and it may represent the N-terminal portion of the protein detected with the Gp-9-specific antibody, while the smaller band running between 25-37 kDa on the blot detected with the c-Myc antibody might represent the C-terminal portion of the protein. The low abundance of Gp9 on blots probed with the c-Myc antibody also imply that the c-Myc tag is cleaved off Gp9, or is inaccessible to the antibody.



Construct	Predicted full length size	Predicted size without c-Myc tag	Predicted size without C-terminal fusions
ER	67	65.33	65.33
ER-ELP	78.47	76.8	65.33
ER-HFBI	74.78	73.63	65.33
Chloroplast	66.52	65.33	65.33

C

Figure 3.1: Expression cassettes for Gp9 transient expression in *N. benthamiana*. A) A schematic representing the gene constructs targeting the endoplasmic reticulum (ER) or the chloroplasts. Two of the cassettes targeting the ER also contain sequences coding for an elastinlike polypeptide (ELP) or a hydrophobin (HFBI) tag with a linker (L). Expression and targeting elements incorporated into the expression vectors are: the double-enhanced 35S

cauliflower mosaic virus promoter (2X35S); tCUP translational enhancer from tobacco; tobacco Pathogenesis-Related-1b signal peptide (PR1b) and the KDEL ER retrieval tetrapeptide; nopaline synthase terminator (*nos*T); and *Xpress* and *c-Myc* tags for detection and purification. A sequence coding for the RuBisCo small subunit transit peptide (*rbcS*) was used for targeting to the chloroplasts. B) A zoomed-out schematic of the constructs in the T-DNA. LB and RB are the left and right borders of the T-DNA sequence. *nptII* codes for neomycin phosphotransferase conferring resistance to kanamycin, driven by a nopaline synthase promoter (*nos*P) and terminator (*nos*T). Schematic represents the same constructs in A) with different N-terminal fusions (*NTFs*) and C-terminal fusions (*CTFs*). Schematics are not to scale. C) Predicted molecular weights of Gp9 constructs **(**values are in kilodaltons (kDa))

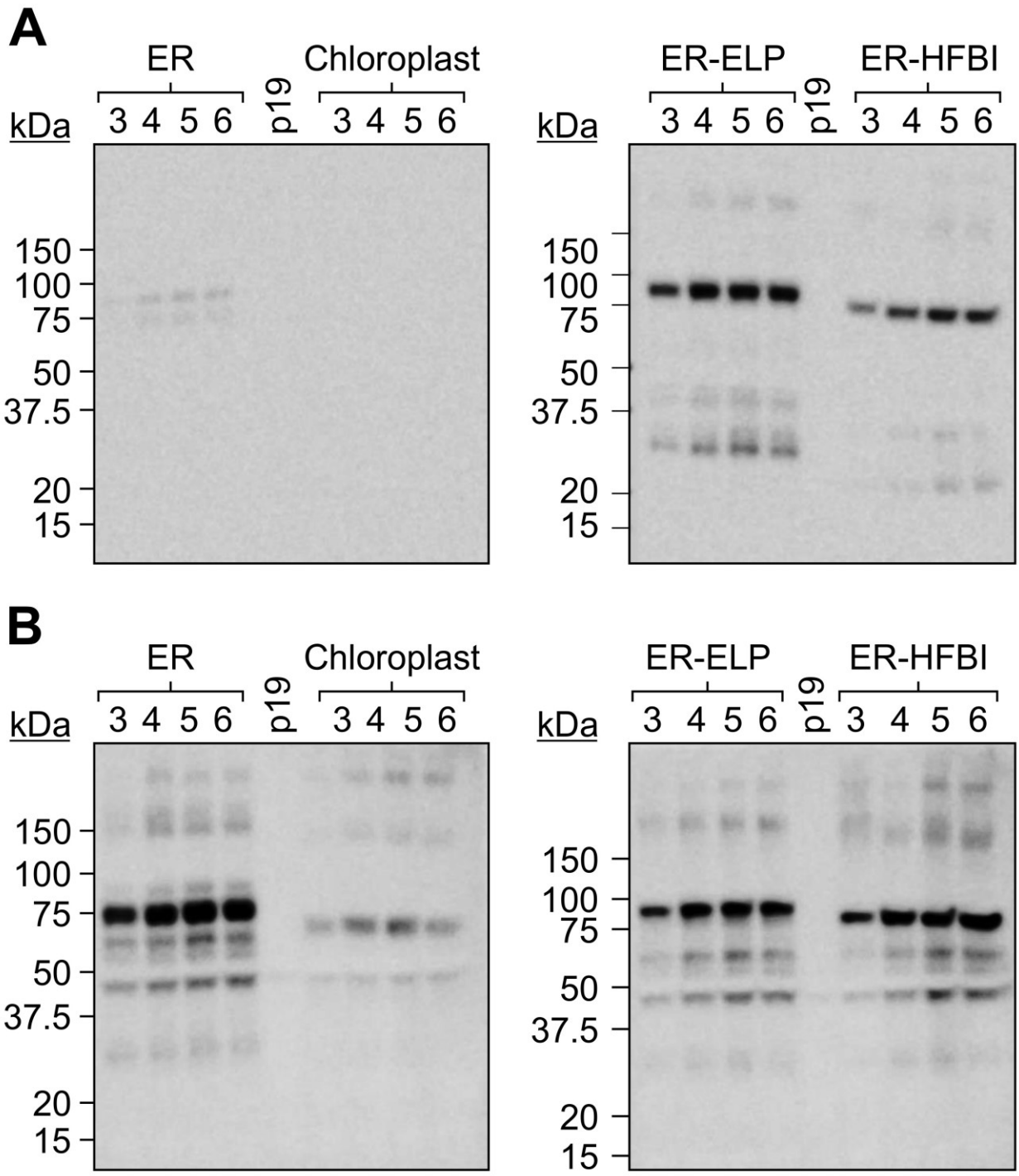


Figure 3.2: A time course of Gp9 accumulation levels in *Nicotiana benthamiana*. A)

Immunoblots of total soluble protein (TSP) extracted from *N. benthamiana* leaf tissue infiltrated with *gp9* constructs targeting the ER, the chloroplasts, or the ER fused to an ELP tag (ER-ELP)

or fused to a hydrophobin tag (ER-HFBI). Four plants were infiltrated with each construct. Tissue was sampled from infiltrated leaves from 3 to 6dpi. Equal volumes of TSP from each of the four replicates were pooled, and 20 µg was loaded per well. Immunoblots were probed with A) an anti-c-Myc antibody or B) an anti-Gp9 antibody. p19: TSP from plants infiltrated with p19 serving as a negative control.

3.3.2 Gp9 accumulation in *N. benthamiana*

Since more protein was detected using the anti-Gp9-antibody, this antibody was used to accurately quantify Gp9 accumulation via densitometry analysis. For the purpose of protein quantitation, only bands representing the full length Gp9 monomer were quantified.

Immunoblots using individual replicates revealed that Gp9 accumulation increases from the 3rd dpi, peaks on day 4 or 5, and subsequently decreases (Figure 3.3). ER-targeted Gp9 accumulated in significantly higher amounts than the other proteins reaching on average of 1.64 ± 0.09% of total soluble protein (TSP) on day 5. When recombinant protein accumulation is calculated in terms of micrograms of Gp9 per gram of fresh leaf weight (FLW), Gp9 accumulates to an average of 235.01 ± 21.12 µg/g of FLW on day 4 (Figure 3.3B). The presence of either the ELP or HFBI tag appears to significantly decrease Gp9 accumulation on 3-5 dpi ($P<0.05$). Gp9-ELP accumulates on average to 0.98 ± 0.05% of TSP on day 5 or 135.86 ± 25.56 µg/g of FLW on day 4, roughly 0.6 times or 1.7 times less, respectively, than when unfused. The presence of the HFBI tag caused Gp9 to accumulate in even lower amounts on average to 0.79 ± 0.09% of TSP (0.48 times less) or 116.49 ± 16.29 µg/g of FLW (2 times less) on day 4. Gp9 accumulated significantly less when targeted to the chloroplasts compared to ER-targeted proteins on days 4 and 5 ($P < 0.05$). Accumulation reached an average of 0.21 ± 0.01% of TSP or 28.45 ± 1.65 µg/g of fresh leaf weight on the fifth dpi.

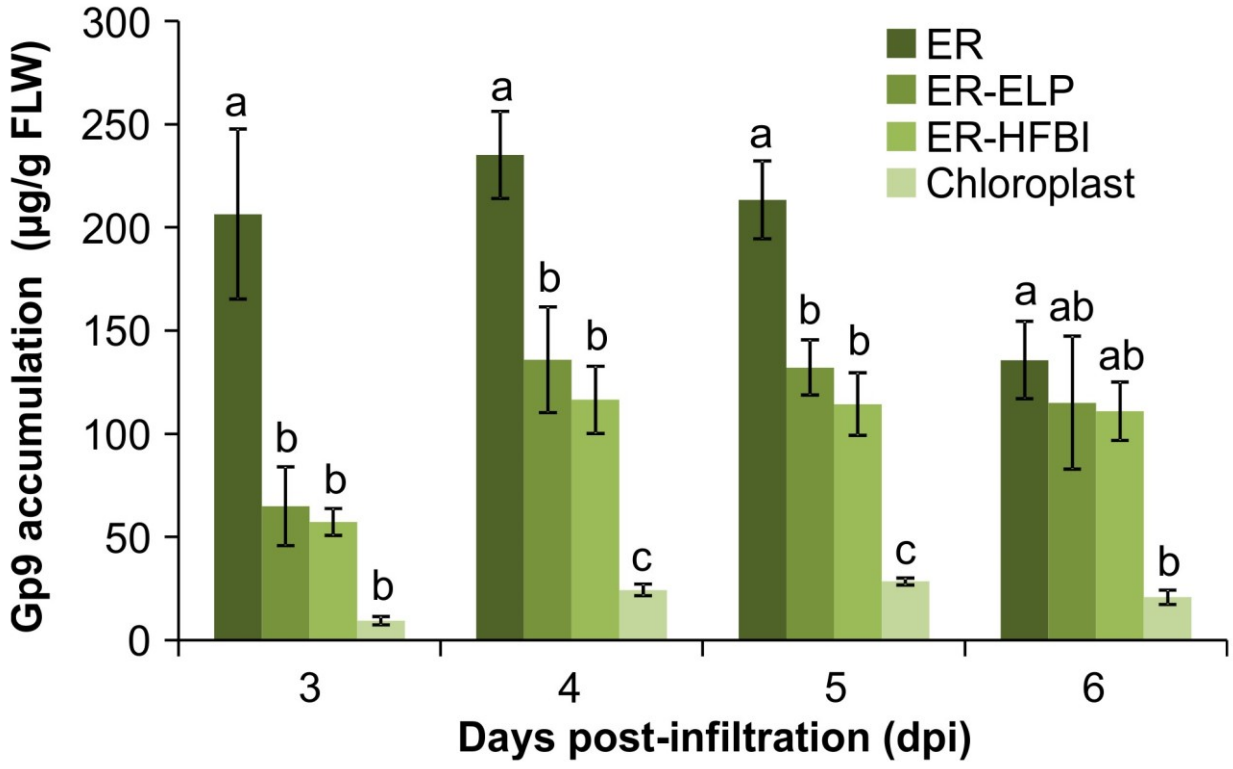


Figure 3.3: Quantification of Gp9 accumulation in *N. benthamiana* over 3, 4, 5, and 6dpi. Quantification was performed on total soluble protein extracts from infiltrated *N. benthamiana* tissue using a standard curve of known amounts of purified Gp9-ELP. Gp9 was targeted to the ER, the chloroplasts, or the ER fused to an ELP tag (ER-ELP) or fused to a hydrophobin tag (ER-HFBI). Immunoblots were probed with an anti-Gp9 antibody. Accumulation levels of Gp9 shown as A) a percentage of total soluble protein (TSP) or B) in µg per g of fresh leaf weight (FLW). Error bars represent the standard error of the mean value of four biological replicates. Treatments which do not share a letter are significantly different ($P < 0.05$) as determined by a one way ANOVA and the Tukey test.

3.3.3 Gp9 purification and characterization

Gp9 polypeptides are found in monomeric, dimeric, and trimeric intermediates before forming the stable, native trimer (Benton *et al.*, 2002). The trimeric intermediate species or protrimer

consists of associated subunits which have not completely folded forming a transient, less-stable precursor to the trimer (Goldenberg and King, 1982). We successfully purified Gp9-ELP using a c-Myc-tag purification kit (Figure 3.4A) and decided to investigate if purified Gp9-ELP is present in either of these states by avoiding complete denaturation (unheated sample) and by avoiding reducing agents such as dithiothreitol (DTT) in the sample buffer to keep the disulfide bonds oxidized (unheated, no DTT). While most of the Gp9-ELP was found in the monomeric form when it was reduced and denatured, there was very little monomer present when heat denaturation was omitted. Instead, an intense band was observed below 250 kDa, which corresponds to the expected size for the 215 kDa trimer (Figure 3.4A, B). Higher banding was observed when samples were electrophoresed under non-reducing conditions (Figure 3.4B). Gp9 contains eight cysteine residues and has disulfide bonds while existing as a protrimer, despite the fact that all are reduced during conversion to the native trimeric state (Robinson and King, 1997). Consequently, bands running above 250 kDa could represent the protrimer intermediate running slower than the trimer. Generally, an oxidizing environment is needed for trimer folding due to the presence of the disulfide bonds in the protrimer, yet increasing concentrations of DTT increases the conversion of protrimer to trimer (Robinson and King, 1997). Even larger bands were visualized and could represent higher-order multimers or aggregates (Altschul *et al.*, 1997). Our results suggest that Gp9 is present in the stable trimeric form when purified and is expected to remain functional to some degree.

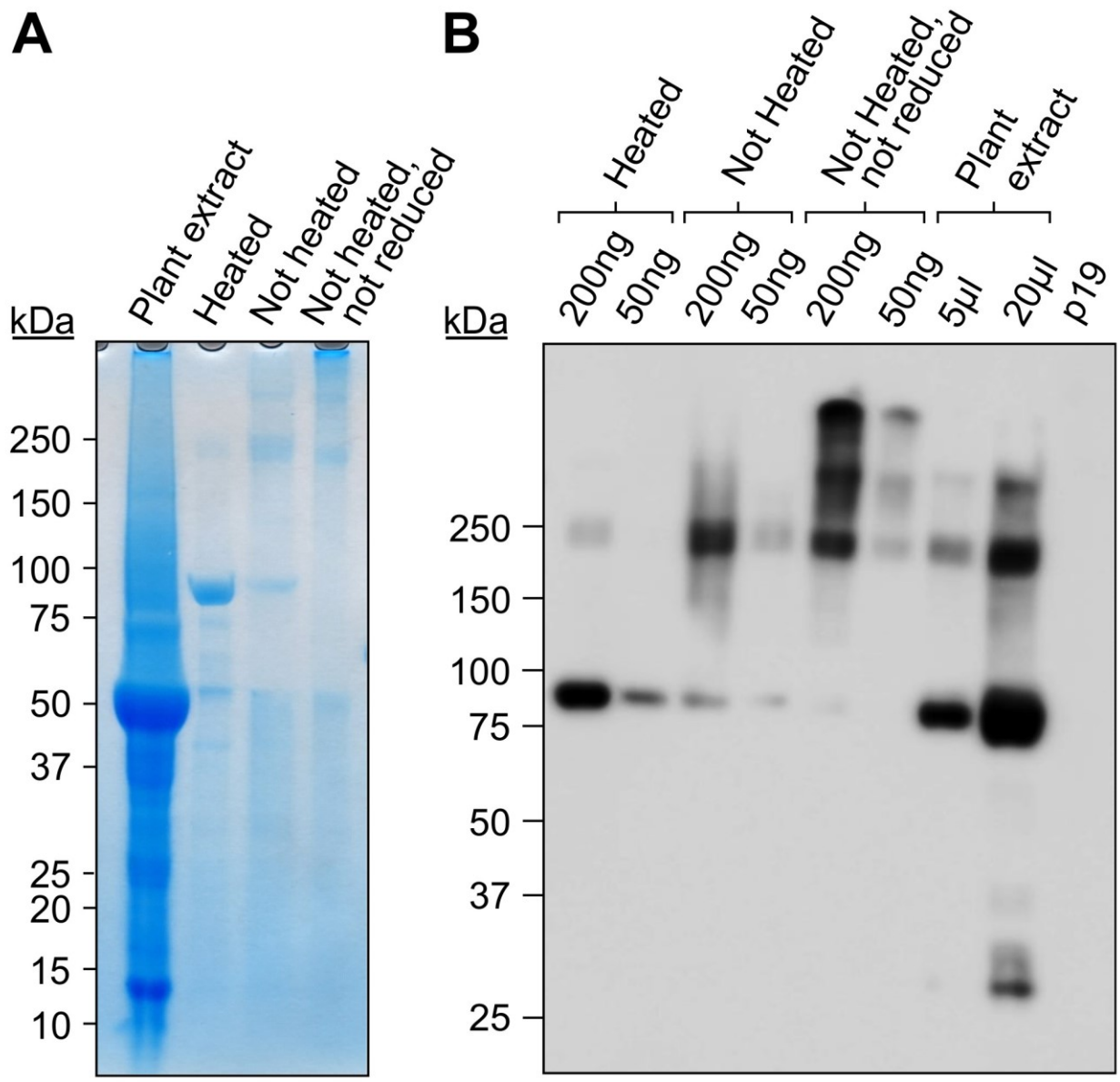


Figure 3.4: Purified Gp9-ELP is found in higher molecular weight protein complexes. Purified Gp9-ELP was electrophoresed under the following conditions: in sample buffer containing DTT and heat-denatured by boiling for 10 minutes, in sample buffer containing DTT with no heat denaturation, or in sample buffer without DTT and no heat denaturation. The plant extract containing Gp9-ELP was also loaded as a positive control in sample buffer containing DTT and heat-denatured by boiling for 10 minutes. Twenty microliters of extract from plants

587 infiltrated with p19 were used as a negative control. Gels were either coomassie stained (A) or 588 immunodetected (B) using an anti-c-Myc antibody.

3.3.4 Gp9-ELP binding to *S.* Typhimurium

To determine if the plant produced Gp9 can bind to *S.* Typhimurium, 500 ng of purified Gp9ELP was spotted onto a nitrocellulose membrane, blocked and then probed for 1 hour with 10^8 cfu/ml of GFP expressing *S.* Typhimurium. BSA and *E. coli*-produced Gp9 were used as negative and positive controls, respectively. Gp9-ELP was able to bind *S.* Typhimurium at similar levels as the *E.coli*-produced Gp9 as measured by fluorescence (Figure 3.5) suggesting that the plant-produced protein is functional.

