Identification and characterisation of lysin enzymes as potential therapeutics for the treatment of *Clostridium difficile*

ABDULLAH ALYOUSEF

A thesis submitted to Cardiff University for the degree of Doctor of Philosophy

March 2013

Cardiff University

Cardiff School of Pharmacy & Pharmaceutical Sciences

Redwood Building
King Edward VII Avenue
Cardiff CF10 3NB

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed  (candidate) Date 22-06-2013

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed  (candidate) Date 22-06-2013

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed  (candidate) Date 22-06-2013

STATEMENT 3 : PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed  (candidate) Date 22-06-2013

Acknowledgements

Firstly, I would like to thank my supervisors Professor L. W. J. Baillie and Dr. Jean-Yves Maillard for their help, support and guidance throughout my PhD studies. You encouraged me to do my best and your advice on article writing is unforgettable.
Secondly, I would like to thank Dr. Jon Brazier and Dr. Val Hall for their kind donation of strains from the national anaerobic reference laboratory in Cardiff. I am thankful to Dr. Andrea Brancale for his help with protein modeling. Huge thanks to my friends and colleagues at Cardiff University, who helped me along my way and made it an excellent place to work.

I would like to thank King Saud University (Riyadh) and Royal Embassy of Saudi Arabia (London) for their help and support during my study.

On a personal note, I want to thank all my family, especially my father and mother, who have helped and encouraged me during my whole academic life, and I would also like to express great thanks to my wife ANFAL and my daughter LAYAL, whose patient love enabled me to complete this work .

Abdullah Alyousef

Publications and Presentations relating to these studies

Publication

Joshi, L. T., Phillips, D.S., Williams, C.F., Alyousef, A., and Baillie, L. 2012. Contribution of spores to the ability of *Clostridium difficile* to adhere to surfaces. *Applied and Environmental Microbiology* 78(21), pp. 7671-7679.

Patent

Regarding characterization of bacteriophage endolysins as potential therapeutics for the treatment of *Clostridium difficile*, as a part of postgraduate study at Cardiff University. (2013) (pending).

Presentations

Isolation and characterization of lytic phage for treatment of *Clostridium difficile*. Welsh School of Pharmacy, Postgraduate Day, Cardiff, 2011. **Poster presentation**

Evaluation of bacteriophages and their products as potential therapeutics for the treatment of *Clostridium difficile*. Welsh School of Pharmacy, Postgraduate Day, Cardiff, 2012. **Oral presentation**

Summary

Clostridium difficile is the most common cause of hospital acquired infections. While the current treatments of choice, antibiotics, are generally effective in promoting recovery ,the increased incidence of *C. difficile* infections and treatment failure associated with antibiotic resistance combined with the emergence of hypervirulent strains highlights the need to develop therapeutic approaches that specifically target the pathogen without causing collateral damage to the protective microbiota. Several non-antibiotic approaches are currently being investigated, such as bacteriophage therapy. For this reason, we attempted to isolate *C. difficile* specific lytic bacteriophages which could form the basis of a treatment for *C. difficile*. While we were unable to isolate lytic phages, we were able to isolate twelve temperate bacteriophages from twenty-three clinical isolates of *C. difficile* using mitomycin C. Unfortunately we failed to identify a susceptible host strain capable of supporting the replication of these phages. This failure may in part be due to repeated episodes of phage infection over time, which have resulted in the emergence of "phage resistant" species mediated by systems such as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR). Employing a PCR-based approach using primers specific for the lysin genes of five previously isolated *C. difficile* phages, we found evidence to suggest that repeated bacteriophage infection is a common event for clinical isolates of *C. difficile.* Our inability to isolate a lytic bacteriophage prompted us to adopt an alternative approach in which we used endolysin enzyme of five previously identified *C. difficile* phages as recombinant protein. These lysins showed broad spectrum activity against the vegetative forms of a large collection of *C. difficile* ribotypes with little or no activity against other species, supporting their potential as therapeutic agents. We also identified a genome associated lysin (CD630, YP_001088405), which lysed vegetative *C. difficile* in a similar manner to the phage derived lysins. We also cloned and expressed a spore cortex lytic enzyme (SleC) which targeted the cortex of *C. difficile* spores. Unfortunately this enzyme was inactive against intact spores, suggesting that the outer layers of the spore act as a permeability barrier. The results of this study showed *in vitro* the applicability of endolysins against the vegetative form of *C. difficile* and the activity of spore cortex lytic enzyme against coatless spores, offering interesting perspectives for evaluation of the antibacterial activity of a mixture of endolysin and spore cortex lytic enzyme.

List of contents

List of tables

Chapter One: Introduction

Chapter Two: General materials and methods

Chapter Three: Isolation of bacteriophages infecting *Clostridium difficile*

Chapter Four: The cloning and expression of the lysin Genes of *Clostridium difficile* and its bacteriophages

Chapter Five: The distribution of endolysins amongst clinical isolates of *Clostridium difficile*

Chapter Six: Germination-specific cortex-lytic enzymes of *Clostridium difficile*

List of figures

Chapter One: Introduction

Chapter Two: General materials and methods

Chapter Three: Isolation of bacteriophages infecting *Clostridium difficile*

Chapter Four: The cloning and expression of the lysin Genes of *Clostridium difficile* and its bacteriophages

Chapter Five: The distribution of endolysins amongst clinical isolates of *Clostridium difficile*

Chapter Six: Germination-specific cortex-lytic enzymes of *Clostridium difficile*