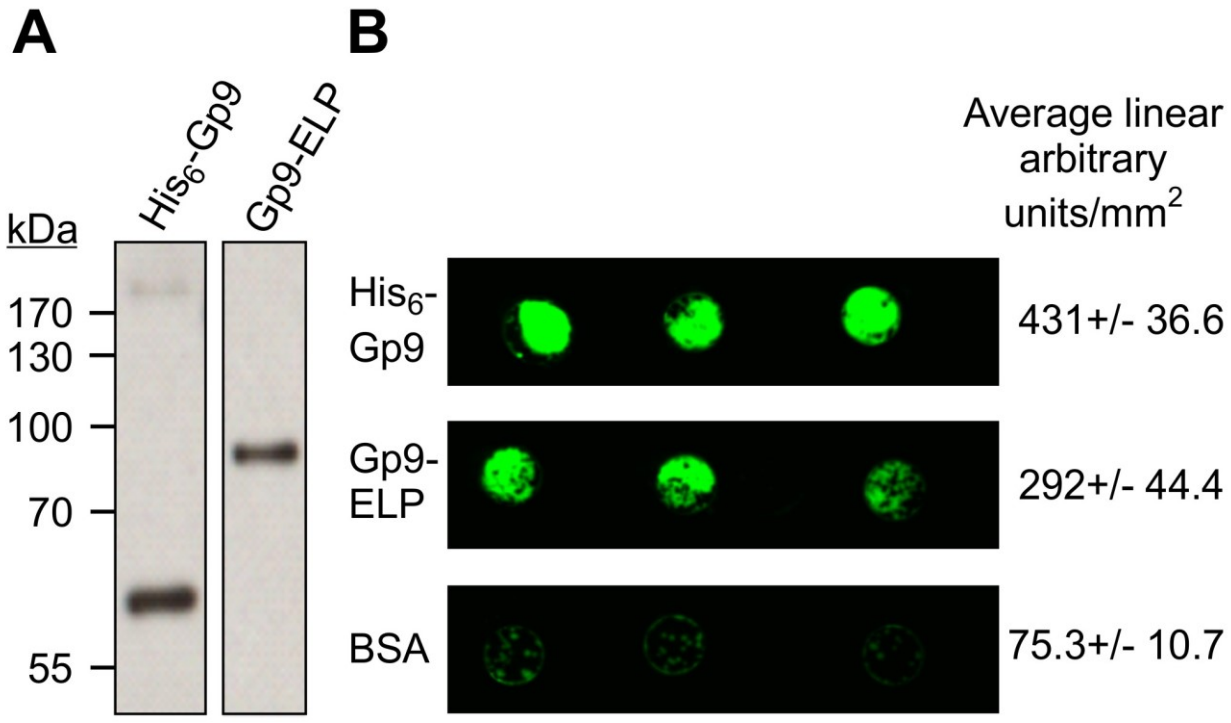


Figure 3.5: *S.* Typhimurium binding activity to plant produced Gp9-ELP A) The *E. coli* produced His_6-Gp9 (62 kDa) and the tobacco produced Gp9-ELP (82 kDa) both react with anti-Gp9 antibodies. B) Either 500 ng of His_6-Gp9, BSA or Gp9-ELP protein were spotted onto

nitrocellulose and then probed with 10^8 cfu/ml GFP-expressing *S.* Typhimurium and allowed to grow overnight. Two biological replicates were done, each with three technical replicates in each. One representative image and densitometry analysis is shown.

3.3.5 The effect of oral administration of Gp9-ELP on *Salmonella* colonization.

Due to the demonstrated binding activity of Gp9-ELP, we investigated the effects of orally administering plant tissue containing Gp9-ELP on *Salmonella* colonization in chicks. The chicks were orally gavaged with 35 mg of resuspended plant tissue containing 7.7 μg Gp9-ELP 1, 18 and 42 hours after gavaging with *S.* Typhimurium. Chicks were culled 47 h after bacterial infection and the *Salmonella* CFU in the cecal contents were enumerated (Figure 3.6). Gp9-ELP showed approximately 1-log reduction in *S.* Typhimurium colonization compared to the untreated control birds (P=0.058). These results are promising considering plant leaves were fed directly to the birds without further purification, and each dose contained much less Gp9-ELP compared with the Waseh et al. (2010) study where a dose consisted of 30 μg of purified Gp9.

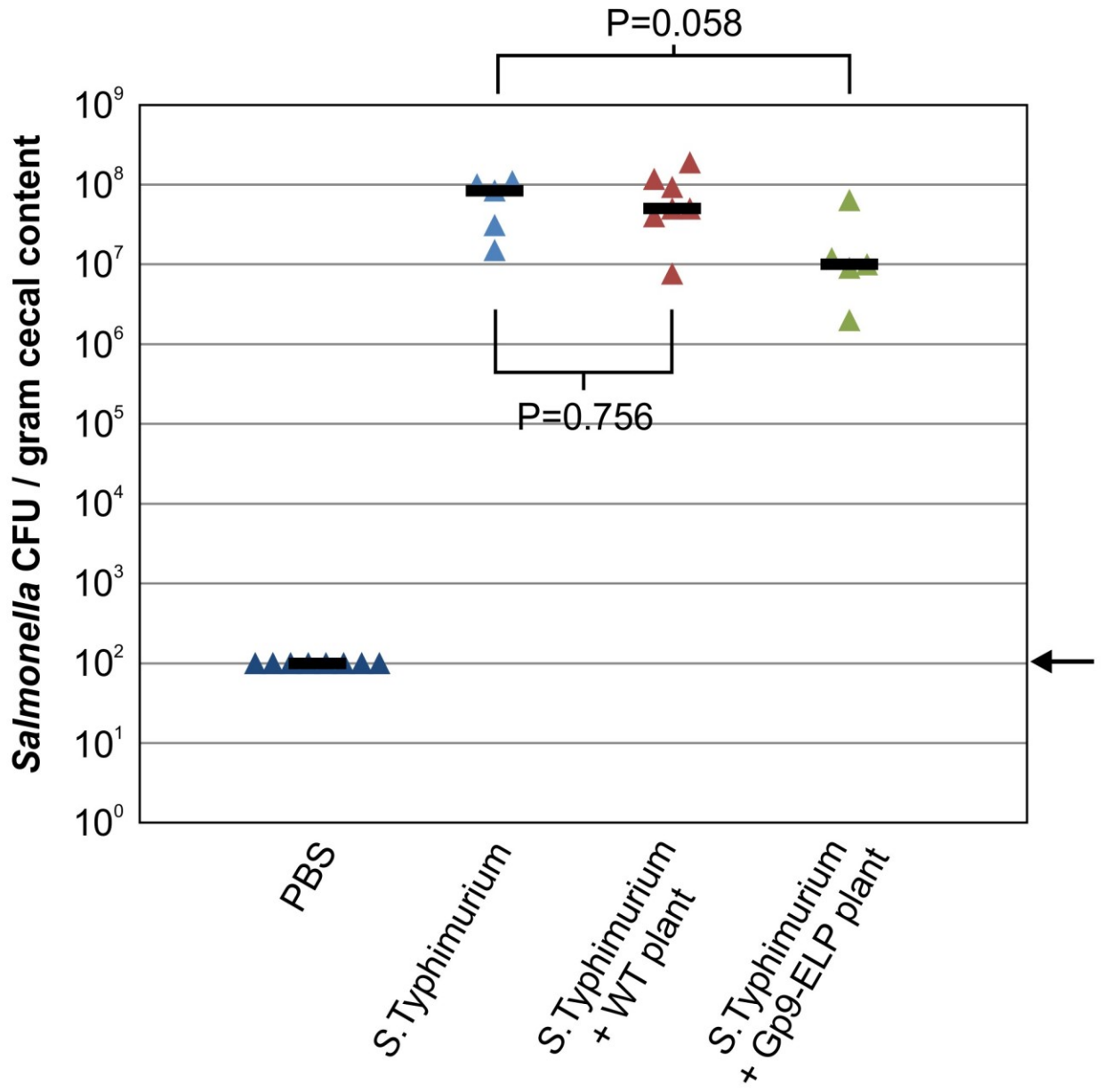


Figure 3.6: Treatment of *S.* Typhimurium infected chickens with Gp9-ELP. SPF leghorn chickens were orally gavaged with PBS or *S.* Typhimurium as indicated. One hour later (and at 18 and 42 h), two of the groups were orally gavaged with *N. benthamiana* extracts with or without Gp9-ELP, while the control groups received PBS. Chickens were fed 7.7 µg of Gp9-ELP per dose. *Salmonella* colonization levels of the chicken ceca were determined 47 h post infection. The median is shown for each group and p-values were determined for the plant-fed groups in

comparison to the *S.* Typhimurium group. The arrow represents the detection limit of the experiment.

3.4 Discussion

This study demonstrated that the receptor binding domain of the P22 bacteriophage tailspike protein, Gp9, could be transiently expressed in *N. benthamiana* at reasonable levels and is able to bind to *S.* Typhimurium. The results presented here suggest that the correct choice of antibody for protein detection is essential to accurately quantify recombinant protein accumulation. When probing with an antibody specific to a C-terminal Myc tag, Gp9 appeared to accumulate to the highest levels when fused to an ELP or HFBI tag reaching 33.2 µg/g or 20 µg/g of FLW, respectively, supporting the hypothesis of this study that these fusion tags would increase accumulation levels (Patel *et al.*, 2007; Conley *et al.*, 2009; Joensuu *et al.*, 2010). These results were expected as ELP and HFBI tags promote the formation and distribution of protein bodies (PBs) which are thought to protect recombinant protein from hydrolysis and protease cleavage (Conley *et al.*, 2011; Saberianfar *et al.*, 2015).

However, probing with a Gp9-specific antibody gave contrasting results. When the Gp9-specific antibody was used for immunodetection, unfused ER-targeted Gp9 displayed about 50% more accumulation than the ELP and HFBI fusions. This result was unexpected, and suggests that Gp9 is stable in *N. benthamiana* and the additional ELP or HFBI tags do not increase protein stability further. Indeed, ELP tags have negligible effects on already high accumulating proteins such as GFP (Conley *et al.*, 2009). It appears that the addition of an ELP or HFBI tag actually decreases Gp9 accumulation when targeted for retrieval to the ER of *N. benthamiana* and this should be investigated in future studies. These results also imply that the c-Myc tag is partially cleaved or

inaccessible causing us to underestimate Gp9 accumulation levels when detecting with an anti-c-Myc tag antibody. It is possible that the ELP and HFBI tags protect the c-Myc tag from cleavage, possibly through protein body formation, and therefore the fusion constructs were able to be detected with the c-Myc antibody. However, when probing with the Gp9 antibody the ER and chloroplast constructs produce a slightly higher banding pattern than expected if the cMyc is indeed cleaved off (Figure 3.2; Figure 3.1C). It can be postulated then that these proteins may fold into a conformation that makes the c-Myc inaccessible. Regardless, as many recombinant proteins are designed to have tags for easy detection, these results are a cautionary note on the limits of short tags, and that perhaps both N- and C-terminal tags might be used to improve chances of accurate quantitation, while protein-specific antibodies are most probably the best option if available. Our constructs also contain an N-terminal Xpress tag (Figure 3.1), however commercially available antibodies have not detected any recombinant protein in crude extracts. (unpublished).

Gp9 also accumulated in the chloroplasts, but to a lower level than when targeted to the secretory pathway. Generally, chloroplast targeting is a promising strategy for increasing recombinant protein accumulation (Hyunjong *et al.*, 2005). Nevertheless, some proteins do not accumulate well when targeted to these organelles. For example, one study described lower accumulation levels of zeolin, a chimeric storage protein, when targeted to the chloroplasts compared to the ER. A visualized degradation pattern of zeolin provided evidence for protein degradation in the chloroplasts (Bellucci *et al.*, 2007). On the other hand, Oey et al. were able to express a phage lysin protein in tobacco chloroplasts with accumulation levels reaching 70% of TSP (Oey *et al.*, 2009). This protein was produced by transforming the plastid genome and the gene was codon optimized for plastid expression. The phage lytic protein was shown to be extremely stable in the

plastids, which is understandable as phages have evolved substantial resistance to bacterial proteases. By analogy, the Gp9 should be stable in the plastids, and therefore it is possible that the protein is unable to efficiently translocate into the chloroplasts. Therefore, future work could focus on transforming the plastid genome of *N. tabacum* for high level, stable expression of Gp9 without having to rely on chloroplast protein import.

Like the full length Gp9, shortened versions lacking the amino acid N-terminus domain maintain the stability and enzymatic activity of the full length parent proteins (Miller et al., 1998). To the best of our knowledge, this is the first study to transiently express this truncated protein in plants. Previously, truncated tailspikes have been expressed in *E. coli*, isolated and shown to reduce *Salmonella* colonization in the chicken GI tract. It is believed that the tailspikes bind to *Salmonella* and retard its motility and binding capability (Waseh *et al.*, 2010). Consequently, plant-produced Gp9 has the potential to serve as a prophylactic to control *Salmonella.* These proteins can potentially be added to chicken or other livestock feed, reducing contamination and reducing infections in humans. While the plant produced Gp9 did not yield statistically significant results in the chicken experiment, the results are indeed promising We were able to purify Gp9-ELP using a Myc column and were therefore able to show that Gp9-ELP exists as a trimer and remains active by binding to *S.* Typhimurium. However, it is still possible that the ELP tag could alter Gp9 conformation in vivo, influencing activity and future work could focus on non-fused Gp9. In vitro, ELPs have been reported to reduce protein activity compared to nonfused proteins (Shamji *et al.*, 2007; Kaldis *et al.*, 2013). It is possible that the physiological temperature of chickens could influence the solubility of Gp9-ELP. ELPs undergo a reversible phase transition from soluble to insoluble aggregates when heated past a certain transition temperature. However, this transition temperature can be manipulated by the altering the number

and/or the residue composition of the peptide repeats in order to adjust the solubility of Gp9ELP, thus allowing for better therapeutic efficacy at physiological temperatures (Hassouneh *et al.*, 2012). Furthermore, it is likely that the 1 log reduction in *Salmonella* colonization observed in this study compared to the 2 log reduction in the Waseh et al. (2010) study is due to the lower Gp9 dose administered to the birds (30 μg of purified His_6-Gp9 versus 7.7 μg Gp9-ELP in 35 mg of lyophilized plant tissue per dose) as well as the fact that the Gp9-ELP protein is contained within plant cells and thus the chicken's gut digestive enzymes need to digest through the plant cell wall in order to release the protein. In the previous study, it was shown that if the first dose was delivered at 18 hours instead of 1 hour post-infection, the benefits of the treatment were less. In our study, the protein would be released from the plant tissue sometime after the treatment is administered, so it is possible that this delay is partially responsible for the lower reduction in infection. Future studies will compare administration of higher levels of Gp9 protein and providing the feed throughout the experiment in order to more closely resemble the natural conditions in which this protein would be used.

One of our goals in this study was to demonstrate that Gp9 could be directly administered while in minimally-processed plant tissue, without having to rely on taxing purification and formulation techniques. It can be postulated that purified Gp9 may better reduce *Salmonella* colonization in chickens as previously reported (Waseh *et al.*, 2010). However, there is a slight, albeit non-significant, reduction in *S.* Typhimurium CFUs when wild type plant tissue was administered alone. This suggests that some innate factor in the plant tissue slurry is influencing *Salmonella* colonization and that administering Gp9 while in unprocessed plant material serves as a functional advantage rather than a hindrance. While future work should mainly focus on

increasing Gp9 accumulation in *N. benthamiana,* this study serves as another successful example of engineering plant factories for the production of a functional therapeutic.

Acknowledgements

We would like to thank Angelo Kaldis, Hong Zhu, Bernadette Beadle, and Cory Wenzel for their technical support and Alex Molnar for help with the figures. This work was funded by Agriculture and Agri-Food Canada's A-base grant #1107, and the Natural Sciences and Engineering Research Council of Canada Strategic Grant #397260. CMS is an Alberta Innovates Technology Futures iCORE Strategic Chair in Bacterial Glycomics.

3.5 References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Atterbury, R.J., Bergen, M.A.P. Van, Ortiz, F., Lovell, M.A., Harris, J.A., Boer, A. De, *et al.* (2007) Bacteriophage Therapy To Reduce Salmonella Colonization of Broiler Chickens. *Appl Environ Microbiol* **73**: 4543–4549

Baxa, U., Steinbacher, S., Miller, S., Weintraub, A., Huber, R., and Seckler, R. (1996) Interactions of phage P22 tails with their cellular receptor, Salmonella O-antigen polysaccharide. *Biophys J* **71**: 2040–2048

Bellucci, M., Marchis, F. De, Nicoletti, I., and Arcioni, S. (2007) Zeolin is a recombinant storage protein with different solubility and stability properties according to its localization in the endoplasmic reticulum or in the chloroplast. *J Biotechnol* **131**: 97–105.

Benton, C.B., King, J., and Clark, P.L. (2002) Characterization of the Protrimer Intermediate in the Folding Pathway of the Interdigitated β-Helix Tailspike Protein †. *Biochemistry* **41**: 5093–5103.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254

Chen, H.-M., Wang, Y., Su, L.-H., and Chiu, C.-H. (2013) Nontyphoid Salmonella Infection: Microbiology, Clinical Features, and Antimicrobial Therapy. *Pediatr Neonatol* **54**: 147–152

Conley, A.J., Joensuu, J.J., Jevnikar, A.M., Menassa, R., and Brandle, J.E. (2009) Optimization of elastin-like polypeptide fusions for expression and purification of recombinant proteins in plants. *Biotechnol Bioeng* **103**: 562–573

Conley, A.J., Joensuu, J.J., Menassa, R., Brandle, J.E., Stoger, E., Ma, J., *et al.* (2009) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol* **7**: 48

Conley, A.J., Joensuu, J.J., Richman, A., and Menassa, R. (2011) Protein body-inducing fusions for high-level production and purification of recombinant proteins in plants. *Plant Biotechnol J* **9**: 419–433

Cook, A., Odumeru, J., Lee, S., and Pollari, F. (2012) *Campylobacter*, *Salmonella*, *Listeria monocytogenes*, Verotoxigenic *Escherichia coli*, and *Escherichia coli* Prevalence, Enumeration, and Subtypes on Retail Chicken Breasts with and without Skin. *J Food Prot* **75**: 34–40.

Fanelli, M.J., Sadler, W.W., Franti, C.E., and Brownell, J.R. (1971) Localization of Salmonellae within the Intestinal Tract of Chickens. *Avian Dis* **15**: 366

Fischer, R., Schillberg, S., Hellwig, S., Twyman, R.M., and Drossard, J. (2012) GMP issues for recombinant plant-derived pharmaceutical proteins. *Biotechnol Adv* **30**: 434–439.

Goldenberg, D., and King, J. (1982) Trimeric intermediate in the in vivo folding and subunit assembly of the tail spike endorhamnosidase of bacteriophage P22. *Proc Natl Acad Sci U S A* **79**: 3403–7

Goode, D., Allen, V.M., and Barrow, P.A. (2003) Reduction of Experimental *Salmonella* and *Campylobacter* Contamination of Chicken Skin by Application of Lytic Bacteriophages. *Appl Environ Microbiol* **69**: 5032–5036

Guarner, F., and Malagelada, J.-R. (2003) Gut flora in health and disease. *Lancet* **361**: 512–519.

Hammack, T. (2011) Salmonella species in Bad Bug Book--Foodborne Pathogenic Microorganisms and Natural Toxins. US Food and Drug Administration, College Park, MD, .

Hassouneh, W., MacEwan, S.R., and Chilkoti, A. (2012) Chapter nine – Fusions of Elastin-Like Polypeptides to Pharmaceutical Proteins. In *Methods in Enzymology*. pp. 215–237.

Hyunjong, B., Lee, D.-S., and Hwang, I. (2005) Dual targeting of xylanase to chloroplasts and peroxisomes as a means to increase protein accumulation in plant cells. *J Exp Bot* **57**: 161–169

Joensuu, J.J., Conley, A.J., Lienemann, M., Brandle, J.E., Linder, M.B., and Menassa, R. (2010) Hydrophobin fusions for high-level transient protein expression and purification in *Nicotiana benthamiana*. *Plant Physiol* **152**: 622–33

Kaldis, A., Ahmad, A., Reid, A., McGarvey, B., Brandle, J., Ma, S., *et al.* (2013) High-level production of human interleukin-10 fusions in tobacco cell suspension cultures. *Plant Biotechnol J* **11**: 535–545

Kolotilin, I., Topp, E., Cox, E., Devriendt, B., Conrad, U., Joensuu, J., *et al.* (2014) Plant-based solutions for veterinary immunotherapeutics and prophylactics. *Vet Res* **45**: 117

Kropinski, A.M., Arutyunov, D., Foss, M., Cunningham, A., Ding, W., Singh, A., *et al.* (2011) Genome and proteome of *Campylobacter jejuni* bacteriophage NCTC 12673. *Appl Environ Microbiol* **77**: 8265–8271.

Kwon, K.-C., and Daniell, H. (2015) Low-cost oral delivery of protein drugs bioencapsulated in plant cells. *Plant Biotechnol J* **13**: 1017–1022

MacDonald, J., Doshi, K., Dussault, M., Hall, J.C., Holbrook, L., Jones, G., *et al.* (2015) Bringing plant-based veterinary vaccines to market: Managing regulatory and commercial hurdles. *Biotechnol Adv* **33**: 1572–1581.

Mead, P. (1999) Food-Related Illness and Death in the United States. *Emerg Infect Dis* **5**: 607–625

Miller, W.G., Bates, A.H., Horn, S.T., Brandl, M.T., Wachtel, M.R., and Mandrell, R.E. (2000) Detection on surfaces and in Caco-2 cells of *Campylobacter jejuni* cells transformed with new gfp, yfp, and cfp marker plasmids. *Appl Environ Microbiol* **66**: 5426–5436.

Oey, M., Lohse, M., Kreikemeyer, B., and Bock, R. (2009) Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *Plant J* **57**: 436–445

Patel, J., Zhu, H., Menassa, R., Gyenis, L., Richman, A., and Brandle, J. (2007) Elastin-like polypeptide fusions enhance the accumulation of recombinant proteins in tobacco leaves. *Transgenic Res* **16**: 239–249

Pillay, P., Schlüter, U., Wyk, S. van, Kunert, K.J., and Vorster, B.J. (2014) Proteolysis of

recombinant proteins in bioengineered plant cells. *Bioengineered* **5**: 15–20

Roach, D.R., and Donovan, D.M. (2015) Antimicrobial bacteriophage-derived proteins and therapeutic applications. *Bacteriophage* **5**: e1062590

Robinson, A.S., and King, J. (1997) Disulphide-bonded intermediate on the folding and assembly pathway of a non-disulphide bonded protein. *Nat Struct Biol* **4**: 450–455

Rodrigue, D.C., Tauxe, R. V., and Rowe, B. (1990) International increase in *Salmonella enteritidis*: A new pandemic? *Epidemiol Infect* **105**: 21

Saberianfar, R., Joensuu, J.J., Conley, A.J., and Menassa, R. (2015) Protein body formation in leaves of *Nicotiana benthamiana* : a concentration-dependent mechanism influenced by the presence of fusion tags. *Plant Biotechnol J* **13**: 927–937

Seckler, R., Fuchs, A., King, J., and Jaenicke, R. (1989) Reconstitution of the Thermostable Trimeric. **264**: 11750–11753.

Shamji, M.F., Betre, H., Kraus, V.B., Chen, J., Chilkoti, A., Pichika, R., *et al.* (2007) Development and characterization of a fusion protein between thermally responsive elastin-like polypeptide and interleukin-1 receptor antagonist: Sustained release of a local antiinflammatory therapeutic. *Arthritis Rheum* **56**: 3650–3661

Silhavy, D., Molnár, A., Lucioli, A., Szittya, G., Hornyik, C., Tavazza, M., and Burgyán, J. (2002) A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs. *EMBO J* **21**: 3070–3080

Simpson, D.J., Sacher, J.C., and Szymanski, C.M. (2015) Exploring the interactions between bacteriophage-encoded glycan binding proteins and carbohydrates. *Curr Opin Struct Biol* **34**: 69–77

Singh, A., Arutyunov, D., Szymanski, C.M., and Evoy, S. (2012) Bacteriophage based probes for pathogen detection. *Analyst* **137**: 3405–3421.

Skurnik, M., and Strauch, E. (2006) Phage therapy: Facts and fiction. *Int J Med Microbiol* **296**: 5–14.

Streatfield, S.J. (2007) Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnol J* **5**: 2–15

Su, L.-H., Chiu, C.-H., Chu, C., and Ou, J.T. (2004) Antimicrobial Resistance in Nontyphoid Salmonella Serotypes: A Global Challenge. *Clin Infect Dis* **39**: 546–551

Sulakvelidze, A., Alavidze, Z., and Morris, J.G. (2001) Bacteriophage Therapy. *Antimicrob Agents Chemother* **45**: 649–659

Waseh, S., Hanifi-Moghaddam, P., Coleman, R., Masotti, M., Ryan, S., Foss, M., *et al.* (2010) Orally Administered P22 Phage Tailspike Protein Reduces Salmonella Colonization in Chickens: Prospects of a Novel Therapy against Bacterial Infections. *PLoS One* **5**: e13904.

Chapter 4

Characterization of the *Escherichia coli* typing phage 1 receptor binding protein Gp145 and the capture of *Salmonella* Typhimurium and *E. coli* O157:H7 on paper using cellulose binding module tagged receptor binding proteins

Author Contributions: I conceived, designed and performed the experiments, analyzed the data and wrote the chapter. Denis Arytyunov designed the CBM-Gp9 construct. Bernadette Beadle performed the CBM-Gp9 capture experiments (Figures 5A and B). Christine Szymanski supervised this project and contributed the materials and reagents.

4.1 Introduction

Bacteriophages (phages), the viruses that infect bacteria, are the most abundant form of life on earth with up to an estimated 10^{31} phage particles (Suttle, 2005). These phages attach to their host bacterium through receptor binding proteins (RBPs) which recognize and attach to either a protein or a carbohydrate epitope on the surface of the bacterium (Simpson *et al.*, 2015). In the family *Caudoviridae* these RBPs come in two forms, tailspikes and tail fibres, tailspikes are broader in their structure and often have an enzymatic activity which allows them to cleave their carbohydrate receptor. Tail fibres have a much narrower structure but can also recognize either proteins or carbohydrates (Simpson *et al.*, 2016). In the past RBPs have been used to capture bacteria on surfaces (Singh *et al.*, 2009) or on beads (Poshtiban *et al.*, 2013), they have also been used as therapeutics in chickens to reduce *Salmonella enterica* serovar Typhimurium colonization in the cecum as well as stopping the spread of *S.* Typhimurium to the liver and the spleen (Waseh *et al.*, 2010). Moreover, we have recently shown that this protein can be produced in plants while still retaining its activity (Miletic *et al.*, 2016). We also have developed a new method to discover RBPs without the need for sequencing (Simpson *et al.*, 2016).

One of the best studied phages is P22, a member of the *Podoviridae* family of phages. P22 recognizes its *S.* Typhimurium host by first binding to the lipopolysaccharide (LPS) and then injecting its 42 kb genome followed by entering into either the lytic or the lysogenic cycle. If P22 enters the lytic cycle, it will take over the host cell and produce between 300-500 phage particles in approximately one hour (Calendar, 2006). Due to this ability of P22 to destroy the host cell, intact phages are not optimal for use in detection platforms. The trimeric RBP (Gp9) is the P22 tailspike protein and 6 of these tailspike trimers are located at the tip of the phage tail. Each 69kDa P22 Gp9 monomer binds to the repeating α-D-mannose-(1,4)-α-L-rhamnose-(1,3)-α-D-

galactose O-antigen of the *S.* Typhimurium LPS, and also has endorhamnosidase activity resulting in cleavage of the polysaccharide to assist in phage binding to the bacterial cell surface (Gemski and Stocker, 1967; Iwashita and Kanegasaki, 1976; Steinbacher *et al.*, 1996) at which point it injects its DNA which then takes over the host cell. In these studies, we use a modified version of the P22 RBP which has the first 106 amino acids, comprising the head binding domain, removed and a mutation, D392E, that removes the endorhamnosidase activity and increases the binding affinity.

In an earlier study, we identified Gp145 as the RBP for the *E. coli* O157:H7 typing phage 1 (ECTP1) (Simpson *et al.*, 2016). Gp145 bears homology to the T4 tail fibre Gp37 (Simpson *et al.*, 2016) which is a particularly well studied tail fibre that attaches to both the outer membrane porin protein C and LPS of *E. coli* K12. Both Gp145 and Gp37 have highly conserved histidines that in Gp37 coordinate iron ions (Bartual *et al.*, 2010). The T4 short tail fibre, Gp12, also has structural homology to Gp37 however it binds to LPS only and has a zinc ion coordinated by six histidines (Thomassen *et al.*, 2003). All of these proteins have a homo-trimeric beta helical-like structure.

Creating protein fusions is an excellent way to give proteins new applications. The carbohydrate binding domain, also known as the cellulose binding module, CBM9, is a protein tag that binds to β-D -Glcp-(1-4)-β-D -Glcp and is isolated from the C-terminus of xylanase 10A from *Thermotoga maritima* (Notenboom *et al.*, 2001). This module is commonly used as a tag that allows proteins to bind to paper (Hong *et al.*, 2008; Oliveira *et al.*, 2015). Increasingly the CBM tag is getting attention as a useful fusion peptide to immobilize proteins to an affordable substrate. A cellulose binding module has been used in the past to attach phage proteins such as T4 lysozyme to paper, creating an antimicrobial wound dressing (Abouhmad *et al.*, 2016). The

tag is also used because of cellulose's low non-specific binding which allows for more selective purification of recombinant proteins (Terpe, 2003).

This study focuses on two common pathogens, *S.* Typhimurium and *E. coli* O157:H7. *S.* Typhimurium is a common foodborne pathogen; this Gram-negative non-spore forming gamma proteobacterium belongs to the *Enterobactericae* family along with *E. coli*. A commensal in avian species, *S.* Typhimurium is commonly found in chicken meat as a contaminant from the butchering process (Revolledo *et al.*, 2006; Chambers and Gong, 2011). Estimates place non-typhoidal Salmonella as the annual cause of one million foodborne illnesses in the United States, with 19,000 hospitalizations and 380 deaths (Painter *et al.*, 2013). In Canada, Salmonella has been connected with half of produce related bacterial foodborne outbreaks (Kozak *et al.*, 2013). This is likely due to its ability to contaminate produce by internalizing in plant tissues, entering either through natural openings like stomata, through damaged tissues or they being brought into the tissues along with contaminated water (Deering *et al.*, 2012).

Enterohemorragic *E. coli* (EHEC) is a serious human gut pathogen which causes hemorrhagic colitis and hemolytic uremic syndrome in humans. EHEC halts protein synthesis in the target cell using two different AB5 toxins known as Shiga toxins (Saxena *et al.*, 1989). *E. coli* O157:H7 is believed to cause 63,000 foodborne illnesses in the United States, 2,100 hospitalizations and 20 deaths every year (Painter *et al.*, 2013). Cattle are asymptomatic carriers and are a common source of *E. coli* O157:H7 as their guts lack the receptor for the shiga toxin (Lim *et al.*, 2010). *E. coli* O157:H7 is commonly transmitted to humans through produce, and is able to invade into plant tissues using the same mechanisms as *S.* Typhimurium (Deering *et al.*, 2012).

This study further characterizes the newly identified *E. coli* O157:H7 RBP, Gp145, revealing that it binds to the O antigen of the LPS and that it is also capable of binding to *S*. Typhimurium. We created a fusion protein with Gp145 and with the P22, Gp9 to CBM9 which we then used to capture *E. coli* O157:H7 or *S*. Typhimurium, respectively, on paper from different suspensions.

4.2 Materials and methods

4.2.1 Bacterial strains and protein constructs

The His-Gp9 construct was acquired from Waseh *et al*. (Waseh *et al.*, 2010). His-CBM-Gp9 was synthesized by Bio Basic Lot UA64, it was cloned into the pET30a vector.

The His-Gp145 construct was created using Vent polymerase, the forward primer 5'-GCTGGTGGTACCATGGCAGTAGGTGAAATTCAAATTAGTGCC-3' and the reverse primer 5'-GGCGCTAAGCTTTTAAGCTGTTCTCTGCCACACATACACTACAAAAG-3' to clone the gene out of the plasmid containing the gene from the library from Simpson *et al* 2016 (Simpson *et al.*, 2016). The gene was ligated using T4 DNA ligase (Thermo Scientific) into the pET30a vector at the KpnI and HindII restriction sites of the vector.

His-CBM-Gp145 was created by cloning out the CBM module from His-CBM-Gp9 using the forward primer 5'-TCTATAGATCTAATGGTGGCGACCGCGAAATATGGC-3' and the reverse primer 5'-AATTAAGGTACCGGATCCTCCTCCTCCGG-3' and then ligated into the construct using the BglI and KpnI restriction sites.

E. coli O157:H7 C-8299-83, O157:H7 A8188-B3 and ECTP1 were acquired from the Félix d'Herelle Reference Center. *E. coli* O157:H7 C-8299-83 is the propagating strain for ECTP1 and was used for propagation, plaque assays, far Westerns and ELISAs, however we were unable to create a Green Flourescent Protein (GFP) producing version of this bacterium so *E. coli*

O157:H7 A8188-B3 was used instead for the capture of *E. coli* on paper, ECTP1 infect both strains. *S.* Typhimurium ATCC 19585 was originally acquired from the American Type Culture Collection. Both *E. coli* O157:H7 C-8299-83 and *S.* Typhimurium were transformed with pWM1007 (Miller *et al.*, 2000) through electroporation in order to make these bacteria fluorescent.

4.2.2 Protein expression and purification

His-Gp9 and His-Gp145 were transformed into chemically competent *E. coli* BL21(DE3) and grown at 37°C to an OD_{600} of 0.5 and then induced with 0.1 mM IPTG for 3 hours. The cultures were centrifuged, resuspended in IMAC A (30mM imidazole, 1M NaCl in 1x Phosphate Buffered Saline (PBS) pH 7.4) with Complete Tablets (Roche) and lysed with a French press at 14000psi. The protein was purified out of the lysate using a Ni-NTA (Qiagen) column, the buffer IMAC A was used for washing and IMAC B (same as IMAC A except with 500 mM imidazole) was used to elute the protein. The proteins were then cleaned on a PD-10 desalting column (GE Healthcare Life Sciences) according to manufacturer's instructions. Both CBM-Gp9 and CBM-Gp145 were expressed and purified the same way as unfused proteins above except the induction period was 2 hours instead of 3.

4.2.3 Protein stability assay

In order to test the stability and protease resistance of Gp145, 5µg of the protein was boiled for 10 minutes in the presence of 1% sodium dodecyl sulfate (SDS) and 0.8% β-mercaptoethanol (BME) or left unboiled in the same buffer. Gp145 was also incubated with 1 µg of trypsin, chymotrypsin or proteinase K for 1 hour at 37°C in 100 mM Tris, 10mM $CaCl_2$ at a pH of 8 in order to determine its level of protease resistance. BSA was treated with the same conditions as a control.

4.2.4 ELISA

The ELISAs were performed using Corning Costar 96 well assay plates. Bacteria were resuspended in PBS at an OD_{600} of 0.5, 100 µL was incubated overnight in each well at 4°C, the wells were blocked with 200 µL of 5% skim milk in PBST, the wells were probed with 100 ng of the Gp145 and then with a 1:1000 dilution of His-probe-(H3)-HRP form Santa Cruz Biotechnology. The ELISA was developed using TMB (Cell Signaling Technologies) and stopped with 2 M H_2SO_4. The wells were then measured at 450 nm.

4.2.5 Plaque assay

Fifty µL of an overnight culture of *S.* Typhimurium or *E. coli* O157:H7 was added to 5 mL of 0.6% LB molten agar (55°C) and then poured on 1.5% LB agar plates. Once the agar was solidified, 10 µL of either P22 or ECTP1 at 10^8 pfu/mL was spotted on the plate, these spots were allowed to dry and the plates were incubated at 37°C overnight.

4.2.6 Far Western

Overnight cultures of bacteria were resuspended in PBS at an OD_{600} of 0.5. These samples were boiled in SDS loading buffer with β-mercaptoethanol for 10 minutes and either treated with proteinase K (Thermo Scientific) overnight at 55°C or frozen. The samples were separated on a 12% SDS-PAGE and run at 150 V for 80 minutes. The gel was transferred to a PVDF membrane, which was then blocked in 5% skim milk in PBST, follow by incubation in 100 ng/mL of His-Gp145 in 5% skim milk followed by a 1:1000 dilution of His-probe-(H3)-HRP in skim milk. The membrane was developed using Western Lightning Plus-ECL (PerkinElmer) and photographed using a ChemiDoc MP (BioRad).

Silver staining was done using the Tsai-Frasch method (Tsai and Frasch, 1982).

4.2.7 Capturing bacteria on paper

Hole punch sized pieces of blotting paper 703 (VWR) were incubated with 20 µg of CBM-Gp9, CBM-Gp145, Gp9, Gp145 or no protein at all, in 50 µL of PBS for 30 min, the papers were then washed 3x with PBS and blocked for 1 hour with 10% BSA, washed 3x with PBS and then incubated for 30 min with 100 cfu of GFP expressing *S.* Typhimurium for the Gp9 constructs, or 100 cfu of GFP expressing *E. coli* O157:H7 for the Gp145 constructs. Afterwards, the papers were washed 3x in PBS and placed on LB-Kn agar plates overnight at 37°C. After overnight growth, the membranes were imaged with a FujiFilm FLA-5000 system using the 473 nm laser at 400 V for excitation and LPB (Y510) filter for emission.

This same experiment was repeated using chicken cecal contents diluted in PBS at a 1:1 ratiowith a GFP expressing *S.* Typhimurium at a concentration of 10^6 cfu/ml.