Abbreviations

CDAD ----------*Clostridium difficile*-associated diarrhea
CDI ----------*Clostridium difficile* infection
PG---------- - Peptidoglycan
RNA ----------Ribonucleic acid
NAG ----------*N*-acetylglucosamine
NAM---------- *N*-acetylmuramic acid
DNA ----------Deoxyribonucleic acid
PCR----------Polymerase chain reaction
EDTA----------Ethylenediaminetetraacetic acid
E.coli----------*Escherichia coli*
Taq DNA ----------*Thermus aquaticus* DNA plymerase
SDW ----------Serial Distilled Water
PEG----------polyethylene glycol
CRISPR----------clustered regularly interspaced short palindromic repeats
CAS----------CRISPR-associated sequence
ATCC ----------American Type Culture Collection
NCTC ----------National Collection of Type Culture
EDTA----------Ethylenediamine-tetraacetic acid
KCl ----------Potassium chloride
$MgCl_2$----------Magnesium chloride
SDS----------Sodium dodecyl sulphate
TAE----------Tris acetic acid EDTA buffer
LB----------Luria-Bertani medium
OD---------- Optical density
PBS----------Phosphate-buffered saline
CFU----------Colony forming units
BHI ----------Brain-heart infusion medium
ORF----------Open reading frame
SCLE ----------Spore cortex lytic enzyme
TEM ----------Transmission electron microscope
NCBI ----------National Centre for Biotechnology Information
BLAST ----------Basic Local Alignment Search Tool
BLASTn ---------- Nucleotide BLAST
BLASTp ---------- Protein BLAST
Pfam ----------protein families' database
BCA ----------Bicinchoninic acid
Ni-NTA ----------Nickel-nitrilotriacetic acid
IPTG ----------Isopropyl β-D-1-thiogalactopyranoside
TBE ----------Tris-borate EDTA
SOC medium----------Super Optimal broth with Catabolite repression
CCFA----------Cycloserine-Cefoxitin-Fructose Agar
CDMN----------*Clostridium difficile* Moxalactam Norfloxacin
SCLEs---------- spore cortex-lytic enzymes
CFLEs ---------- cortical fragment-lytic enzymes
GSP ---------- germination-specific proteases
UV ---------- ----- Ultraviolet

Chapter One

Introduction

1.1 Historical review

The first mention of *C. difficile* in the literature was by Hall and O'Toole in 1935 when they considered it part of the normal body microbiota of the newborn (Brazier and Borriello, 2000). It was originally called *Bacillus difficile* due to it being difficult to isolate , however, as the organism is a strictly anaerobic, Gram-positive spore-forming bacillus, it was subsequently classified as a member of the *Clostridium* genus (Brazier and Borriello, 2000). Tedesco *et al.* (1974) found that pseudomembranous colitis (PMC) occurred in approximately 10% of clindamycin-treated patients. In 1977, Bartlett and his colleague described clindamycin-induced colitis in hamsters, and were able to isolate an unidentified *Clostridium* species which was subsequently identified as *C. difficile* and implicated as the causative agent of the condition (Bartlett, 2002). Subsequently Larson *et al.* (1977) identified the presence of cytotoxins in the stools of patients with PMC. *C. difficile* is now recognised as the causative agent of PMC and of 15% to 25% of cases of antibiotic-associated diarrhea (Gerding *et al.*, 1995).

1.2 *Clostridium difficile* characterisation

C. difficile is ubiquitous in nature, being especially prevalent in humans. About 40-60% of neonates and 1-3% of healthy adults carry *C. difficile* asymptomatically (Kuijpers and Surawicz, 2008).

The bacterium is a Gram-positive, spore-forming, anaerobic bacillus (Brazier and Borriello, 2000). It's Gram stain is predominantly positive; however, some old colonies may exhibit Gram variability. Cells are between 3-5 μm in length and are usually round-ended. Organisms produce large oval subterminal spores, usually in the stationary or decline phase (Brazier and Borriello, 2000). Colony morphology is variable depending on the strain, inoculation media and incubation time (Brazier and Borriello, 2000). However, colonies cultured on blood-based media following 24 to 48 hours incubation show irregular, lobate or rhizoid edges, are grey in colour and opaque with no haemolysis and a fried egg appearance can be seen due to spore formation (Brazier and Borriello, 2000). Colonies are between 4.5 mm (blood agar) to 7.5mm (cycloserine cefoxitin fructose agar (CCFA) in diameter (George *et al.*, 1979). The vegetative form of the bacterium exhibits sluggish motility due to the presence of peritrichous flagellae. Some strains have multiple polar fimbriae which are up to 6 μm long and 4–9 nm in diameter (Hafiz and Oakley, 1976; Borriello *et al.*, 1988).

In the presence of oxygen, the vegetative form of *C. difficile* can remain viable for up to 15 minutes in a dry environment (Buggy *et al*., 1983) and for up to 3 hours in a moist environment (Jump *et al*., 2007).

The first complete genome sequence of a *C. difficile* strain was published in 2006 by the Sanger Centre, UK (Sebaihia *et al.,* 2006). This was of the *C. difficile* strain 630(NAP1/027), a hypervirulent and multidrug-resistant strain which isolated from a patient with severe PMC in Zurich, Switzerland.

The genome of strain has a single circular chromosome of 4,290,252 bp (G+C = 29.06%), and a circular plasmid of 7,881 bp (G+C = 27.9%). Analysis of the genome revealed that approximately 11% is comprised of mobile genetic elements, such as conjugative transposons, prophages, a *skin* element and other unknown mobile elements. These mobile elements could provide *C. difficile* 630 with the genes responsible for antimicrobial resistance and other virulence factors (Sebaihia *et al.,* 2006).

1.3 Spores of *C. difficile*

Spores were first described in 1876 by Robert Koch and Ferdinand Cohn (Brock, 1975; Labbe, 2005). In 1933, Bayne-Jones and Petrilli produced pictures showing sporulation and different morphological changes (Bayne-Jones and Petrilli, 1933; Labbe, 2005). Spores are dormant with a non-reproductive structure which is formed following a multiphase process involving asymmetric cell division, followed by engulfment of the new small cell containing a single DNA chromosome by the mother cell (Errington, 2003), and are normally surrounded by several envelope layers.

Structurally, the bacterial spore contains a core, a cortex and inner and outer spore coats (Sousa *et al.,* 1976). A loosely or closely fitting outermost structure has been observed around the spore in a number of species of *Bacillus* and *Clostridium* (Figure 1.1). This layer was given the name "exosporium" by Fluigge in 1886 (Gerhardt and Ribi, 1964).