GFP expressing *S.* Typhimurium was added to 50 mL of buffered peptone water at a concentration of 10^8, 10^6, 10^4 or 10^2 cfu/mL and incubated with a commercially purchased chicken breast for 1 hour at room temperature. The washings from the chicken breast were incubated with CBM-Gp9 attached to blotting paper as above.

4.3 Results

4.3.1 Gp145 is SDS and protease resistant.

Gp145 was able to retain its Oligomeric state in the presence of SDS, only after boiling for 10 minutes did it begin unfold into its monomeric form. Moreover, when treated with trypsin or chymotrypsin, the protein was only partially degraded. Even in the presence of proteinase K, Gp145 showed resistance to degradation. BSA however, showed no difference between the

boiled and unboiled samples, but the protease treated samples were completely degraded (Figure 4.1).

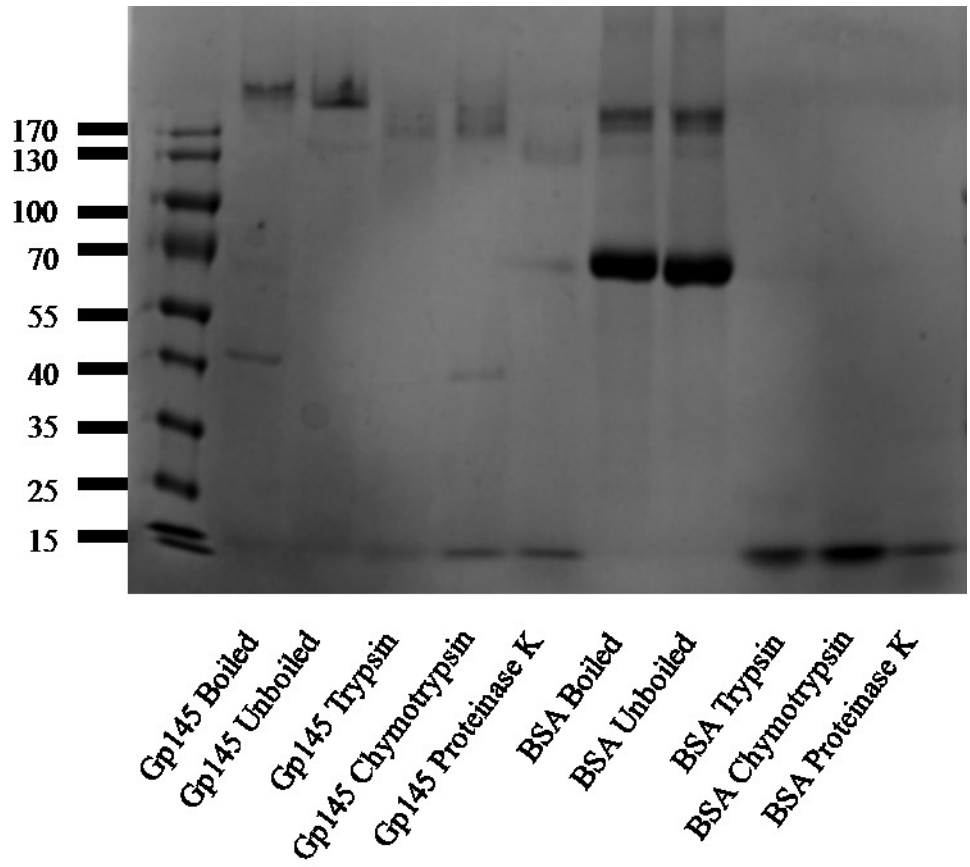


Figure 4.1 SDS-PAGE gel of protein degradation. Five ug of purified His-Gp145 or BSA was either boiled for 10 minutes, left unboiled in SDS loading dye with BME or treated with 1 µg of trypsin, chymotrypsin or proteinase K for 1 hour at 37°C. The MW of His-Gp145 is 45 kDa and the MW of BSA is 69 kDa.

4.3.2 Gp145 binds to *E. coli* O157:H7 and *S.* Typhimurium but not to APEC or DH5a.

An ELISA was performed testing whether Gp145 could bind to *S.* Typhimurium, *E. coli* O157:H7, avian pathogenic *E. coli* (APEC) O78 or *E. coli* DH5α. It was discovered that Gp145

could bind to both *E. coli* O157:H7 and *S.* Typhimurium. Gp145 was also able to bind to proteinase K treated *E. coli* O157:H7 but not to proteinase K treated *S.* Typhimurium. The protein was unable to bind to either APEC O78 or to *E. coli* DH5α regardless of whether or not the cells were treated with proteinase K (Figure 4.2 A). Plaque assays performed using the whole phage showed that ECTP1 cannot lyse *S.* Typhimurium despite the fact that Gp145 can bind to it (Figure 4.2 B).

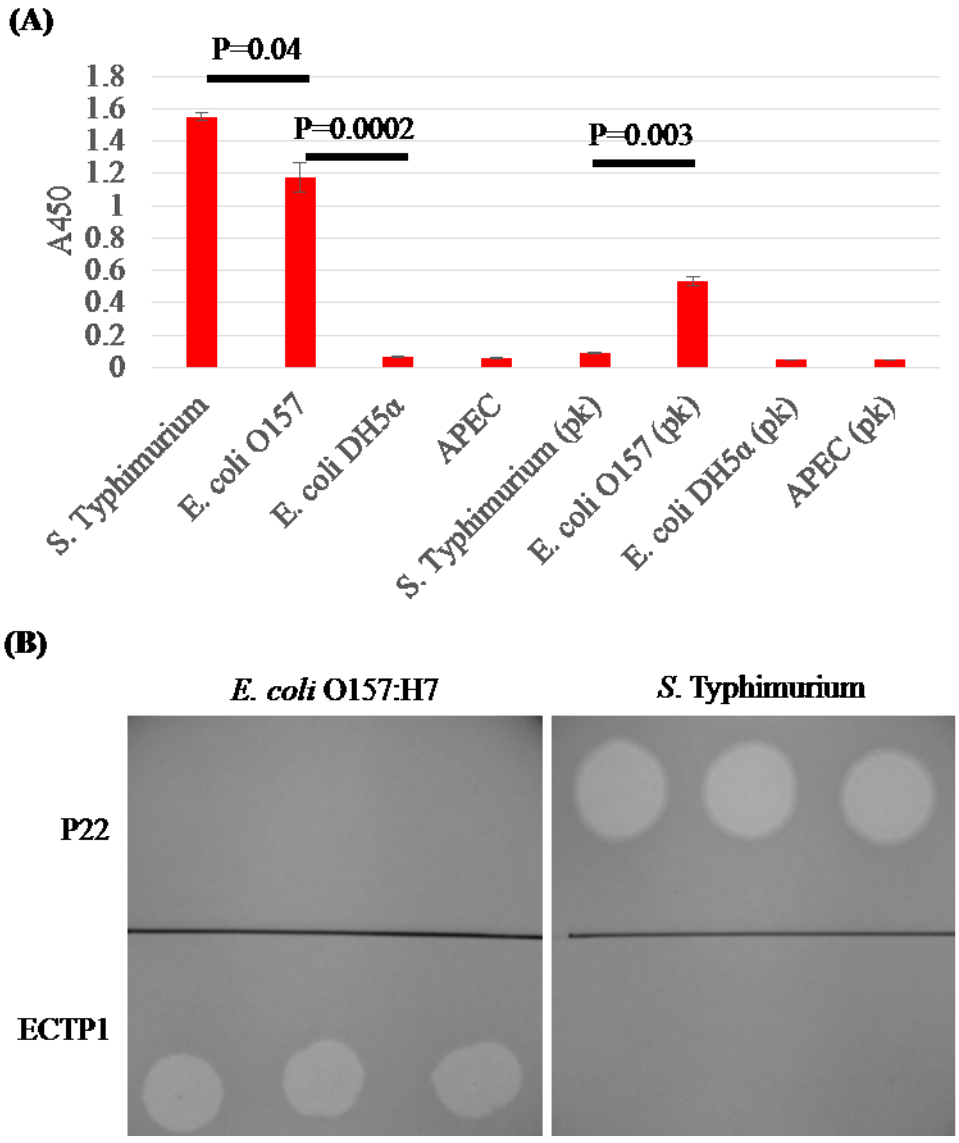


Figure 4.2 (A) An ELISA of the binding range of Gp145. Absorbance at 450nm of ELISA probed with His-Gp145 and α-His-HRP and developed with TMB and H_2SO_4. Wells were coated with *S*. Typhimurium, *E. coil* O157:H7, *E. coli* DH5α or APEC either untreated or treated with proteinase K (pk). Error bars indicate standard error, n=3. **(B) Plaque assay of P22 and ECTP1 on *E. coli* O157**:H7 **and *S*. Typhimurium**. P22 forms plaques on *S*. Typhimurium, but not on *E. coli* O157:H7. Likewise, ECTP1 forms plaques on *E. coli* O157:H7, but not on *S*. Typhimurium.

4.3.3 Gp145 binds to LPS in *E. coli* O157:H7 but not in *S*. Typhimurium.

In order to characterize the receptor of Gp145, whole cell lysates of *S*. Typhimurium, *E. coli* O157:H7, *E. coli* DH5α and APEC were run on a 12% SDS-PAGE gel and transferred to a PVDF membrane. A far Western was performed using His-Gp145 and α-His-HRP antibodies. A signal was visible around the 15 kDa marker which is where the LPS inner core is likely migrate, this signal was also found in the proteinase treated *E. coli* O157:H7 lane indicating that the receptor in *E. coli* O157:H7 is not a protein. No signal was visible at any molecular weight in either of the *S*. Typhimurium lanes. As expected, there was also no visible signal in any of the *E. coli* DH5α or APEC lanes (Figure 4.3).

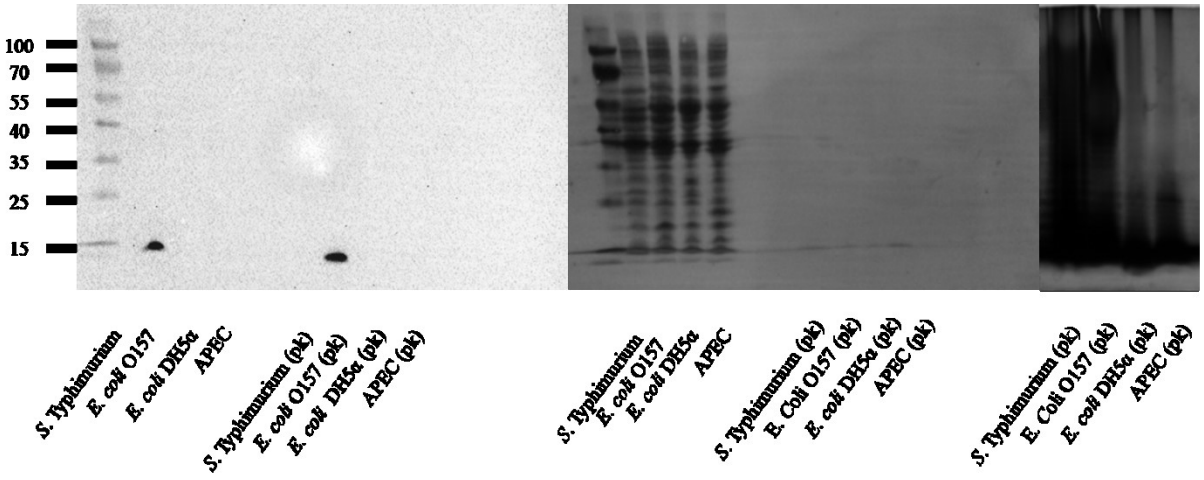


Figure 4.3 Far Western and coomassie stained membrane showing Gp145 binding. Left, far-Western of whole cell lysates using His-Gp145 and α-His-HRP as probes, pk denotes proteinase k treatement of the sample. Middle, coomassie stain of the same membrane showing complete degradation of all the proteins in the proteinase k treated samples. Right, silver stain of proteinase k treated samples.

4.3.4 CBM-Gp9 and CBM-Gp145 form SDS resistant species.

Like their parent proteins CBM-Gp9 and CBM-Gp145 form multimeric structures that are resistant to SDS but become monomers by boiling in SDS (Figure 4.3). Both proteins appear slightly above their predicted size, 82 kDa for CBM-Gp9 and 64 kDa for CBM-Gp145.

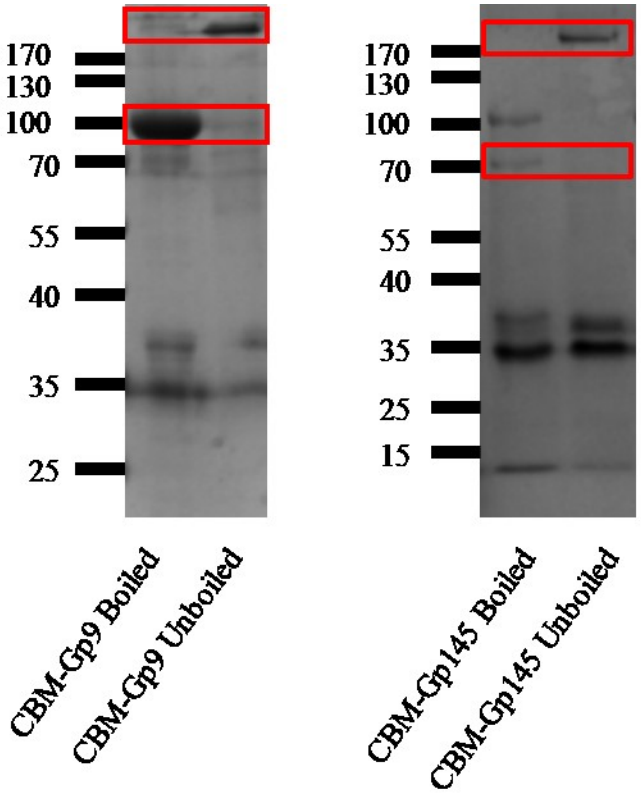


Figure 4.4 SDS resistance of CBM constructs. CBM-Gp9 (left) and CBM-Gp145 (right) were run on a SDS-PAGE gel either with or without boiling in SDS. Both unboiled samples appear at a much larger molecular weight than the boiled samples.

4.3.5 CBM-Gp9 captures *S.* Typhimurium and CBM-Gp145 captures *E. coli* O157:H7 on paper.

The CBM-Gp9 constructs bound to blotting paper and were able to capture *S.* Typhimurium from PBS, diluted cecal contents and from buffered peptone water that had washed chicken breasts (chicken juices). The discs were able to capture *S.* Typhimurium from PBS with bacterial concentrations as low as 100 cfu out of 100 µL. Moreover, the paper was able to capture *S.* Typhimurium from cecal contents where the bacteria were at a concentration of 10^6 cfu/mL (Figure 4.5 A). The protein bound to paper was also able to capture *S.* Typhimurium from chicken juices at concentrations as low as 100 cfu/ml (Figure 4.5 B). The CBM-Gp145 protein was also able to capture its target bacterium, *E. coli* O157:H7 at a concentration of 100 cfu in 100 µl in PBS (Figure 4.5 C).

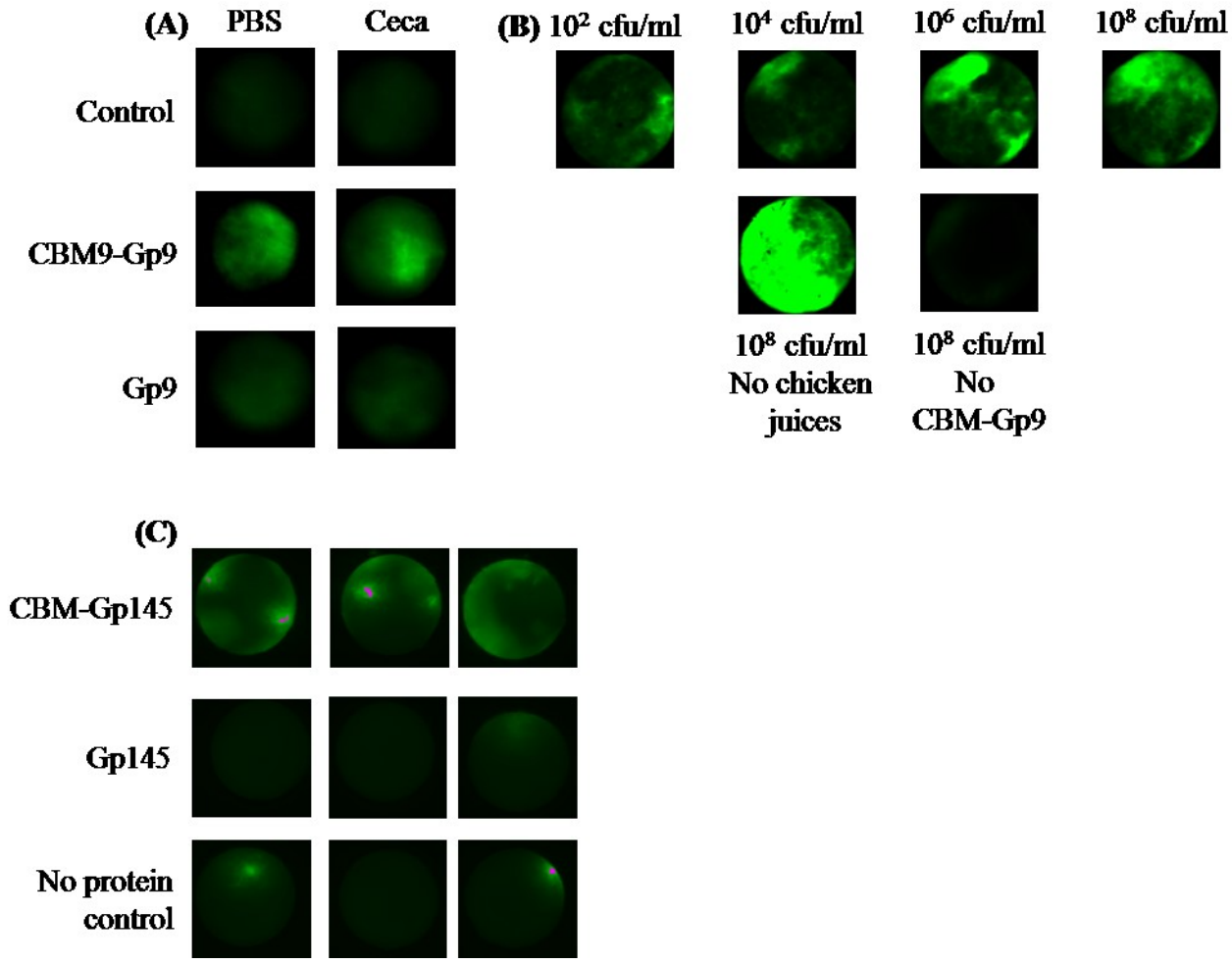


Figure 4.5 (A) *S.* Typhimurium captured from PBS and ceca. CBM-Gp9 was able to capture S. Typhimurium from PBS (10^3 cfu/ml) and diluted ceca (10^6 cfu/ml) better than Gp9 alone or paper alone. **(B) *S.* Typhimurium captured from chicken washes.** CBM-Gp9 capture of *S.* Typhimurium from water that had been used to wash raw chicken at bacterial concentrations of 10^2, 10^4, 10^6 and 10^8 cfu/ml, and from buffered water that had not been incubated with chicken breasts.. **(C) CBM-Gp145 captures *E. coli* O157**:H7 **from PBS**. The CBM-Gp145, Gp145 and blocked paper capture of *E. coli* O157:H7 at a concentration of 10^3 cfu/ml on paper.