Furthermore, Koch (1877) described the exosporium as a "round transparent mass which appeared like a small, light ring surrounding the spore" (Gerhardt and Ribi, 1964), and

acting as the primary physiological barrier between the spore and the environment (Gerhardt and Ribi, 1964; Holt and Leadbetter, 1969; Ohye and Murrell, 1973). It was suggested that the exosporium may play a role in resistance to protease and lytic enzymes (Sousa *et al.,* 1976) or may play role in spore attachment to surfaces and human gut epithelial cells (Joshi *et al.,* 2012). In terms of the chemical composition of the spore layers, the outer and inner spore coats are mainly protein, while the cortex is mainly peptidoglycan; the core contains proteins, DNA, RNA and ribosomes (Lawley *et al.,* 2009). The exosporium layer is composed of phospholipid, lipid, protein and polysaccharide (Labbe, 2005).

It has been found that *C. difficile* spores have extra germination activity in the presence of cholate derivatives, which are usually found in bile (Sorg and Sonenshein, 2008), which may explain why the spores undergo germination in the intestinal tract. However, spores do not possess a known germinant receptor such as those present in *Bacilli* and many *Clostridia* species and perhaps they employee a novel mechanisms for response to germinates (Lawley *et al.,* 2009). Recently reported proteomic and genomic analysis of *C. difficile* spores has revealed that five spore genes unique to *C. difficile* (CD0939 , CD0940 , CD1898 ,CD2010 and CD2926) are encoded within various mobile elements, including genomic islands , transposons and prophage 1 and 2 which may directly contribute to spore composition (Lawley *et al.,* 2009).*Clostridium difficile* spores may survive on contaminated surfaces for an extended period of time and are resistant to many chemical and physical agents (Setlow, 2007).

It is this feature that makes *C. difficile* an enormous challenge for infection control systems in hospitals and health care facilities (Nerandzic and Donskey, 2010).

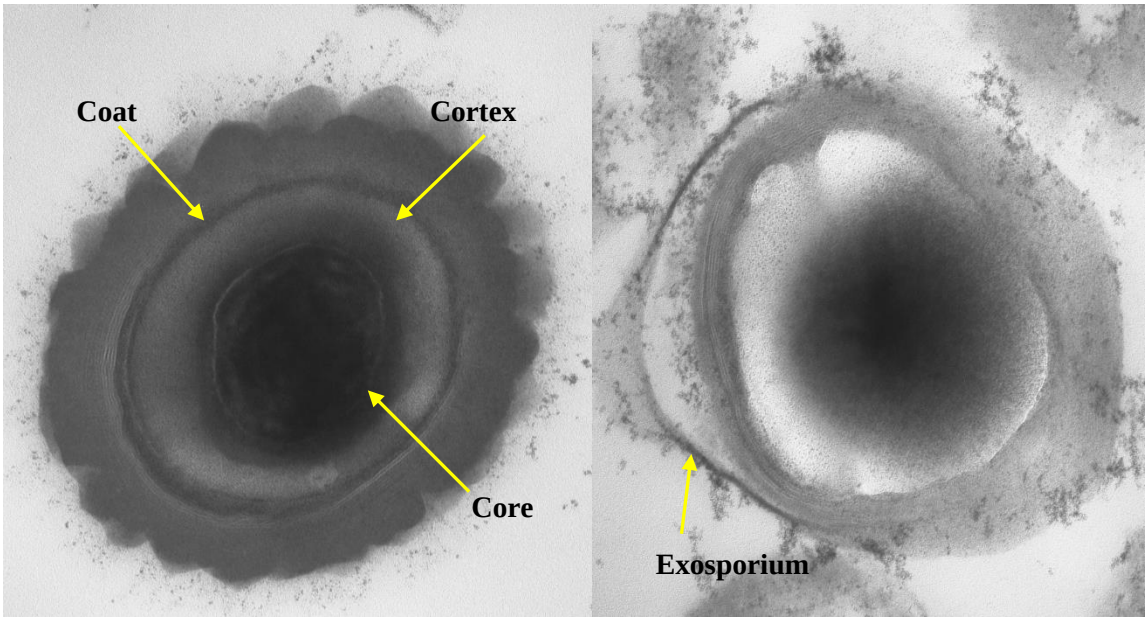


Figure 1.1. Transmission electron microscopy of sectioned *C. difficile* DS1813 spores (ribotype 027), showing the spore's ultrastructure including the core, cortex, coat and exosporium. Thin sections were double-stained with 2% uranyl acetate and 2% lead citrate and viewed on a Philips transmission electron microscope, EM 280, at 55,000x .

1.4 Sporulation and germination of spores

Sporulation is a multiphase process resulting in the development of a spore from a vegetative cell under conditions of environmental stress, such as a reduction in available nutrients. Germination is defined as the processes that result in the irreversible loss of spore-specific characteristics during the transition to a metabolically active vegetative cell following exposure to sense-specific effectors called germinants (Gould, 1970; Russell, 1990; Burns *et al*., 2010).

The vast majority of studies that have described the events of sporulation and germination have been performed on *Bacillus* species (especially *Bacillus subtilis* and *Bacillus cereus*). Little is known about *Clostridium* spore germination; however, *Bacillus* and *Clostridium* have some similarities in the sporulation process (Burns *et al*., 2010; Nerandzic and Donskey, 2010).

Spore sporulation

Sporulation is classically divided into seven distinct morphological stages (Gombas and Labbe, 1981; Russell, 1990; Labbe, 2005; McDonnell, 2007; De Hoon, *et al*., 2010). Stage 0 represents the vegetative cell, followed by stage I, the presporulation phase, where DNA is present as an axial filament rather than the normal spherical shape seen in normal cells (Figure 1.2). This is followed by stage II, where vegetative cells undergo septation in which asymmetric cell formation occurs. In stage III, the mother cell engulfs the new small cell containing a single DNA chromosome to form a forespore, which lies in the vegetative cell cytoplasm and is bounded by inner and outer forespore membranes. A thin layer of the vegetative cell wall lies between them, which is composed of peptidoglycan formed in stage IV; this layer is known as the cortex. The cortex is composed of modified peptidoglycan from the vegetative cell. This modification includes a repeating subunit of muramic lactam without any attached amino acids. Moreover, the alanine subunit is an L-alanyl residue that forms a tetrapeptide subunit with the sequence L-ala-D-glu-meso-DAP-D-ala with very little cross-linking between the tetrapeptide chains.