4.4 Discussion

Bacteriophage RBPs bind strongly and specifically to their host bacterium and these properties give them tremendous potential as diagnostic tools (Singh *et al.*, 2012; Simpson *et al.*, 2015). In this study, we further characterized the newly discovered RBP Gp145, a tail fibre from an *E. coli* O157:H7 infecting *Myoviridae*. The ability of this RBP to bind selectively to both *E. coli* O157:H7 and *S.* Typhimurium while not binding to DH5α or APEC O78 makes it a useful tool in detecting these human pathogens and shows that the protein has great potential for use as part of an alone or as part of a suite of species specific RBPs.

Gp145 retains its monomeric form in the presence of SDS demonstrating the incredible stability of this protein which is consistent with what is seen with other already characterized RBPs (Goldenberg *et al.*, 1982). Moreover, the resistance of this protein to trypsin, chymotrypsin and even proteinase K shows the considerable hardiness of this protein. While tail fibres are less studied than tailspikes, it is known that some tail fibres have protease resistance domains, T4's Gp12 possess trypsin and chymotrypsin resistant domains (van Raaij *et al.*, 2001; van Raaij *et al.*, 2001) as does Gp37 (Bartual *et al.*, 2010) these domains are often used in crystallization experiments.

Gp145 is able to bind to proteinase k treated *E. coli* O157:H7 both in an ELISA and in a far Western, this indicates that its receptor is not a protein. Since, LPS is a dominant epitope on the surface of *E. coli*, is not susceptible to proteases, and often migrates close to the 10-15 kDa marker on an SDS-PAGE gel, it is likely the receptor of Gp145 is LPS. However, the far Western also does not show signs of laddering that is common with LPS due to the varying number of repeats of the O-antigen, perhaps indicating that Gp145 binds preferentially to the LPS core or to a single repeat of the O-antigen. While Gp145 is able to capture *S.* Typhimurium, it does not bind to any structures in either of the *S.* Typhimurium lanes in the far Western. It may

be that Gp145 binds to a three-dimensional epitope that is denatured by SDS-PAGE and Western transfer. Interestingly, ECTP1 does not infect *S.* Typhimurium despite binding to it, it may be that *S.* Typhimurium does not have the necessary primary/secondary receptor or it may have some other means of phage defense such as a restriction modification system or a CRISPR system (Fabre *et al.*, 2012) that prevent infection. These results are consistent with T4's Gp37 which binds to LPS and a protein, OmpC, of *E. coli* K12. Gp145 appears to have two receptors, one is the LPS of *E. coli* O157:H7 and the other is likely a protein found on the outer membrane of *S.* Typhimurium. In the ELISA, Gp145 showed greater binding to the whole cells of *E. coli* O157:H7 compared to the proteinase K treated cells, this may be due to Gp145 binding to both the LPS as well as to an unidentified protein on the whole cells of *E. coli* O157:H7, this protein may be more abundant in *S.* Typhimurium resulting in higher binding to that bacterium. It may also be simply that the whole cells bound better to the bottom of the ELISA wells when compared to the proteinase K treated lysate resulting in more receptors being available for Gp145 to bind.

Both CBM-Gp9 and CBM-Gp145 were expressed and active, although both protein preparations have an unknown band at 35 kDa. This contaminant may be a degradation product of the CBM fusion however the CBM portion of the peptide is 22 kDa and Gp9 and Gp145 are 60 kDa and 41 kDa respectively so any degradation would have to occur within the RBPs, which are both extremely stable. In the future it may be best to create constructs with C-terminal His fusions, this was not done in these experiments because fusion tags on the C-terminus of Gp9 have interfered with binding in our experience, however we did see some activity with a C-terminal elastin-like polypeptide fusion in chapter 3, using a different expression system may also reduce

the problem. In the future all CBM fusions should be purified on a size exclusion column to remove any contaminants.

CBM-Gp9 and CBM-Gp145 were able to capture *S.* Typhimurium and *E. coli* O157:H7 respectively on paper. CBM-Gp9 was also effective is complex solutions such as chicken cecal contents and chicken breast washes that are more representative of actual samples that could be tested for this pathogen. Thus, these fusions proteins are a promising technology for the inexpensive detection of pathogens in produce. Future studies could attempt to enrich bacteria such as *E. coli* O157:H7 out of water by attaching the CBM fusions to a filter. Now that we have the capability to identify RBPs from different phages and fuse them to cellulose binding domains to anchor these proteins to paper, the possibilities for RBP exploitation are indeed exciting.

Acknowledgements

I would like to thanks Bernadette Beadle for her help purifying proteins and with Figures 5A and 5B.

4.5 References

Abouhmad, A., Mamo, G., Dishisha, T., Amin, M.A., and Hatti-Kaul, R. (2016) T4 lysozyme fused with cellulose-binding module for antimicrobial cellulosic wound dressing materials. *J Appl Microbiol* **121**: 115–125

Bartual, S.G., Otero, J.M., Garcia-Doval, C., Llamas-Saiz, A.L., Kahn, R., Fox, G.C., and Raaij, M.J. van (2010) Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proc Natl Acad Sci U S A* **107**: 20287–92

Calendar, R. (2006) *The Bacteriophages*. 2nd ed., Oxford University Press, Oxford. Chambers, J.R., and Gong, J. (2011) The intestinal microbiota and its modulation for Salmonella control in chickens. *Food Res Int* **44**: 3149–3159.

Deering, A.J., Mauer, L.J., and Pruitt, R.E. (2012) Internalization of *E. coli* O157:H7 and Salmonella spp. in plants: A review. *Food Res Int* **45**: 567–575.

Fabre, L., Zhang, J., Guigon, G., Hello, S. Le, Guibert, V., Accou-Demartin, M., *et al.* (2012) CRISPR Typing and Subtyping for Improved Laboratory Surveillance of Salmonella Infections. *PLoS One* **7**: e36995

Gemski, P., and Stocker, B.A. (1967) Transduction by bacteriophage P22 in nonsmooth mutants of *Salmonella* typhimurium. *J Bacteriol* **93**: 1588–1597.

Goldenberg, D.P., Berget, P.B., and King, J. (1982) Maturation of the tail spike endorhamnosidase of Salmonella phage P22. *J Biol Chem* **257**: 7864–7871.

Hong, J., Ye, X., Wang, Y., and Zhang, Y.-H.P. (2008) Bioseparation of recombinant cellulose-binding module-proteins by affinity adsorption on an ultra-high-capacity cellulosic adsorbent. *Anal Chim Acta* **621**: 193–199.

Iwashita, S., and Kanegasaki, S. (1976) Enzymic and molecular properties of base-plate parts of bacteriophage P22. *Eur J Biochem* **65**: 87–94.

Kozak, G.K., MacDonald, D., Landry, L., and Farber, J.M. (2013) Foodborne Outbreaks in Canada Linked to Produce: 2001 through 2009. *J Food Prot* **76**: 173–183

Lim, J.Y., Yoon, J., and Hovde, C.J. (2010) A brief overview of *Escherichia coli* O157:H7 and its plasmid O157. *J Microbiol Biotechnol* **20**: 5–14

Miletic, S., Simpson, D.J., Szymanski, C.M., Deyholos, M.K., and Menassa, R. (2016) A Plant-Produced Bacteriophage Tailspike Protein for the Control of Salmonella. *Front Plant Sci* **6**: 1221.

Miller, W.G., Bates, A.H., Horn, S.T., Brandl, M.T., Wachtel, M.R., and Mandrell, R.E. (2000) Detection on surfaces and in Caco-2 cells of *Campylobacter jejuni* cells transformed with new gfp, yfp, and cfp marker plasmids. *Appl Environ Microbiol* **66**: 5426–5436.

Notenboom, V., Boraston, A.B., Kilburn, D.G., and Rose, D.R. (2001) Crystal Structures of the Family 9 Carbohydrate-Binding Module from *Thermotoga maritima* Xylanase 10A in Native and Ligand-Bound Forms. *Biochemistry* **40**: 6248–6256 http://dx.doi.org/10.1021/bi0101704.

Oliveira, C., Carvalho, V., Domingues, L., and Gama, F.M. (2015) Recombinant CBM-fusion technology — Applications overview. *Biotechnol Adv* **33**: 358–369.

Painter, J.A., Hoekstra, R.M., Ayers, T., Tauxe, R. V., Braden, C.R., Angulo, F.J., and Griffin, P.M. (2013) Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 1998–2008. *Emerg Infect Dis* **19**: 407–415

Poshtiban, S., Javed, M.A., Arutyunov, D., Singh, A., Banting, G., Szymanski, C.M., and Evoy, S. (2013) Phage receptor binding protein-based magnetic enrichment method as an aid for real time PCR detection of foodborne bacteria. *Analyst* **138**: 5619–5626

Raaij, M.J. van, Schoehn, G., Burda, M.R., and Miller, S. (2001) Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. *J Mol Biol* **314**: 1137–1146

Raaij, M.J. van, Schoehn, G., Jaquinod, M., Ashman, K., Burda, M.R., and Miller, S. (2001) Identification and Crystallisation of a Heat- and Protease-Stable Fragment of the Bacteriophage T4 Short Tail Fibre. *Biol Chem* **382**: 1049–1055

Revolledo, L., Ferreira, A.J.P., and Mead, G.C. (2006) Prospects in Salmonella Control: Competitive Exclusion, Probiotics, and Enhancement of Avian Intestinal Immunity. *J Appl Poult Res* **15**: 341–351

Saxena, S.K., O'Brien, A.D., and Ackerman, E.J. (1989) Shiga toxin, Shiga-like toxin II variant, and ricin are all single-site RNA N-glycosidases of 28 S RNA when microinjected into Xenopus oocytes. *J Biol Chem* **264**: 596–601

Simpson, D., Sacher, J., and Szymanski, C. (2016) Development of an Assay for the Identification of Receptor Binding Proteins from Bacteriophages. *Viruses* **8**: 17

Simpson, D.J., Sacher, J.C., and Szymanski, C.M. (2015) Exploring the interactions between bacteriophage-encoded glycan binding proteins and carbohydrates. *Curr Opin Struct Biol* **34**: 69–77

Singh, A., Arutyunov, D., Szymanski, C.M., and Evoy, S. (2012) Bacteriophage based probes for pathogen detection. *Analyst* **137**: 3405–3421.

Singh, A., Glass, N., Tolba, M., Brovko, L., Griffiths, M., and Evoy, S. (2009) Immobilization of bacteriophages on gold surfaces for the specific capture of pathogens. *Biosens Bioelectron* **24**: 3645–3651.

Steinbacher, S., Baxa, U., Miller, S., Weintraub, A., Seckler, R., and Huber, R. (1996) Crystal structure of phage P22 tailspike protein complexed with Salmonella sp. O-antigen receptors. *Proc Natl Acad Sci U S A* **93**: 10584–10588

Suttle, C. a (2005) Viruses in the sea. *Nature* **437**: 356–361.

Terpe, K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol* **60**: 523–533

Thomassen, E., Gielen, G., Schütz, M., Schoehn, G., Abrahams, J.P., Miller, S., and Raaij, M.J. van (2003) The Structure of the Receptor-binding Domain of the Bacteriophage T4 Short Tail Fibre Reveals a Knitted Trimeric Metal-binding Fold. *J Mol Biol* **331**: 361–373.

Tsai, C.-M., and Frasch, C.E. (1982) A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. *Anal Biochem* **119**: 115–119

Waseh, S., Hanifi-Moghaddam, P., Coleman, R., Masotti, M., Ryan, S., Foss, M., *et al.* (2010) Orally Administered P22 Phage Tailspike Protein Reduces Salmonella Colonization in Chickens: Prospects of a Novel Therapy against Bacterial Infections. *PLoS One* **5**: e13904.

Chapter 5

Conclusions and future directions

5.1 The importance of research

Over the last decade, phage research has exhibited a resurged interest, especially in the area of phage therapy. One alternative to using the whole phage is to use select phage proteins in order to isolate specific features of a phage without having to deal with a whole organism, which may include unwanted features such as bacterial lysis invoking resistance development or transfer of pathogenicity genes. Receptor binding proteins (RBPs) are some of the most biotechnologically relevant phage proteins, as they have evolved and continue to evolve to specifically and effectively bind their bacterial hosts. This constant evolutionary arms race dynamic between phages and bacteria over time leads to a strong selection for variant RBPs resulting in a considerable diversity of binding targets, making them attractive in commercial exploitation and important for understanding the lifecycle of any phage (Ferris *et al.*, 2007). Utilizing phage RBPs as surrogate antibodies opens up a great amount of potential for improvement of existing diagnostic technologies as well as the development of new ones. The advantages of using RBPs include their greater stability and their superior ability to bind to carbohydrate receptors. Immobilizing RBPs on chips for use in surface plasmon resonance spectroscopy analysis of binding interactions (Naidoo *et al.*, 2012) and to magnetic microbeads, we have further demonstrated the utility of phage RBPs (Singh *et al.*, 2011).

RBPs also have therapeutic applications, for instance, *S.* Typhimurium colonization of the chicken ceca, as well as the bacterial spread to the liver and spleen can be reduced by oral administration of recombinant P22 RBP (Waseh *et al.*, 2010). Like most RBPs discovered so far, the phage P22 RBP is resistant to chicken cecal proteases, moreover, it agglutinates *S.* Typhimurium cells and reduces their motility *in vitro* (Waseh *et al.*, 2010). The protein shows similar agglutination levels as antibodies recognizing the *S.* Typhimurium LPS (Sigurskjold *et*

al., 1991; Cygler *et al.*, 1991; Waseh *et al.*, 2010), but is capable of binding to a wider range of *S.* Typhimurium serotypes.

However, the fact remains that phage RBPs and the bacterial receptors they recognize need to be well characterized and carefully chosen in order to ensure the success of any technology. Therefore, it is important to be able to rapidly identify and characterize RBPs. It is also necessary to be able to develop inexpensive methods to produce and deliver these therapeutics. Finally, affordable diagnostic platforms need to be created to be able to utilize these proteins as an affordable diagnostic. This thesis has attempted to achieve several of these goals through the development of an RBP discovery assay, the expression of RBPs in plants and the fusion of RBPs to the Cellulose Binding module (CBM9) tag allowing for the creation of an inexpensive diagnostic technology.

5.2 Main conclusions

Chapter 2 describes a method for the discovery of RBPs from phages that is independent of sequencing and bioinformatic tools. In this method, the phage DNA is extracted, fragmented, ligated into an expression vector, transformed into *E. coli* and then transferred to a nitrocellulose membrane where recombinant proteins are overexpressed. The membrane is then probed for binding to the phage host bacterium.

Contrary to what was expected, the inserts appear to be using the vector promoter rather than the native phage promoter. However, protein expression most likely involves the native start codon of the individual genes, as opposed to that of the vector, since the inserts were found to be translated in different frames.

This assay allows for the discovery of truly novel and unrelated RBPs since the assay is based solely on the ability to bind to a host bacterium, thus it represents an improvement over previous homology or synteny based methods. A comparable assay has been published for the discovery of phage lysins from both isolated phages and from metagenomic samples (Schmitz *et al.*, 2008). With this in mind, it seems likely that DNA from an environmental sample could be used to discover RBPs that originate from phages that cannot be propagated under laboratory conditions, which would help elucidate the characteristics of these phages. However, this assay is currently limited to phage of culturable bacteria and requires any chaperone to be in close proximity to the RBP in the phage genome. This method could be combined with other detection methods, such as antibodies or through the biotinylation of the bacteria, to detect bound organisms to the nitrocellulose membranes, circumventing the need to culture the target organism.

This assay should be applicable for a phage that does not infect *Enterobactericae* since a wealth of proteins from other species have been expressed in *E. coli* including RBPs from Gram-positive bacteria (Arutyunov *et al.*, 2014) (Smith *et al.*, 2005). Therefore, it is likely that this assay will be amenable to RBP discovery from phages targeting other species. We did attempt to repeat this assay using the *Campylobacter jejuni* phage NCTC 12673, however while the assay worked very well when using known RBP Gp047 constructs we had difficulty creating libraries with the phage DNA and the libraries that we did create did not discover any RBPs including Gp047 which we knew to be in the genome. It is possible that *E. coli* was unable to use the phage promoters in this case reducing the likelihood of a RBP being expressed in a random library.

Chapter 3 shows that the P22 RBP can be transiently expressed in *N. benthamiana* at reasonable levels and is able to bind to *S.* Typhimurium. Interestingly, we discovered that Gp9 without either the Elastin Like Polypeptide (ELP) or Hydrophobin I (HFBI) tags accumulated at a higher

concentration than either ELP or HFBI tagged Gp9, it may be that the presence of a tag on the C-terminus interferes with the folding of the protein, since the formation of the β-helical structure at the C-terminal region is critical for the correct folding of the protein (Weigele *et al.*, 2005). It also seems likely that the c-Myc tag was degraded in the construct without the ELP or HFBI tags resulting in different amounts detected depending on whether the α-c-Myc or the α-Gp9 antibodies were used. This study was the first time a phage RBP was expressed in plants.

The plant extract was able to slightly reduce the colonization of *S.* Typhimurium in chickens when the protocol from Waseh et al. was used (Waseh *et al.*, 2010), however we were unable to achieve as high a concentration of Gp9 in the plant extract as used in the previous study and so the reduction in colonization was also less and did not reach statistical significance. Another potential cause of the reduced effect we see may be the fact that the tailspikes are contained within the tissues of the plant delaying their release into the gut of the chicken, allowing the *S.* Typhimurium an opportunity to colonize the intestines before encountering the RBPs, in the previous study a delay in the administration of Gp9 greatly reduced the effect of the treatment and this may be what is occurring here. We were also unable to purify enough protein from the plants to use in the experiment but this would be an excellent future experiment to further this study. While these experiments were done with infiltrated leaves creating a line of Gp9 expressing *N. benthamiana* would give enough plant matter to purify any desired amount of protein.

Chapter 4 builds upon the previous work in chapter 2 by further characterizing the newly discovered RBP Gp145, a tail fibre from an *E. coli* O157:H7 infecting *Myoviridae*. Gp145 was able to bind to both *E. coli* O157:H7 and *S.* Typhimurium while not binding to *E. coli* DH5α or avian pathogenic *E. coli* O78, making it a useful tool in detecting the human pathogens. Gp145

was shown to be very stable in the presence of SDS, requiring boiling to dissociate it down to its monomeric form, similar to other known RBPs (Goldenberg *et al.*, 1982). The protein was also resistant to trypsin, chymotrypsin, and proteinase K, this protease resistance is consistent with previous studies of tail fibres (Mark J. van Raaij *et al.*, 2001; Mark J van Raaij *et al.*, 2001; Bartual *et al.*, 2010).