In stage V, the synthesis of spore coats and spore-specific molecule dipicolinic acid (DPA) occurs, followed by stage VI in which involves increase thickness of outer spore coat with further development of the cortex and loss of water from the core until a complete endospore is formed.

In stage VII, degradation of the mother cell takes place by the activity of autolysin enzymes in order to release the mature endospore. The spore remains in the dormant stage until the external environment once more becomes favourable for growth.

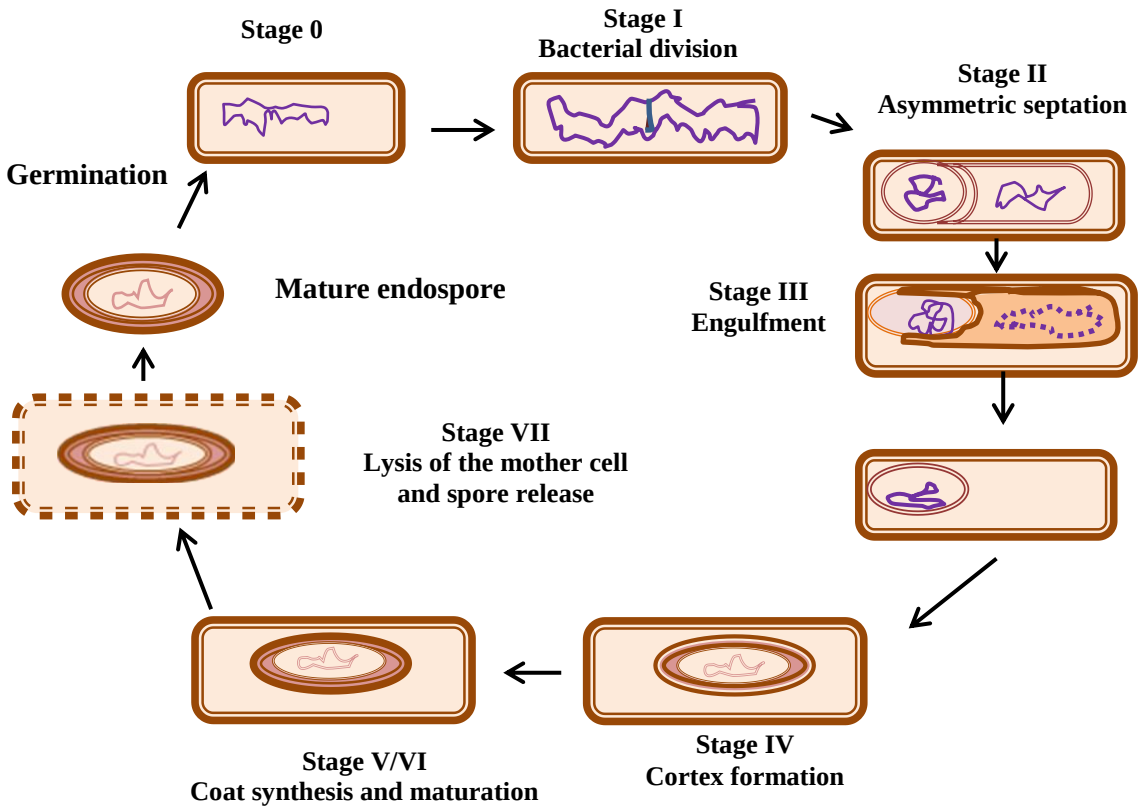


Figure 1.2. Stages of the *B. subtilis* life cycle. Modified from De Hoon, *et al.* (2010).

Spore germination

Spore germination in *Clostridia sp*. proceeds more slowly than in *Bacillus sp.* and often involves a mixture of germinants (Peck, 2009; Xiao *et al*., 2011). For example, *C. perfringens* spore germination can be triggered by a mixture of NaP and inorganic phosphate (NaPi) , a mixture of L-asparagine and KCl , and can also be triggered by L-asparagine, KCl and Ca-DPA (Paredes-Sabja *et al*., 2008a; Xiao *et al*., 2011). In *C. difficile*, spores can be stimulated by cogerminants, cholate derivatives, and glycine in bile (Sorg and Sonenshein, 2008).

During germination, dormant spores change to actively growing cells; this occurs in three stages (activation, germination and outgrowth), and occurs under favourable growth conditions that signal the process of germination (Russell, 1990). Germinants that trigger germination are thought to interact with specific receptors on spores (Burns *et al*., 2010). Activation is the process that prepares the dormant spore to enter germination and can be induced by metabolic and non-metabolic stimuli such as heat, acidity, ions and minerals. Activation is reversible and the spore retains most of its properties.

This step is followed by germination, which involves the release of dipicolinic acid (DPA) and calcium ions and the core layer rehydration. Spore lytic enzymes such as CwlJ and SleB in *B. subtilis* (Popham *et al*., 1996), SleC (Miyata *et al*., 1995) and SleM (Chen *et al.,* 1997) in *C. perfringens* and SleC in *C. difficile* (Burns *et al.,* 2010), are activated and then cortex peptidoglycan degraded, leading to full spore rehydration. ATP synthesised and other enzymes are reactivated. Spores DNA are released after the small acid-soluble proteins (SASPs) are degraded by specific protease.

The outgrowth step follows germination, during which actively dividing cells are generated (Russell, 1990; Makino and Moriyama, 2002; Labbe, 2005; Moir, 2006; Burns *et al*., 2010).

1.5 Pathogenesis

Healthy adults are protected from *C. difficile* colonisation and disease by the presence of the resident colonic microbiota, by a mechanism known as colonisation resistance (Gorbach *et al.,* 1988). This mechanism is likely to involve a number of factors such as the production of hydrogen peroxide, lactic acid, short chain fatty acids and bacitracin by the resident microbiota in addition to competition for mucosal binding sites and nutrients (Payne *et al.,* 2003; Naaber *et al.,* 2004).

However, administration of broad spectrum antibiotics disrupts this protective mechanism, allowing *C. difficile* spores acquired from either exogenous or endogenous sources to colonise and germinate in the small bowel upon exposure to bile acids (Kelly *et al.,* 1994; Kelly and Lamont, 1998). The resulting vegetative bacteria adhere to the cells lining the gut with the aid of flagella and proteases, so that *C. difficile* can then adhere to enterocytes (Denève *et al.,* 2009).

Pathogenic strains of *C. difficile* produce two distinct toxins. Toxin A (*tcdA*), which consists of 2710 residues (308.0 kDa), is an enterotoxin, whereas toxin B (*tcdB*), containing 2366 residues (269.6 kDa), has been described as a cytotoxin (Taylor *et al.,* 1981). However, some virulent strains may only produce toxin B (Alfa *et al.,* 2000).

These A^-B^+ strains are capable of causing disease and have been involved in outbreaks among hospitalised patients (Wren *et al.,* 2009). These toxins are the principal virulence factors of the pathogen and act by glucosylating small Rho proteins using UDP-glucose as a sugar donor; this damages the human colonic mucosa (Jank *et al*., 2007). The actions of these toxins on gut cells result in the production of proinflammatory interleukins and tumour necrosis factor-α*,* which increase vascular permeability and cause the opening of epithelial cell junctions and epithelial cell apoptosis. PMC with watery diarrhea forms as a consequence of connective tissue degradation (Castgliuolo *et al.,* 1998; Jank *et al*., 2007).

In terms of clinical presentation, a patient's response will range from an asymptomatic carrier state to active infection depending on host-specific factors (Borriello, 1998). The major factors predisposing patients to the development of symptomatic *C. difficile*-associated diarrhea are the type and duration of antibiotic therapy, the age of the patient and prolonged hospitalisation (Bignardi , 1998). Elderly patients in hospitals and nursing homes are at high risk of colonisation, with a rate of 73% having been reported (Denève *et al.,* 2009).

Other factors have been described, such as admission to intensive care units, the number and severity of underlying diseases and treatment with proton-pump inhibitors (Dial *et al.,* 2005). Also at risk are patients who have recently received immunosuppressive therapy or undergone gastrointestinal surgery, have a nasogastric tube or are sharing a hospital room with a *C. difficile*-infected patient (McFarland, 1998; Greenstein *et al.,* 2008).

1.6 Incidence of *C. difficile* infection

According to the health protection agency (HPA) report, dating 15th February 2013, the number of laboratory reports in England, Wales and Northern Ireland mentioning *C. difficile* infection decreased for the five consecutive year by 25 % from 19807 in 2011 to 14910 in 2012 after increasing every year since records began in 1999, as shown in Figure 1.3.

According to the European Centre for Disease Prevention and Control (ECDPC), the cost of treating a single case of *C. difficile* infection in England is estimated at between €5,000 and €15,000. The cost is expected to almost double over the next four decades.

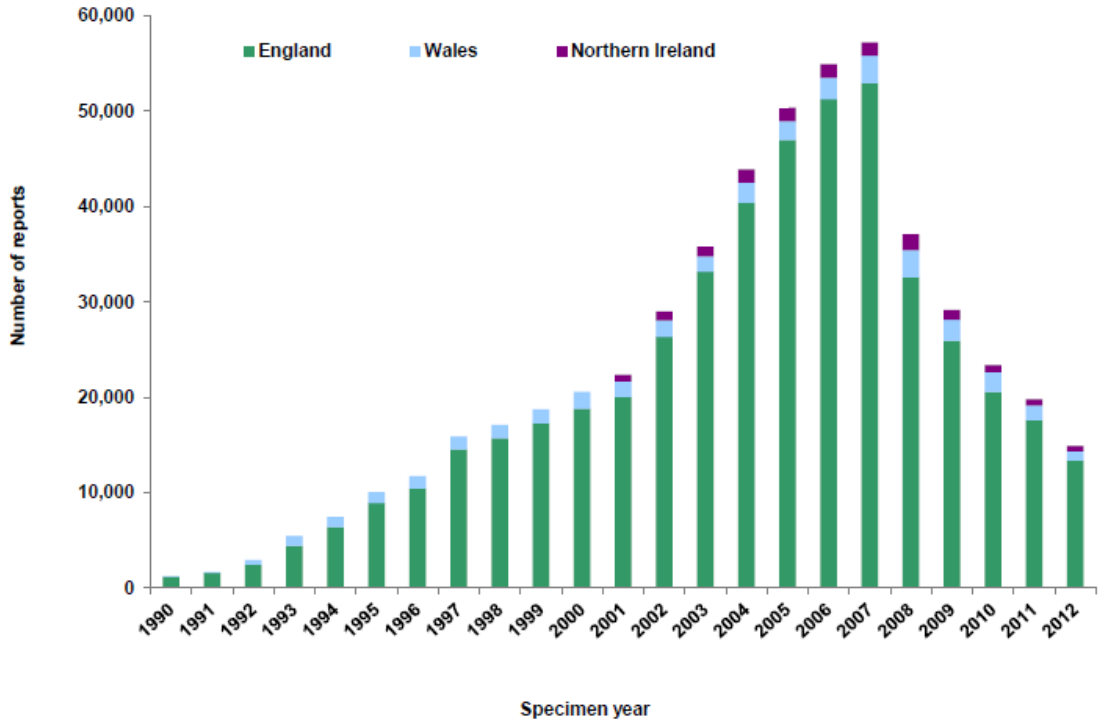


Figure 1.3. Voluntary laboratory reports of *C. difficile* positive faecal specimens: England, Wales and Northern Ireland (1990 – 2012).

http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317138039648, accessed on 10-06-2013).

1.7 Treatment strategies

The most important step is to discontinue offending agent, if possible, while giving fluids and electrolytes to maintain hydration. With this conservative strategy, up to 25% of patients with *C. difficile*-associated diarrhea (CDAD) recover without further therapy (Barbut *et al.,* 2000).

Metronidazole is recommended as a first-line treatment for CDAD, at a dose of 500 mg orally three or four times a day for up to 14 days, depending on the severity of the disease. Metronidazole is an inexpensive drug with a high positive response rate (Wenisch *et al.,* 1996). Metronidazole is one of the most important drugs for the treatment of anaerobic infections. It has a low molecular weight that allows the drug to diffuse passively into bacterial cells and across cell membranes. The agent is then reduced by the pyruvate :ferredoxin oxidoreductase system of bacterial cells, resulting in the formation of a cytotoxic nitro radical anion which interacts with the host cell's DNA, resulting in DNA strand breakage (Edwards, 1993; Cudmore *et al.,* 2004).