Based on Gp145's ability to bind to proteinase K treated *E. coli* O157:H7, it is most likely that the phage recognizes the LPS of that organism, however, in *S.* Typhimurium there is no binding by Gp145 to the proteinase-treated samples only to whole cells suggesting that Gp145 recognizes a three-dimensional protein structure on *S.* Typhimurium and possibly on *E. coli* O157:H7 as well. Interestingly Gp145's binding range is different than the whole phage's host range since ECTP1 does not infect *S.* Typhimurium. Screening the binding of Gp145 against an *E. coli* O157:H7 LPS mutant library could help to discover the carbohydrate residues that are recognized by Gp145, initially Gp145 was screened against a *S.* Typhimurium LPS mutant library but the results were inconclusive. A pull down could be used to discover the receptor in *S.* Typhimurium, this was done with *E. coli* O157:H7 but no obvious candidates for binding were discovered. If the receptor on *S.* Typhimurium were discovered, a mutant in *E. coli* O157:H7 could be made to determine if that had any effect on binding. Once both receptors were found, the affinity of Gp145 for those receptors could be measured using surface plasmon resonance.

Both of the CBM9 RBP fusions were able to capture their target bacterium on paper and CBM-Gp9 was able to capture *S.* Typhimurium from chicken ceca and from washes of raw chicken. Further optimization of expression of both of these constructs is advisable for future work with these constructs. One method that has worked well with the mycobacterial binding RBP Gp10, is

to reduce the incubation temperature during expression, expressing the protein at 30°C or room temperature may yield better results.

The bands that appear at around 35 kDa could be sent for identification through mass spectrometry to discover the point of breakage and the construct could be modified to alter the site of degradation.

5.3 Future directions

5.3.1 Engineering protease sensitive receptor binding proteins to become resistant to host proteases

While phage RBPs are often stable in the environment in which the phage is naturally found, RBPs when expressed in isolation, may have reduced stability. Directed evolution has been used in the past to create robust enzymes for industrial applications (Turner, 2009); these same techniques could be used to make RBPs more stable under many desired conditions. Directed evolution could be used to create mutant RBPs which would then be screened under conditions of stress that would be encountered in the target environment. The NCTC 12673 phage protein Gp047 which binds to the flagella of *C. jejuni* would work well in a proof-of-principle study. This protein has been shown to be able to bind to the majority of C. jejuni and C. coli species (Javed *et al.*, 2013) making it useful for the detection of these pathogens, however, it is susceptible to degradation by trypsin and is unstable when fed to chickens (unpublished results). Directed evolution would be used to create a library of mutant versions of *gp047* in *E. coli*, lift these colonies onto a membrane, induce expression of the mutant genes, lyse the cells and then incubate them with a diluted solution of chicken gut proteases (unable to digest the flagellar receptor). After several washing steps with protease inhibitors, the membranes would be

incubated with *C. jejuni* cells in order to determine which inserts encode proteins that are resistant to the proteases, since only Gp047 derivatives able to withstand the protease treatment would be able to bind to the bacterium. Candidate genes would then be used as the starting material to repeat the process using more concentrated protease solutions. DNA shuffling is an excellent technique to evolve the protein and is one of the most common techniques currently used for directed evolution (Stemmer, 1994; Crameri *et al.*, 1998). This method could be used to create the library of mutant *gp047*genes, since this gene has a number of homologues in Genbank which can be used to shuffle the DNA with (Figure 5.1). DNA shuffling creates a mutant library by creating short random fragments of several homologues of the gene of interest and then subjecting them to a partial PCR reaction to create mosaics of the target genes. These gene variants can then be ligated into a vector and expressed in *E. coli*. Once refined, this technique could be used to make RBPs resistant to other types of stresses that tend to denature proteins such as heat, pH, oxidation or heavy metals, providing a way to customize proteins for functionality under any given medically and/or industrially relevant condition.

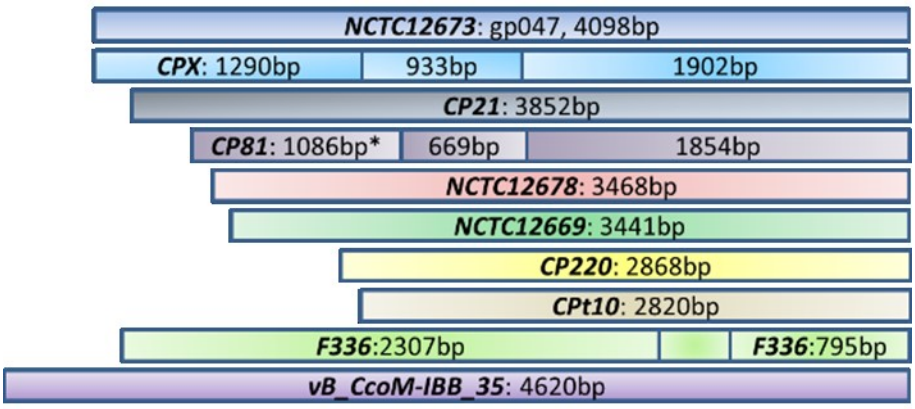


Figure 5.1 Schematic representation of campylobacter phage sequences with homology to NCTC 12673 Gp047. Homologues from phages NCTC 12678, NCTC 12669 and F336 were amplified and sequenced using primers specific for conserved regions of Gp047. Homology to

CPX, CP21, CP81, CP220, CPt10 and vB_CcoM-IBB was discovered using BLAST (Javed *et al.*, 2013).

5.3.2 Engineer receptor binding proteins to recognize desired receptors

Bacteriophages have been shown to be able to naturally change RBP binding specificity to infect new hosts or to continue to infect hosts that have lost the previous phage receptor (ie developed a mechanism for resistance) (Meyer *et al.*, 2012).The RBP of the *E. coli* phage Ox2 was shown to be able to change from protein to carbohydrate receptor specificity by changing as little as one amino acid (Drexler *et al.*, 1991). Also, it was shown that a single amino acid change in the RBP from *E. coli* phage HK620 increased its cell-binding affinity 1000-fold (Broeker *et al.*, 2013). One *Bordetella* phage, BPP-1, is known to use an error-prone reverse transcriptase to alter the nucleotide code of a hypervariable region of its RBP gene to expand its host range and to counter-evolve against changes in its host receptor (Liu *et al.*, 2002). Bacteriophage have also been found to use multiple RBPs as a way to increase their host range (Schwarzer *et al.*, 2012). Using a directed evolution approach to create a diverse RBP library and combining this technique with T7 phage display could emulate the process already found in nature to create RBPs that bind to a desired epitope. A similar technique has been developed to create custom-binding antibodies, however, the use of this technique with RBPs is more promising because RBPs are more robust than antibodies and whole phage particles which are susceptible to environmental factors such as desiccation.

Magnetic beads could be tagged with a desired polysaccharide or protein receptor and mixed with a library of mutant RBP-presenting T7 phages. The beads could then be washed and the phages would be eluted off and plated onto a lawn of *E. coli* to find phages that were able to bind to the receptor that was conjugated to the bead. This method has the benefits of being able to find

RBPs that bind to a target sugar or protein and also to select for high-affinity binding events if more stringent washing conditions are used to elute off low-affinity binding proteins. Also, just as with the protease-resistance assay, this assay allows for selection of RBPs that will bind in the environment in which they are to be used.

5.4 Final remarks

Bacteriophage RBPs are a promising technology with applications in both therapeutics and diagnostics. The discoveries made and technologies developed in this thesis expand on the knowledge of the useful proteins and open up new areas for their application and further discovery.

5.5 References

Arutyunov, D., Singh, U., El-Hawiet, A., Seckler, H. dos S., Nikjah, S., Joe, M., *et al.* (2014) Mycobacteriophage cell binding proteins for the capture of mycobacteria. *Bacteriophage* **4**: e960346 http://dx.doi.org/10.4161/21597073.2014.960346.

Bartual, S.G., Otero, J.M., Garcia-Doval, C., Llamas-Saiz, A.L., Kahn, R., Fox, G.C., and Raaij, M.J. van (2010) Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proc Natl Acad Sci U S A* **107**: 20287–92

Broeker, N.K., Gohlke, U., M??ller, J.J., Uetrecht, C., Heinemann, U., Seckler, R., *et al.* (2013) Single amino acid exchange in bacteriophage HK620 tailspike protein results in thousand-fold increase of its oligosaccharide affinity. *Glycobiology* **23**: 59–68.

Crameri, A., Raillard, S.A., Bermudez, E., and Stemmer, W.P. (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**: 288–291 http://dx.doi.org/10.1038/34663.

Cygler, M., Rose, D.R., and Bundle, D.R. (1991) Recognition of a cell-surface oligosaccharide of pathogenic Salmonella by an antibody Fab fragment. *Science* **253**: 442–445.

Drexler, K., Dannull, J., Hindennach, I., Mutschler, B., and Henning, U. (1991) Single mutations in a gene for a tail fiber component of an *Escherichia coli* phage can cause an extension from a protein to a carbohydrate as a receptor. *J Mol Biol* **219**: 655–663.

Ferris, M.T., Joyce, P., and Burch, C.L. (2007) High frequency of mutations that expand the host range of an RNA virus. *Genetics* **176**: 1013–1022.

Goldenberg, D.P., Berget, P.B., and King, J. (1982) Maturation of the tail spike endorhamnosidase of Salmonella phage P22. *J Biol Chem* **257**: 7864–7871.

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**.

Liu, M., Deora, R., Doulatov, S.R., Gingery, M., Eiserling, F.A., Preston, A., *et al.* (2002) Reverse Transcriptase-Mediated Tropism Switching in Bordetella Bacteriophage. *Science (80-)* **295**: 2091–2094

Meyer, J.R., Dobias, D.T., Weitz, J.S., Barrick, J.E., Quick, R.T., and Lenski, R.E. (2012) Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* **335**: 428–432

Naidoo, R., Singh, A., Arya, S.K., Beadle, B., Glass, N., Tanha, J., *et al.* (2012) Surface-immobilization of chromatographically purified bacteriophages for the optimized capture of bacteria. *Bacteriophage* **2**: 15–24

Raaij, M.J. van, Schoehn, G., Burda, M.R., and Miller, S. (2001) Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. *J Mol Biol* **314**: 1137–1146

Raaij, M.J. van, Schoehn, G., Jaquinod, M., Ashman, K., Burda, M.R., and Miller, S. (2001) Identification and Crystallisation of a Heat- and Protease-Stable Fragment of the Bacteriophage

T4 Short Tail Fibre. *Biol Chem* **382**: 1049–1055

Schmitz, J.E., Daniel, A., Collin, M., Schuch, R., and Fischetti, V.A. (2008) Rapid DNA library construction for functional genomic and metagenomic screening. *Appl Environ Microbiol* **74**: 1649–1652.

Schwarzer, D., Buettner, F.F.R., Browning, C., Nazarov, S., Rabsch, W., Bethe, A., *et al.* (2012) A Multivalent Adsorption Apparatus Explains the Broad Host Range of Phage phi92: a Comprehensive Genomic and Structural Analysis. *J Virol* **86**: 10384–10398

Sigurskjold, B.W., Altman, E., and Bundle, D.R. (1991) Sensitive titration microcalorimetric study of the binding of Salmonella O-antigenic oligosaccharides by a monoclonal antibody. *Eur J Biochem* **197**: 239–246

Singh, A., Arutyunov, D., McDermott, M.T., Szymanski, C.M., and Evoy, S. (2011) Specific detection of *Campylobacter jejuni* using the bacteriophage NCTC 12673 receptor binding protein as a probe. *Analyst* **136**: 4780–4786.

Smith, N.L., Taylor, E.J., Lindsay, A.-M., Charnock, S.J., Turkenburg, J.P., Dodson, E.J., *et al.* (2005) Structure of a group A streptococcal phage-encoded virulence factor reveals a catalytically active triple-stranded -helix. *Proc Natl Acad Sci* **102**: 17652–17657

Stemmer, W.P.C. (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**: 389–391 http://www.nature.com/doifinder/10.1038/370389a0.

Turner, N.J. (2009) Directed evolution drives the next generation of biocatalysts. *Nat Chem Biol* **5**: 567–573.

Waseh, S., Hanifi-Moghaddam, P., Coleman, R., Masotti, M., Ryan, S., Foss, M., *et al.* (2010) Orally Administered P22 Phage Tailspike Protein Reduces Salmonella Colonization in Chickens: Prospects of a Novel Therapy against Bacterial Infections. *PLoS One* **5**: e13904.

Weigele, P.R., Haase-Pettingell, C., Campbell, P.G., Gossard, D.C., and King, J. (2005) Stalled Folding Mutants in the Triple β-Helix Domain of the Phage P22 Tailspike Adhesin. *J Mol Biol* **354**: 1103–1117.

6 Bibliography

Abouhmad, A., Mamo, G., Dishisha, T., Amin, M.A., and Hatti-Kaul, R. (2016) T4 lysozyme fused with cellulose-binding module for antimicrobial cellulosic wound dressing materials. *J Appl Microbiol* **121**: 115–125

Ackermann, H. (2009) *Bacteriophages*. Humana Press, Totowa, NJ. http://link.springer.com/10.1007/978-1-60327-164-6.

Ahmad, A., O. Pereira, E., J. Conley, A., S. Richman, A., and Menassa, R. (2010) Green Biofactories: Recombinant Protein Production in Plants. *Recent Pat Biotechnol* **4**: 242–259

Ahmed, R., Bopp, C., Borczyk, A., and Kasatiya, S. (1987) Phage-typing scheme for *Escherichia coli* O157:H7. *J Infect Dis* **155**: 806–809

Allison, G.E., and Verma, N.K. (2000) Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol* **8**: 17–23.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Andres, D., Gohlke, U., Broeker, N.K., Schulze, S., Rabsch, W., Heinemann, U., *et al.* (2013) An essential serotype recognition pocket on phage P22 tailspike protein forces *Salmonella enterica* serovar Paratyphi A O-antigen fragments to bind as nonsolution conformers. *Glycobiology* **23**: 486–94

Andres, D., Hanke, C., Baxa, U., Seul, A., Barbirz, S., and Seckler, R. (2010) Tailspike interactions with lipopolysaccharide effect DNA ejection from phage P22 particles in vitro. *J Biol Chem* **285**: 36768–36775.

Arakawa, T., Yu, J., Chong, D.K.X., Hough, J., Engen, P.C., and Langridge, W.H.R. (1998) A plant-based cholera toxin B subunit–insulin fusion protein protects against the development of autoimmune diabetes. *Nat Biotechnol* **16**: 934–938

Arutyunov, D., Singh, U., El-Hawiet, A., Seckler, H. dos S., Nikjah, S., Joe, M., *et al.* (2014) Mycobacteriophage cell binding proteins for the capture of mycobacteria. *Bacteriophage* **4**: e960346

Atterbury, R.J., Bergen, M.A.P. Van, Ortiz, F., Lovell, M.A., Harris, J.A., Boer, A. De, *et al.* (2007) Bacteriophage Therapy To Reduce Salmonella Colonization of Broiler Chickens. *Appl Environ Microbiol* **73**: 4543–4549

Barbirz, S., Müller, J.J., Uetrecht, C., Clark, A.J., Heinemann, U., Seckler, R., *et al.* (2008) Crystal structure of *Escherichia coli* phage HK620 tailspike: podoviral tailspike endoglycosidase modules are evolutionarily related. *Mol Microbiol* **69**: 303–316.

Bartual, S.G., Garcia-Doval, C., Alonso, J., Schoehn, G., and Raaij, M.J. van (2010) Two-chaperone assisted soluble expression and purification of the bacteriophage T4 long tail fibre protein gp37. *Protein Expr Purif* **70**: 116–121

Bartual, S.G., Otero, J.M., Garcia-Doval, C., Llamas-Saiz, A.L., Kahn, R., Fox, G.C., and Raaij, M.J. van (2010) Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proc Natl Acad Sci U S A* **107**: 20287–92

Baxa, U., Steinbacher, S., Miller, S., Weintraub, A., Huber, R., and Seckler, R. (1996) Interactions of phage P22 tails with their cellular receptor, Salmonella O-antigen polysaccharide. *Biophys J* **71**: 2040–2048

Bellucci, M., Marchis, F. De, Nicoletti, I., and Arcioni, S. (2007) Zeolin is a recombinant storage protein with different solubility and stability properties according to its localization in the endoplasmic reticulum or in the chloroplast. *J Biotechnol* **131**: 97–105.

Benton, C.B., King, J., and Clark, P.L. (2002) Characterization of the Protrimer Intermediate in the Folding Pathway of the Interdigitated β-Helix Tailspike Protein †. *Biochemistry* **41**: 5093–5103

Benzer, S. (1961) On the Topography of the Genetic Fine Structure. *PNAS* **47**: 62–509

Berrington, A.W., Tan, Y.C., Srikhanta, Y., Kuipers, B., Ley, P. van der, Peak, I.R., and Jennings, M.P. (2002) Phase variation in meningococcal lipooligosaccharide biosynthesis genes. *FEMS Immunol Med Microbiol* **34**: 267–275.

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

Bondy-Denomy, J., and Davidson, A.R. (2014) When a virus is not a parasite: The beneficial effects of prophages on bacterial fitness. *J Microbiol* **52**: 235–242

Born, Y., Fieseler, L., Klumpp, J., Eugster, M.R., Zurfluh, K., Duffy, B., and Loessner, M.J. (2014) The tail-associated depolymerase of *Erwinia amylovora* phage L1 mediates host cell adsorption and enzymatic capsule removal, which can enhance infection by other phage. *Environ Microbiol* **16**: 2168–2180

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254

Briers, Y., Walmagh, M., Puyenbroeck, V. Van, Cornelissen, A., Cenens, W., Aertsen, A., *et al.* (2014) Engineered Endolysin-Based "Artilysins" To Combat Multidrug-Resistant Gram-Negative Pathogens. *MBio* **5**: e01379-14-e01379-14

Broeker, N.K., Gohlke, U., M??ller, J.J., Uetrecht, C., Heinemann, U., Seckler, R., *et al.* (2013) Single amino acid exchange in bacteriophage HK620 tailspike protein results in thousand-fold increase of its oligosaccharide affinity. *Glycobiology* **23**: 59–68.

Bruynoghe, R., and Maisin, J. (1921) Essais de thérapeutique au moyen du bacteriophage. *CR Soc Biol* **85**: 120–121.

Calendar, R. (2006) *The Bacteriophages*. 2nd ed., Oxford University Press, Oxford.

Casjens, S.R. (2008) Diversity among the tailed-bacteriophages that infect the Enterobacteriaceae. *Res Microbiol* **159**: 340–348.