The antibiotic vancomycin is recommended as a second-line therapy. It is a glycopeptide that forms a complex with the D-Ala-D-Ala termini of peptidoglycan precursors in the outer surface of the bacterial membrane, and this complex results in cell wall synthesis inhibition (Nagarajan, 1991). The standard treatment with this drug is 125 to 500 mg orally four times a day for 10 to 14 days. It is also an effective treatment, with a response rate greater than 90% (Fekety *et al.,* 1989; Bartlett, 2002). However, due to the higher cost of oral vancomycin therapy and increased concern about selection for vancomycin resistant *Enterococcus* (VRE), metronidazole is preferred as the initial agent of choice

(Bartlett, 2002). Despite this, a recent study has revealed that vancomycin gives a better therapeutic response than metronidazole with a lower relapse rate in patients with severe *Clostridium difficile*-associated infection (CDAI) (Bartlett, 2006). Alternative antibiotics have been described, including oral therapy with teicoplanin, bacitracin or fusidic acid (Bartlett, 2002).

Resolution of fever, usually within the first two days, is a sign of a good response to therapy, and diarrhea should resolve within two to five days. Treatment should be continued for an average of 10 to 15 days, and at least five days are required to determine whether the treatment has failed or not (Florea *et al.*, 2003). The use of antiperistaltic drugs to reduce diarrhea should be avoided as they can result in the loss of metronidazole (Fekety, 1997).

Unfortunately, 15 to 25% of patients with *C. difficile* infection will suffer from recurrent infection following antibiotic treatment (Liang, 2003). This may be due to the germination of *C. difficile* spores which have survived the initial course of treatment (original organism) or as a result of reinfection by a new strain of the organism (Kelly *et al.*, 1994). Metronidazole remains the drug of choice in cases of recurrent infection (Gerding, 2000).

Concerns over the increased incidence of CDAD, the emergence of strains with enhanced virulence and an increase in the level of antibiotic resistance has prompted researchers to investigate the utility of employing bacteriophages (denoted as φ) to treat infections caused by this organism. Bacteriophages are viruses that only target bacterial cells and have been successfully used in the former Soviet Union and Eastern European countries

for many years to treat infectious diseases, often in preference to antibiotics (Sulakvelidze *et al.,* 2001; Fortuna *et al*., 2008; Kutateladze and Adamia, 2008).

A major advantage of their use in the context of *C. difficile* is their specificity; they are normally species-specific, meaning that other members of the bacterial microbiota of the gut are unaffected, in contrast to antibiotics. To date, several temperate bacteriophages have been shown to be active against *C. difficile* (Mayer *et al.,* 2008; Meader *et al.,* 2010 Sekulovic *et al.,* 2011). Indeed, the feasibility of employing bacteriophages to treat experimentally infected animals has been demonstrated in hamsters (Ramesh *et al.,* 1999).

Endolysins, which are bacteriophage enzymes that degrade the bacterial peptidoglycan cell wall, resulting in lysis and death of the bacterial cell, have also been tested as antimicrobial agents against *C. difficile*. Endolysin, isolated from φCD27, was found to be active against 30 different *C. difficile* strains, including the major epidemic ribotype 027(B1, NAP1) (Mayer *et al.,* 2008; Mayer *et al.,* 2011).

In addition to approaches which directly target the bacteria, other potential therapies have been investigated as a means of treating the disease. Rectal infusions of faeces from a normal host have been investigated as means of restoring the normal microbiota of the colon in patients who suffer from recurrent *C. difficile* infection. Despite the good response rates which have been seen with this approach, there is a risk of transmitting retroviruses and other infectious agents from the donor to the recipient (Gustafsson *et al.,* 1999).

A further live organism based approach which is being explored is the use of probiotic agents such as *Saccharomyces boulardii* (Surawicz *et al.,* 1989) or *Lactobacillus rhamnosus* GG. *Lactobacillus acidophilus, Lactobacillus bulgaricus* and *Lactobacillus plantarum* 299v have also been found to be effective in patients with recurrent infections (Pochapin, 2000; Miller, 2007). In addition to directly targeting the bacteria, researchers have investigated the feasibility of inactivating the pathogen's principle virulence factors, its two toxins. Anion resins, such as colestipol or cholestyramine, which bind the toxins of *C. difficile*, have been found to be effective in some patients with recurrent CDAD when used in combination with vancomycin (Poutanen and Simor, 2004). Tolevamer is another toxin sequestering compound which binds the toxins to a styrene sulphonate polymer. The polymer lacks antimicrobial activity and thus should permit the restoration of a normal gut microbiota and thereby prevent recurrent infections. However, this agent is still under investigation (Barker *et al.,* 2006; Louie *et al.,* 2006; Baines *et al.,* 2009).

1.8 Prevention

The most important strategy for preventing *C. difficile* disease is the proper use of antibiotics. In hospital, most patients who develop CDAD have previously been exposed to antimicrobials. In addition, a decrease in CDAD cases in the hospital environment has been shown to be associated with judicious use of antibiotics (McNulty *et al.,* 1997).

According to the WHO Infection control measures, such as hand washing between patient contacts, environmental decontamination, and good hospital cleaning are the only

ways to reduce the spore burden (WHO, 2002). The use of disinfectants is controversial , although vegetative organisms might be killed, some spores will survive. Staff and patient education, as well as adequate nursing levels, are also vital. All of these measures, when correctly implemented, have proven to be effective at reducing the incidence of the disease (Vonberg *et al.,* 2008).

The use of probiotics containing live bacteria such as *Lactobacillus casei*, *L. bulgaricus* and *Streptococcus thermophilus* has been reported to reduce the incidence of antibiotic-associated diarrhea and *C. difficile*-associated diarrhea (Surawicz *et al.,* 1989; Miller, 2007). Hickson *et al.* (2007) stated that this approach has the potential to decrease healthcare costs, morbidity, and mortality, if applied routinely in patients aged over 50.

Vaccination is currently being investigated as a means of preventing infection (Kotloff *et al*., 2001 ; Aboudola *et al.,* 2003). A *C. difficile* toxoid vaccine capable of inducing high-level responses in the sera via anti-toxin A immunoglobulin G (IgG) has been developed and used in healthy volunteers (Aboudola *et al.,* 2003). A vaccine containing formalin-inactivated toxoid A and B has also been developed and has been shown to be highly immunogenic in healthy adults and may be capable of improving patient outcomes (Kotloff *et al*., 2001). Further testing will be required to validate the efficacy of this vaccine (Mattila *et al.,* 2008).

Vaccination is not the only form of immunotherapy. Human monoclonal antibodies (HuMAbs), anti-toxin A HuMAb CDA1 and anti-toxin B HuMAb MDX-1388, which neutralise toxin A and toxin B, respectively, have been developed and have been shown in animal studies to prevent disease (Babcock *et al.,* 2006). A further study in 2008 showed that CDA1 was safe and well-tolerated at single doses after being used in healthy adults, suggesting that this intervention could be used to treat infected individuals (Taylor *et al.,* 2008).

1.9 Bacteriophages

1.9.1 Introduction

Bacteriophages (or phages) are viruses that target and infect only bacterial cells. Phages are ubiquitous and can be found in all habitats populated by bacteria (Sherman, 2008). They are the most abundant life forms on Earth and approximately 5,100 phage types had been isolated and identified by the end of the last century (Ackermann, 2001).

1.9.2 Structure

A phage has an outer protein shell called the head or capsid, which can vary in size and shape (Guttman *et al.,* 2005). Some are icosahedral and others are filamentous. The head, which is present in every phage, acts as a protective covering and encloses the phage's genetic material. The genetic material can be ssRNA, dsRNA, ssDNA or dsDNA, in either a circular or linear arrangement (Guttman *et al.,* 2005). Some phage heads have tails attached to them, through which the nucleic acid passes during infection of the bacterial cell. The tail may be surrounded by a contractile sheath, as in the T4 phage. At the end of the tail, some phages have tail fibres and a base plate; these structures play a role in the attachment of the phage to the bacterial cell (Hanlon, 2007) (Figure 1.4).

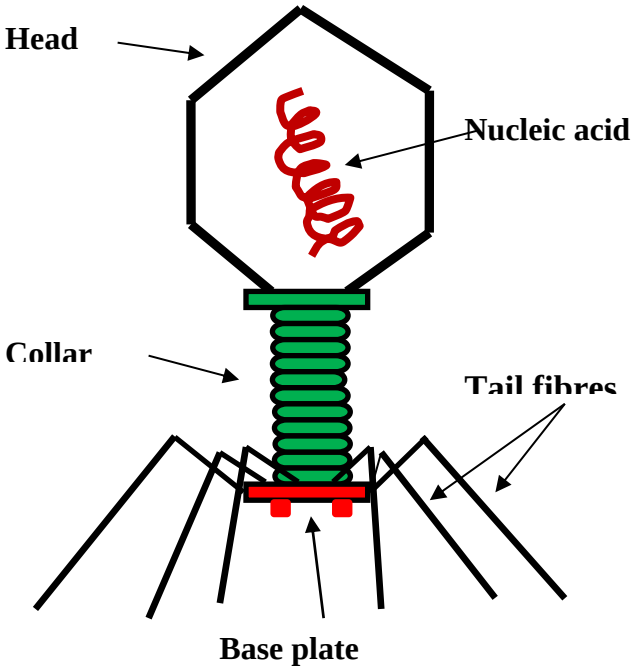


Figure 1.4. Illustration of the basic structural features of bacteriophages. Image is modified from Hanlon (2007).

Phages can be classified based on their morphology, presence or absence of an envelope and type of nucleic acid . To date, phages have been classified into 13 families (Parisien *et al.*, 2008) (Figure 1.5). Approximately 96% of reported phages are "tailed phages", which are composed of an icosahedral head and a tail, with double-stranded DNA (dsDNA), These phages are classified according to the morphological features of the tail into three families: *Podoviridae* (short tail; 13.9%) *Myoviridae* (contractile tail; 24.5%), and *Siphoviridae* (long non-contractile tail; 61.7%) (Ackermann, 1996).

The remaining phage types, constituting only 186 (3.6%) of the 5,100 known phages, are cubic, filamentous or pleomorphic in shape with genomes comprising single or double-

stranded DNA (ssDNA or dsDNA), or single or double-stranded RNA (ssRNA or dsRNA) (Ackermann, 2001).

Phages can also be classified based on their way to escape from their host cell, i.e. filamentous phages and lytic phages (Young *et al*., 2000; Bernhardt *et al.,* 2002). Filamentous phages (shaped like a rod filament) can, because of their unique morphology and morphogenesis, reproduce and disseminate without killing their host (Russel, 1995), whereas lytic phages lyse and destroy the bacterial cell wall, resulting in the death of the host (Borysowski *et al*., 2006).

Lytic phages can be further divided into two classes according to their cell wall lytic mechanisms: phages which employ an endolysin-holin system to destroy the wall of the bacterial cell (this system is encoded in phages that have large double-stranded DNA) and phages that use a single lytic factor which inhibits peptidoglycan synthesis (these usually occur in phages having small single-stranded RNA or DNA) (Young *et al.,* 2000).