Casjens, S.R., and Molineux, I.J. (2012) Short Noncontractile Tail Machines: Adsorption and DNA Delivery by Podoviruses. In *Viral Molecular Machines*. Springer US, pp. 143–179

Chambers, J.R., and Gong, J. (2011) The intestinal microbiota and its modulation for Salmonella control in chickens. *Food Res Int* **44**: 3149–3159.

Chapman, P.A., and Ashton, R. (2003) An evaluation of rapid methods for detecting *Escherichia coli* O157 on beef carcasses. *Int J Food Microbiol* **87**: 279–285.

Chen, H.-M., Wang, Y., Su, L.-H., and Chiu, C.-H. (2013) Nontyphoid Salmonella Infection: Microbiology, Clinical Features, and Antimicrobial Therapy. *Pediatr Neonatol* **54**: 147–152

Conley, A.J., Joensuu, J.J., Jevnikar, A.M., Menassa, R., and Brandle, J.E. (2009) Optimization of elastin-like polypeptide fusions for expression and purification of recombinant proteins in plants. *Biotechnol Bioeng* **103**: 562–573

Conley, A.J., Joensuu, J.J., Menassa, R., Brandle, J.E., Stoger, E., Ma, J., *et al.* (2009) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol* **7**: 48

Conley, A.J., Joensuu, J.J., Richman, A., and Menassa, R. (2011) Protein body-inducing fusions for high-level production and purification of recombinant proteins in plants. *Plant Biotechnol J* **9**: 419–433

Cook, A., Odumeru, J., Lee, S., and Pollari, F. (2012) *Campylobacter, Salmonella, Listeria monocytogenes*, Verotoxigenic *Escherichia coli*, and *Escherichia coli* Prevalence, Enumeration, and Subtypes on Retail Chicken Breasts with and without Skin. *J Food Prot* **75**: 34–40.

Cowley, L.A., Beckett, S.J., Chase-Topping, M., Perry, N., Dallman, T.J., Gally, D.L., and Jenkins, C. (2015) Analysis of whole genome sequencing for the *Escherichia coli* O157:H7 typing phages. *BMC Genomics* **16**: 271

Crameri, A., Raillard, S.A., Bermudez, E., and Stemmer, W.P. (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**: 288–291

Crick, F.H.C., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961) General Nature of the Genetic Code for Proteins. *Nature* **192**: 1227–1232

Cygler, M., Rose, D.R., and Bundle, D.R. (1991) Recognition of a cell-surface oligosaccharide of pathogenic Salmonella by an antibody Fab fragment. *Science* **253**: 442–445.

d'Herelle, F. (1917) Sur un microbe invisible antagoniste des bacilles dysentériques. *CR Acad Sci Paris* **165**: 373–375.

Danner, M., Fuchs, A., Miller, S., and Seckler, R. (1993) Folding and assembly of phage P22 tailspike endorhamnosidase lacking the N-terminal, head-binding domain. *Eur J Biochem* **215**: 653–661

Davidson, A.R., Cardarelli, L., Pell, L.G., Radford, D.R., and Maxwell, K.L. (2012) Long Noncontractile Tail Machines of Bacteriophages. In *Viral Molecular Machines*. pp. 115–142

Deering, A.J., Mauer, L.J., and Pruitt, R.E. (2012) Internalization of *E. coli* O157:H7 and Salmonella spp. in plants: A review. *Food Res Int* **45**: 567–575.

Doore, S.M., and Fane, B.A. (2016) The microviridae: Diversity, assembly, and experimental evolution. *Virology* **491**: 45–55.

Drexler, K., Dannull, J., Hindennach, I., Mutschler, B., and Henning, U. (1991) Single mutations in a gene for a tail fiber component of an *Escherichia coli* phage can cause an extension from a protein to a carbohydrate as a receptor. *J Mol Biol* **219**: 655–663.

Drexler, K., Riede, I., Montag, D., Eschbach, M.-L., and Henning, U. (1989) Receptor specificity of the *Escherichia coli* T-even type phage Ox2: Mutational alterations in host range mutants. *J Mol Biol* **207**: 797–803.

Drulis-Kawa, Z., Majkowska-Skrobek, G., Maciejewska, B., Delattre, A.-S., and Lavigne, R. (2012) Learning from Bacteriophages - Advantages and Limitations of Phage and Phage-Encoded Protein Applications. *Curr Protein Pept Sci* **13**: 699–722.

Duckworth, D.H. (1976) "Who discovered bacteriophage?". *Bacteriol Rev* **40**: 793–802

El-Hawiet, A., Kitova, E.N., Arutyunov, D., Simpson, D.J., Szymanski, C.M., and Klassen, J.S. (2012) Quantifying ligand binding to large protein complexes using electrospray ionization mass spectrometry. *Anal Chem* **84**: 3867–3870

Espejo, R.T., and Canelo, E.S. (1968) Properties of bacteriophage PM2: A lipid-containing bacterial virus. *Virology* **34**: 738–747

Fabre, L., Zhang, J., Guigon, G., Hello, S. Le, Guibert, V., Accou-Demartin, M., *et al.* (2012) CRISPR Typing and Subtyping for Improved Laboratory Surveillance of Salmonella Infections. *PLoS One* **7**: e36995

Fabrega, A., and Vila, J. (2013) *Salmonella enterica* Serovar Typhimurium Skills To Succeed in the Host: Virulence and Regulation. *Clin Microbiol Rev* **26**: 308–341

Fanelli, M.J., Sadler, W.W., Franti, C.E., and Brownell, J.R. (1971) Localization of Salmonellae within the Intestinal Tract of Chickens. *Avian Dis* **15**: 366

Ferris, M.T., Joyce, P., and Burch, C.L. (2007) High frequency of mutations that expand the host range of an RNA virus. *Genetics* **176**: 1013–1022.

Fischer, R., Schillberg, S., Hellwig, S., Twyman, R.M., and Drossard, J. (2012) GMP issues for recombinant plant-derived pharmaceutical proteins. *Biotechnol Adv* **30**: 434–439.

Forterre, P., Filée, J., and Myllykallio, H. (2013) *Origin and Evolution of DNA and DNA Replication Machineries*. Landes Bioscience, Austin, Texas.

Freiberg, A., Morona, R., Bosch, L. Van den, Jung, C., Behlke, J., Carlin, N., *et al.* (2003) The tailspike protein of Shigella phage Sf6. A structural homolog of Salmonella phage P22 tailspike protein without sequence similarity in the beta-helix domain. *J Biol Chem* **278**: 1542–1548.

Garcia-Doval, C., and Raaij, M.J. van (2012) Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Proc Natl Acad Sci* **109**: 9390–9395

Garcia-Doval, C., and Raaij, M.J. van (2013) Bacteriophage receptor recognition and nucleic Acid transfer. *Subcell Biochem* **68**: 489–518.

Garneau, J.E., Dupuis, M.-V., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., *et al.* (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**: 67–71.

Ge, P., Scholl, D., Leiman, P.G., Yu, X., Miller, J.F., and Zhou, Z.H. (2015) Atomic structures of a bactericidal contractile nanotube in its pre- and postcontraction states. *Nat Struct Mol Biol* **22**: 377–382

Gebhart, D., Lok, S., Clare, S., Tomas, M., Stares, M., Scholl, D., *et al.* (2015) A modified R-type bacteriocin specifically targeting *Clostridium difficile* prevents colonization of mice without affecting gut microbiota diversity. *MBio* **6**: 10.1128/mBio.02368-14.

Gemski, P., and Stocker, B.A. (1967) Transduction by bacteriophage P22 in nonsmooth mutants of *Salmonella* typhimurium. *J Bacteriol* **93**: 1588–1597.

George, D.G., Yeh, L.-S.L., and Barker, W.C. (1983) Unexpected relationships between bacteriophage lambda hypothetical proteins and bacteriophage T4 tail-fiber proteins. *Biochem Biophys Res Commun* **115**: 1061–1068

Goldenberg, D., and King, J. (1982) Trimeric intermediate in the in vivo folding and subunit assembly of the tail spike endorhamnosidase of bacteriophage P22. *Proc Natl Acad Sci U S A* **79**: 3403–7

Goldenberg, D.P., Berget, P.B., and King, J. (1982) Maturation of the tail spike endorhamnosidase of Salmonella phage P22. *J Biol Chem* **257**: 7864–7871.

Goode, D., Allen, V.M., and Barrow, P.A. (2003) Reduction of Experimental Salmonella and Campylobacter Contamination of Chicken Skin by Application of Lytic Bacteriophages. *Appl Environ Microbiol* **69**: 5032–5036

Guarner, F., and Malagelada, J.-R. (2003) Gut flora in health and disease. *Lancet* **361**: 512–519.

Habann, M., Leiman, P.G., Vandersteegen, K., Bossche, A. Van den, Lavigne, R., Shneider, M.M., *et al.* (2014) Listeria phage A511, a model for the contractile tail machineries of SPO1-related bacteriophages. *Mol Microbiol* **92**: 84–99.

Hammack, T. (2011) *Salmonella species in Bad Bug Book--Foodborne Pathogenic Microorganisms and Natural Toxins*. US Food and Drug Administration, College Park, MD, .

Hashemolhosseini, S., Stierhof, Y.-D., Hindennach, I., and Henning, U. (1996) Characterization of the Helper Proteins for the Assembly of Tail Fibers of Coliphages T4 and lambda. *J Bacteriol* **178**: 6258–6265.

Hassouneh, W., MacEwan, S.R., and Chilkoti, A. (2012) Chapter nine – Fusions of Elastin-Like Polypeptides to Pharmaceutical Proteins. In *Methods in Enzymology*. pp. 215–237.

Hatfull, G.F., and Hendrix, R.W. (2011) Bacteriophages and their genomes. *Curr Opin Virol* **1**: 298–303.

Hellwig, S., Drossard, J., Twyman, R.M., and Fischer, R. (2004) Plant cell cultures for the production of recombinant proteins. *Nat Biotechnol* **22**: 1415–1422

Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E., and Holden, D.W. (1995)

Simultaneous identification of bacterial virulence genes by negative selection. *Science* **269**: 400–3

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

Herskowitz, I., and Hagen, D. (1980) The Lysis-Lysongeny Decision of Phage lambda: Explict Programming and Responsiveness. *Ann Rev Genet* **14**: 399–445.

Hiatt, A., Caffferkey, R., and Bowdish, K. (1989) Production of antibodies in transgenic plants. *Nature* **342**: 76–78

Hong, J., Ye, X., Wang, Y., and Zhang, Y.-H.P. (2008) Bioseparation of recombinant cellulose-binding module-proteins by affinity adsorption on an ultra-high-capacity cellulosic adsorbent. *Anal Chim Acta* **621**: 193–199.

Hyunjong, B., Lee, D.-S., and Hwang, I. (2005) Dual targeting of xylanase to chloroplasts and peroxisomes as a means to increase protein accumulation in plant cells. *J Exp Bot* **57**: 161–169

Ignacio-Espinoza, C.J., Solonenko, S.A., Sullivan, M.B., Ignacio-Espinoza, J.C., Solonenko, S.A., and Sullivan, M.B. (2013) The global virome: Not as big as we thought? *Curr Opin Virol* **3**: 566–571

Israel, V. (1977) E proteins of bacteriophage P22. I. Identification and ejection from wild-type and defective particles. *J Virol* **23**: 91–7

Iwashita, S., and Kanegasaki, S. (1976) Enzymic and molecular properties of base-plate parts of bacteriophage P22. *Eur J Biochem* **65**: 87–94.

Javed, M.A., Ackermann, H.W., Azeredo, J., Carvalho, C.M., Connerton, I., Evoy, S., *et al.* (2014) A suggested classification for two groups of Campylobacter myoviruses. *Arch Virol* **159**: 181–190.

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**.

Joensuu, J.J., Conley, A.J., Lienemann, M., Brandle, J.E., Linder, M.B., and Menassa, R. (2010) Hydrophobin fusions for high-level transient protein expression and purification in *Nicotiana benthamiana*. *Plant Physiol* **152**: 622–33

Kaldis, A., Ahmad, A., Reid, A., McGarvey, B., Brandle, J., Ma, S., *et al.* (2013) High-level production of human interleukin-10 fusions in tobacco cell suspension cultures. *Plant Biotechnol J* **11**: 535–545

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845–858

Kiem, S., Oh, W.S., Peck, K.R., Lee, N.Y., Lee, J.Y., Song, J.H., *et al.* (2004) Phase variation of biofilm formation in *Staphylococcus aureus* by IS 256 insertion and its impact on the capacity adhering to polyurethane surface. *J Korean Med Sci* **19**: 779–782.

Kolotilin, I., Topp, E., Cox, E., Devriendt, B., Conrad, U., Joensuu, J., *et al.* (2014) Plant-based

solutions for veterinary immunotherapeutics and prophylactics. *Vet Res* **45**: 117

Kopertekh, L., Jüttner, G., and Schiemann, J. (2004) Site-specific recombination induced in transgenic plants by PVX virus vector expressing bacteriophage P1 recombinase. *Plant Sci* **166**: 485–492.

Kozak, G.K., MacDonald, D., Landry, L., and Farber, J.M. (2013) Foodborne Outbreaks in Canada Linked to Produce: 2001 through 2009. *J Food Prot* **76**: 173–183

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

Kropinski, A.M., Arutyunov, D., Foss, M., Cunningham, A., Ding, W., Singh, A., *et al.* (2011) Genome and proteome of Campylobacter jejuni bacteriophage NCTC 12673. *Appl Environ Microbiol* **77**: 8265–8271.

Kutter, E., and Sulakvelidze, A. (2005) *Bacteriophages : biology and applications*. CRC Press, Boca Raton.

Kutter, E.M., Kuhl, S.J., and Abedon, S.T. (2015) Re-establishing a place for phage therapy in western medicine. **10**: 685–688 http://www.futuremedicine.com/doi/10.2217/fmb.15.28. Accessed June 21, 2016.

Kwon, K.-C., and Daniell, H. (2015) Low-cost oral delivery of protein drugs bioencapsulated in plant cells. *Plant Biotechnol J* **13**: 1017–1022

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

Laemmli, U.K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nat Publ Gr* **227**: 680–685.

Lander, G.C., Khayat, R., Li, R., Prevelige, P.E., Potter, C.S., Carragher, B., and Johnson, J.E. (2009) The P22 Tail Machine at Subnanometer Resolution Reveals the Architecture of an Infection Conduit. *Structure* **17**: 789–799

Lang, A.S., Zhaxybayeva, O., and Beatty, J.T. (2012) Gene transfer agents: phage-like elements of genetic exchange. *Nat Rev* **10**: 472–482

Leiman, P.G., and Shneider, M.M. (2012) Contractile Tail Machines of Bacteriophages. Springer US, pp. 93–114

Leonard, P., Hearty, S., Quinn, J., and O'Kennedy, R. (2004) A generic approach for the detection of whole *Listeria monocytogenes* cells in contaminated samples using surface plasmon resonance. .

Levine, M. (1957) Mutations in the temperate phage P22 and lysogeny in Salmonella. *Virology* **3**: 22–41

Lim, J.Y., Yoon, J., and Hovde, C.J. (2010) A brief overview of *Escherichia coli* O157:H7 and its plasmid O157. *J Microbiol Biotechnol* **20**: 5–14

Lindberg, A.A. (1973) Bacteriophage receptors. *Annu Rev Microbiol* **27**: 205–241.

Liu, M., Deora, R., Doulatov, S.R., Gingery, M., Eiserling, F.A., Preston, A., *et al.* (2002) Reverse Transcriptase-Mediated Tropism Switching in Bordetella Bacteriophage. *Science (80-)* **295**: 2091–2094

Lobocka, M.B., Rose, D.J., Plunkett, G., Rusin, M., Samojedny, A., Lehnherr, H., *et al.* (2004) Genome of Bacteriophage P1. *J Bacteriol* **186**: 7032–7068

Loessner, M.J., Kramer, K., Ebel, F., and Scherer, S. (2002) C-terminal domains of *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and high-affinity binding to bacterial cell wall carbohydrates. *Mol Microbiol* **44**: 335–349

Lwoff, A., and Gutmann, A. (1950) Recherches sur un Bacillus-Megatherium Lysogene. *Ann L Inst PASTEUR* **78**: 711–739.

MacDonald, J., Doshi, K., Dussault, M., Hall, J.C., Holbrook, L., Jones, G., *et al.* (2015) Bringing plant-based veterinary vaccines to market: Managing regulatory and commercial hurdles. *Biotechnol Adv* **33**: 1572–1581.

Mandal, P., Biswas, A., Choi, K., and Pal, U. (2011) Methods for Rapid Detection of Foodborne Pathogens: An Overview. *Am J Food Technol* **6**: 87–102

Maniloff, J., Kampo, G.J., and Dascher, C.C. (1994) Sequence analysis of a unique temperate phage: mycoplasma virus L2. *Gene* **141**: 1–8.

Marcó, M.B., Moineau, S., and Quiberoni, A. (2012) Bacteriophages and dairy fermentations. *Bacteriophage* **2**: 149–158

Marcus, S.L., Brumell, J.H., Pfeifer, C.G., and Finlay, B.B. (2000) Salmonella pathogenicity islands: big virulence in small packages. *Microbes Infect* **2**: 145–156.

Marvin, D. (1998) Filamentous phage structure, infection and assembly. *Curr Opin Struct Biol* **8**: 150–158

Mead, P. (1999) Food-Related Illness and Death in the United States. *Emerg Infect Dis* **5**: 607–625

Meyer, J.R., Dobias, D.T., Weitz, J.S., Barrick, J.E., Quick, R.T., and Lenski, R.E. (2012) Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* **335**: 428–432

Michel, C.J., Jacq, B., Arquès, D.G., and Bickle, T.A. (1986) A remarkable amino acid sequence homology between a phage T4 tail fibre protein and ORF314 of phage λ located in the tail operon. .

Michel-Briand, Y., and Baysse, C. (2002) The pyocins of Pseudomonas aeruginosa. *Biochimie* **84**: 499–510.

Miletic, S., Simpson, D.J., Szymanski, C.M., Deyholos, M.K., and Menassa, R. (2016) A Plant-Produced Bacteriophage Tailspike Protein for the Control of Salmonella. *Front Plant Sci* **6**: 1221.

Miller, S., Schuler, B., and Seckler, R. (1998) Phage P22 tailspike protein: removal of head-binding domain unmasks effects of folding mutations on native-state thermal stability. *Protein Sci* **7**: 2223–2232.

Miller, W.G., Bates, A.H., Horn, S.T., Brandl, M.T., Wachtel, M.R., and Mandrell, R.E. (2000) Detection on surfaces and in Caco-2 cells of *Campylobacter jejuni* cells transformed with new gfp, yfp, and cfp marker plasmids. *Appl Environ Microbiol* **66**: 5426–5436.

Montag, D., Schwarz, H., and Henning, U. (1989) A component of the side tail fiber of *Escherichia coli* bacteriophage lambda can functionally replace the receptor-recognizing part of a long tail fiber protein of the unrelated bacteriophage T4. *J Bacteriol* **171**: 4378–84

Morita, M., Tanji, Y., Orito, Y., Mizoguchi, K., Soejima, A., and Unno, H. (2001) Functional analysis of antibacterial activity of *Bacillus amyloliquefaciens* phage endolysin against Gram-negative bacteria. *FEBS Lett* **500**: 56–59

Mutoh, N., Furukawa, H., and Mizushima, S. (1978) Role of lipopolysaccharide and outer membrane protein of *Escherichia coli* K-12 in the receptor activity for bacteriophage T4. *J Bacteriol* **136**: 693–699.

Naidoo, R., Singh, A., Arya, S.K., Beadle, B., Glass, N., Tanha, J., *et al.* (2012) Surface-immobilization of chromatographically purified bacteriophages for the optimized capture of bacteria. *Bacteriophage* **2**: 15–24

Nguyen, H.T., Leelavathi, S., and Reddy, V.S. (2004) Bacteriophage T7 RNA polymerase-directed, inducible and tissue-specific over-expression of foreign genes in transgenic plants. *Plant Biotechnol J* **2**: 301–310

Nomura, M., and Benzer, S. (1961) The nature of the "deletion" mutants in the rII region of phage T4. *J Mol Biol* **3**: 684–692

Notenboom, V., Boraston, A.B., Kilburn, D.G., and Rose, D.R. (2001) Crystal Structures of the Family 9 Carbohydrate-Binding Module from *Thermotoga maritima* Xylanase 10A in Native and Ligand-Bound Forms. *Biochemistry* **40**: 6248–6256

Novotny, M. V, Alley, W.R., and Mann, B.F. (2013) Analytical glycobiology at high sensitivity: current approaches and directions. *Glycoconj J* **30**: 89–117

Nykiforuk, C.L., Boothe, J.G., Murray, E.W., Keon, R.G., Goren, H.J., Markley, N.A., and Moloney, M.M. (2006) Transgenic expression and recovery of biologically active recombinant human insulin from *Arabidopsis thaliana* seeds. *Plant Biotechnol J* **4**: 77–85

Ochman, H., Soncini, F.C., Solomon, F., and Groisman, E.A. (1996) Identification of a pathogenicity island required for Salmonella survival in host cells. *Proc Natl Acad Sci U S A* **93**: 7800–4

Oey, M., Lohse, M., Kreikemeyer, B., and Bock, R. (2009) Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. *Plant J* **57**: 436–445

Olia, A.S., Casjens, S., and Cingolani, G. (2007) Structure of phage P22 cell envelope–penetrating needle. *Nat Struct Mol Biol* **14**: 1221–1226

Oliveira, C., Carvalho, V., Domingues, L., and Gama, F.M. (2015) Recombinant CBM-fusion technology — Applications overview. *Biotechnol Adv* **33**: 358–369.

Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L., and Adhya, S. (2005) Switches in Bacteriophage Lambda Development. *Annu Rev Genet*. **39**:409-29.

Painter, J.A., Hoekstra, R.M., Ayers, T., Tauxe, R. V., Braden, C.R., Angulo, F.J., and Griffin, P.M. (2013) Attribution of Foodborne Illnesses, Hospitalizations, and Deaths to Food Commodities by using Outbreak Data, United States, 1998–2008. *Emerg Infect Dis* **19**: 407–415

Patel, J., Zhu, H., Menassa, R., Gyenis, L., Richman, A., and Brandle, J. (2007) Elastin-like polypeptide fusions enhance the accumulation of recombinant proteins in tobacco leaves. *Transgenic Res* **16**: 239–249

Pillay, P., Schlüter, U., Wyk, S. van, Kunert, K.J., and Vorster, B.J. (2014) Proteolysis of recombinant proteins in bioengineered plant cells. *Bioengineered* **5**: 15–20

Poranen, M.M., Bamford, D.H., Oksanen, H.M., Poranen, M.M., Bamford, D.H., and Oksanen, H.M. (2015) Membrane-Containing Bacteriophages. In *eLS*. John Wiley & Sons, Ltd, Chichester, UK. pp. 1–11

Poshtiban, S., Javed, M.A., Arutyunov, D., Singh, A., Banting, G., Szymanski, C.M., and Evoy, S. (2013) Phage receptor binding protein-based magnetic enrichment method as an aid for real time PCR detection of foodborne bacteria. *Analyst* **138**: 5619–5626

Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D., and Mekalanos, J.J. (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc Natl Acad Sci U S A* **104**: 15508–15513.

Raaij, M.J. van, Schoehn, G., Burda, M.R., and Miller, S. (2001) Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. *J Mol Biol* **314**: 1137–1146

Raaij, M.J. van, Schoehn, G., Jaquinod, M., Ashman, K., Burda, M.R., and Miller, S. (2001) Identification and Crystallisation of a Heat- and Protease-Stable Fragment of the Bacteriophage T4 Short Tail Fibre. *Biol Chem* **382**: 1049–1055

Revolledo, L., Ferreira, A.J.P., and Mead, G.C. (2006) Prospects in Salmonella Control: Competitive Exclusion, Probiotics, and Enhancement of Avian Intestinal Immunity. *J Appl Poult Res* **15**: 341–351

Ritchie, J.M., Greenwich, J.L., Davis, B.M., Bronson, R.T., Gebhart, D., Williams, S.R., *et al.* (2011) An Escherichia coli O157-Specific Engineered Pyocin Prevents and Ameliorates Infection by *E. coli* O157:H7 in an Animal Model of Diarrheal Disease. *Antimicrob Agents Chemother* **55**: 5469–5474

Roach, D.R., and Donovan, D.M. (2015) Antimicrobial bacteriophage-derived proteins and therapeutic applications. *Bacteriophage* **5**: e1062590

Robinson, A.S., and King, J. (1997) Disulphide-bonded intermediate on the folding and assembly pathway of a non-disulphide bonded protein. *Nat Struct Biol* **4**: 450–455

Rodrigue, D.C., Tauxe, R. V., and Rowe, B. (1990) International increase in *Salmonella*

enteritidis: A new pandemic? *Epidemiol Infect* **105**: 21

Rohwer, F. (2003) Global phage diversity. *Cell* **113**: 141.

Saberianfar, R., Joensuu, J.J., Conley, A.J., and Menassa, R. (2015) Protein body formation in leaves of *Nicotiana benthamiana* : a concentration-dependent mechanism influenced by the presence of fusion tags. *Plant Biotechnol J* **13**: 927–937

Salomon, D., and Orth, K. (2015) Type VI secretion system. *Curr Biol* **25**: R265–R266.

Sana, T.G., Berni, B., and Bleves, S. (2016) The T6SSs of Pseudomonas aeruginosa Strain PAO1 and Their Effectors: Beyond Bacterial-Cell Targeting. *Front Cell Infect Microbiol* **6**: 61

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., *et al.* (1977) Nucleotide sequence of bacteriophage φX174 DNA. *Nature* **265**: 687–695

Sanger, F., and Coulson, A.R. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* **94**: 441–448

Santoso, Y., Joyce, C.M., Potapova, O., Reste, L. Le, Hohlbein, J., Torella, J.P., *et al.* (2010) Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc Natl Acad Sci U S A* **107**: 715–720.

Saxena, S.K., O'Brien, A.D., and Ackerman, E.J. (1989) Shiga toxin, Shiga-like toxin II variant, and ricin are all single-site RNA N-glycosidases of 28 S RNA when microinjected into Xenopus oocytes. *J Biol Chem* **264**: 596–601

Schmelcher, M., Donovan, D.M., and Loessner, M.J. (2012) Bacteriophage endolysins as novel antimicrobials. *Future Microbiol* **7**: 1147–71

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. *Appl Environ Microbiol* **76**: 5745–5756

Schmitz, J.E., Daniel, A., Collin, M., Schuch, R., and Fischetti, V.A. (2008) Rapid DNA library construction for functional genomic and metagenomic screening. *Appl Environ Microbiol* **74**: 1649–1652.

Scholl, D., Cooley, M., Williams, S.R., Gebhart, D., Martin, D., Bates, A., and Mandrell, R. (2009) An Engineered R-Type Pyocin Is a Highly Specific and Sensitive Bactericidal Agent for the Food-Borne Pathogen *Escherichia coli* O157:H7. *Antimicrob Agents Chemother* **53**: 3074–3080

Schulz, E.C., Schwarzer, D., Frank, M., Stummeyer, K., Mühlenhoff, M., Dickmanns, A., *et al.* (2010) Structural Basis for the Recognition and Cleavage of Polysialic Acid by the Bacteriophage K1F Tailspike Protein EndoNF. *J Mol Biol* **397**: 341–351.

Schwarzer, D., Browning, C., Stummeyer, K., Oberbeck, A., Mühlenhoff, M., Gerardy-Schahn, R., and Leiman, P.G. (2015) Structure and biochemical characterization of bacteriophage phi92 endosialidase. *Virology* **477**: 133–143

Schwarzer, D., Buettner, F.F.R., Browning, C., Nazarov, S., Rabsch, W., Bethe, A., *et al.* (2012)

A Multivalent Adsorption Apparatus Explains the Broad Host Range of Phage phi92: a Comprehensive Genomic and Structural Analysis. *J Virol* **86**: 10384–10398

Seckler, R., Fuchs, A., King, J., and Jaenicke, R. (1989) Reconstitution of the Thermostable Trimeric. **264**: 11750–11753.

Seul, A., Müller, J.J., Andres, D., Stettner, E., Heinemann, U., and Seckler, R. (2014) Bacteriophage P22 tailspike: structure of the complete protein and function of the interdomain linker. *Acta Crystallogr Sect D Biol Crystallogr* **70**: 1336–1345

Shamji, M.F., Betre, H., Kraus, V.B., Chen, J., Chilkoti, A., Pichika, R., *et al.* (2007) Development and characterization of a fusion protein between thermally responsive elastin-like polypeptide and interleukin-1 receptor antagonist: Sustained release of a local antiinflammatory therapeutic. *Arthritis Rheum* **56**: 3650–3661

Sigurskjold, B.W., Altman, E., and Bundle, D.R. (1991) Sensitive titration microcalorimetric study of the binding of Salmonella O-antigenic oligosaccharides by a monoclonal antibody. *Eur J Biochem* **197**: 239–246

Silhavy, D., Molnár, A., Lucioli, A., Szittya, G., Hornyik, C., Tavazza, M., and Burgyán, J. (2002) A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs. *EMBO J* **21**: 3070–3080

Simpson, D., Sacher, J., and Szymanski, C. (2016) Development of an Assay for the Identification of Receptor Binding Proteins from Bacteriophages. *Viruses* **8**: 17

Simpson, D.J., Sacher, J.C., and Szymanski, C.M. (2015) Exploring the interactions between bacteriophage-encoded glycan binding proteins and carbohydrates. *Curr Opin Struct Biol* **34**: 69–77

Singh, A., Arutyunov, D., McDermott, M.T., Szymanski, C.M., and Evoy, S. (2011) Specific detection of *Campylobacter jejuni* using the bacteriophage NCTC 12673 receptor binding protein as a probe. *Analyst* **136**: 4780–4786.

Singh, A., Arutyunov, D., Szymanski, C.M., and Evoy, S. (2012) Bacteriophage based probes for pathogen detection. *Analyst* **137**: 3405–3421.

Singh, A., Glass, N., Tolba, M., Brovko, L., Griffiths, M., and Evoy, S. (2009) Immobilization of bacteriophages on gold surfaces for the specific capture of pathogens. *Biosens Bioelectron* **24**: 3645–3651.

Singh, A., Poshtiban, S., and Evoy, S. (2013) Recent Advances in Bacteriophage Based Biosensors for Food-Borne Pathogen Detection. *Sensors* **13**: 1763–1786

Skurnik, M., and Strauch, E. (2006) Phage therapy: Facts and fiction. *Int J Med Microbiol* **296**: 5–14.

Smith, N.L., Taylor, E.J., Lindsay, A.-M., Charnock, S.J., Turkenburg, J.P., Dodson, E.J., *et al.* (2005) Structure of a group A streptococcal phage-encoded virulence factor reveals a catalytically active triple-stranded -helix. *Proc Natl Acad Sci* **102**: 17652–17657

Sorek, R., Lawrence, C.M., and Wiedenheft, B. (2013) CRISPR-Mediated Adaptive Immune

Systems in Bacteria and Archaea. *Annu Rev Biochem* **82**: 237–266

Sørensen, M.C.H., Alphen, L.B. van, Fodor, C., Crowley, S.M., Christensen, B.B., Szymanski, C.M., *et al.* (2012) Phase Variable Expression of Capsular Polysaccharide Modifications Allows *Campylobacter jejuni* to Avoid Bacteriophage Infection in Chickens. *Front Cell Infect Microbiol* **2**: 11

Steinbacher, S., Baxa, U., Miller, S., Weintraub, A., Seckler, R., and Huber, R. (1996) Crystal structure of phage P22 tailspike protein complexed with Salmonella sp. O-antigen receptors. *Proc Natl Acad Sci U S A* **93**: 10584–10588

Steinbacher, S., Miller, S., Baxa, U., Budisa, N., Weintraub, A., Seckler, R., and Huber, R. (1997) Phage P22 tailspike protein: crystal structure of the head-binding domain at 2.3 A, fully refined structure of the endorhamnosidase at 1.56 A resolution, and the molecular basis of O-antigen recognition and cleavage. *J Mol Biol* **267**: 865–80

Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. *Science* **265**: 383–386.

Stemmer, W.P.C. (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**: 389–391

Stoger, E., Fischer, R., Moloney, M., and Ma, J.K.-C. (2014) Plant Molecular Pharming for the Treatment of Chronic and Infectious Diseases. *Annu Rev Plant Biol* **65**: 743–768

Streatfield, S.J. (2007) Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnol J* **5**: 2–15

Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R., and Ficner, R. (2005) Crystal structure of the polysialic acid–degrading endosialidase of bacteriophage K1F. *Nat Struct Mol Biol* **12**: 90–96

Su, L.-H., Chiu, C.-H., Chu, C., and Ou, J.T. (2004) Antimicrobial Resistance in Nontyphoid Salmonella Serotypes: A Global Challenge. *Clin Infect Dis* **39**: 546–551

Sulakvelidze, A., Alavidze, Z., and Morris, J.G. (2001) MINIREVIEW Bacteriophage Therapy. **45**: 649–659.

Sulakvelidze, A., Alavidze, Z., and Morris, J.G. (2001) Bacteriophage Therapy. *Antimicrob Agents Chemother* **45**: 649–659

Summers, W.C. (1999) *Félix d'Herelle and the origins of molecular biology*. Yale University Press, .

Suttle, C. a (2005) Viruses in the sea. *Nature* **437**: 356–361.

Székely, A.J., Breitbart, M., Ackermann, H.-W., Adams, M., Lefkowitz, E., King, A., *et al.* (2016) Single-stranded DNA phages: from early molecular biology tools to recent revolutions in environmental microbiology. *FEMS Microbiol Lett* **363**: 1–7

Tang, J., Lander, G.C., Olia, A., Li, R., Casjens, S., Prevelige, P., *et al.* (2011) Peering Down the Barrel of a Bacteriophage Portal: The Genome Packaging and Release Valve in P22. *Structure*

19: 496–502.

Tay, L.-L.L., Huang, P.-J.J., Tanha, J., Ryan, S., Wu, X., Hulse, J., and Chau, L.-K.K. (2012) Silica encapsulated SERS nanoprobe conjugated to the bacteriophage tailspike protein for targeted detection of Salmonella. *Chem Commun (Camb)* **48**: 1024–6

Terpe, K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol* **60**: 523–533

Thomassen, E., Gielen, G., Schütz, M., Schoehn, G., Abrahams, J.P., Miller, S., and Raaij, M.J. van (2003) The Structure of the Receptor-binding Domain of the Bacteriophage T4 Short Tail Fibre Reveals a Knitted Trimeric Metal-binding Fold. *J Mol Biol* **331**: 361–373.

Tolba, M., Ahmed, M.U., Tlili, C., Eichenseher, F., Loessner, M.J., Zourob, M., *et al.* (2012) A bacteriophage endolysin-based electrochemical impedance biosensor for the rapid detection of Listeria cells. *Analyst* **137**: 5749

Tsai, C.-M., and Frasch, C.E. (1982) A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. *Anal Biochem* **119**: 115–119

Turner, N.J. (2009) Directed evolution drives the next generation of biocatalysts. *Nat Chem Biol* **5**: 567–573.

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

Unterweger, D., Kitaoka, M., Miyata, S.T., Bachmann, V., Brooks, T.M., Moloney, J., *et al.* (2012) Constitutive Type VI Secretion System Expression Gives *Vibrio cholerae* Intra- and Interspecific Competitive Advantages. *PLoS One* **7**: e48320

Waddell, T., Head, S., Petric, M., Cohen, A., and Lingwood, C. (1988) Globotriosyl ceramide is specifically recognized by the escherichia coli verocytotoxin 2. *Biochem Biophys Res Commun* **152**: 674–679

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–Ba. *Foodborne Pathog Dis* **7**: 1019–1024

Walter, M., Fiedler, C., Grassl, R., Biebl, M., Rachel, R., Hermo-Parrado, X.L., *et al.* (2008) Structure of the Receptor-Binding Protein of Bacteriophage Det7: a Podoviral Tail Spike in a Myovirus. *J Virol* **82**: 2265–2273

Warren, R.A. (1980) Modified bases in bacteriophage DNAs. *Annu Rev Microbiol* **34**: 137–158.

Waseh, S., Hanifi-Moghaddam, P., Coleman, R., Masotti, M., Ryan, S., Foss, M., *et al.* (2010) Orally Administered P22 Phage Tailspike Protein Reduces Salmonella Colonization in Chickens: Prospects of a Novel Therapy against Bacterial Infections. *PLoS One* **5**: e13904.

Weigele, P.R., Haase-Pettingell, C., Campbell, P.G., Gossard, D.C., and King, J. (2005) Stalled Folding Mutants in the Triple β-Helix Domain of the Phage P22 Tailspike Adhesin. *J Mol Biol*

354: 1103–1117.

Williams, S.R., Gebhart, D., Martin, D.W., and Scholl, D. (2008) Retargeting R-type pyocins to generate novel bactericidal protein complexes. *Appl Environ Microbiol* **74**: 3868–3876.

Wisniewski-Dye, F., and Vial, L. (2008) Phase and antigenic variation mediated by genome modifications. *Antonie Van Leeuwenhoek* **94**: 493–515.

Wolffs, P.F.G., Glencross, K., Thibaudeau, R., and Griffiths, M.W. (2006) Direct Quantitation and Detection of Salmonellae in Biological Samples without Enrichment, Using Two-Step Filtration and Real-Time PCR. *Appl Environ Microbiol* **72**: 3896–3900

Zaleski, P., Wojciechowski, M., and Piekarowicz, A. (2005) The role of Dam methylation in phase variation of *Haemophilus influenzae* genes involved in defence against phage infection. *Microbiology* **151**: 3361–3369

Zinder, N.D., and Lederberg, J. (1952) Genetic exchange in Salmonella. *J Bacteriol* **64**: 679–699