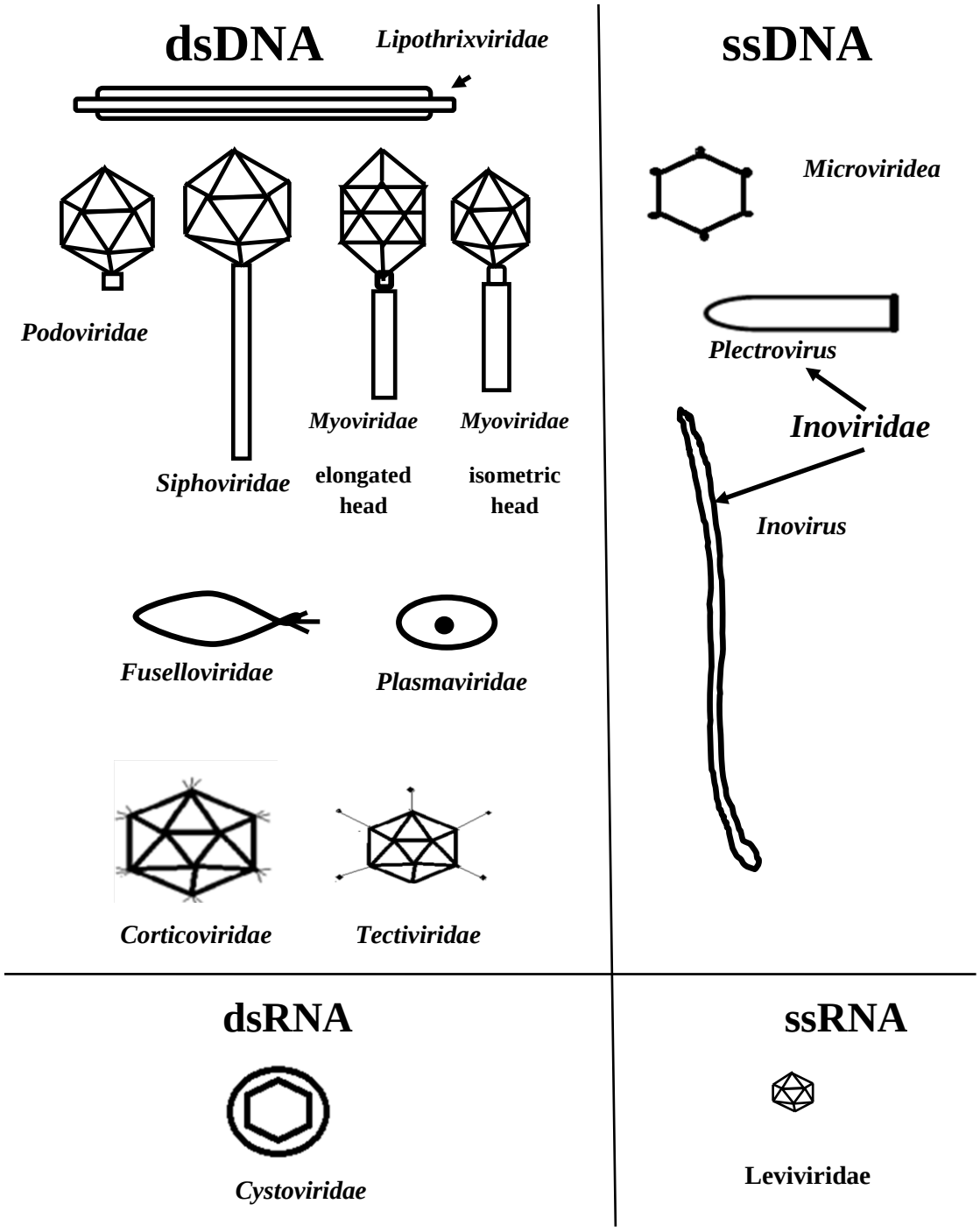


Figure 1.5. The basic structure of families of bacteriophages. Images are modified from Murphy *et al*. (1995).

1.9.3 Replication

Phages have either a lytic or a lysogenic life cycle. Those with a lytic life cycle infect a host cell, resulting in lysis of the host cell, whereas phages with a lysogenic life cycle can either multiply via the lytic cycle or enter a dormant state inside the cell (Sulakvelidze *et al.,* 2001). The life cycle starts with infection of the host cell. The bacteriophage binds to specific receptors, which may be present on the cell envelope, capsule, flagella or even conjugative pili, by their tail fibres or by some other structure in those phages that lack tail fibres (Young *et al.,* 2000). However, this attachment is weak and needs to be supported by irreversible attachment via the base plate.

The main reason for this process is to allow the phage to bind tightly to the bacterial cells Penetration then takes place by which the phage injects its genome into the cell cytoplasm (Guttman *et al.,* 2005).

The lytic life cycle

As already stated, lytic or virulent phages multiply inside their bacterial victim and kill the cell by lysing its cell wall at the end of the reproductive cycle. The ability to specifically target and kill a particular species of bacteria has led to the development of phages as a treatment for infectious diseases.

Following penetration of the bacterial cell wall, the genome of the phage circularises in the cytoplasm of the bacteria. The phage's nucleic acid uses the host cell's machinery to

synthesise large amounts of structural proteins and phage DNA. Progeny phages are formed by the self-assembly of these components which escape from the cell when the cell wall is lysed by the actions of a phage-encoded endolysin (Guttman *et al.,* 2005).

The lysogenic life cycle

Temperate bacteriophage is one which is capable of undergoing a lysogenic life cycle. After infection the temperate phage can either integrate its genome into their host bacterium's and become a non-infectious prophage, or it can replicate using lytic life cycle which result in lysis of host cell (Sulakvelidze *et al.,* 2001).

The insertion of phage DNA into the genome occurs at specific sites and is catalysed by a phage-encoded enzyme, such as att P in lambda phage, which reacts with the host integration site att B in *E. coli* which is catalysed by integrase (Int) and host integration factor (IHF) (Little, 2005). In order to maintain the lysogenic state, a phage-coded protein, called a repressor binds to an operator on the phage DNA and turns off transcription of phage genes. The result is a stable lysogenic phage genome which remains integrated into the host's DNA (Little, 2005). As the bacterial cell replicates, the prophage replicates as part of the genome. When a prophage is induced, phage gene expression is no longer repressed and, as a consequence, viral proteins and copies of the viral genome are generated. Progeny virus particles are formed and released from the host cell after completing the life cycle (Hanlon, 2007) (Figure 1.6).

Release

Lytic phages produce an enzyme called an endolysin which disrupts the peptidoglycan structure of the bacterial cell wall and affects cross-linking. Peptidoglycan has a role to play in defining the structure of the bacteria and prevents the cell from bursting due to the osmotic pressure of the bacterial contents. Weakening of peptidoglycan means that the cell wall can no longer contain the pressure and the cell bursts (Fischetti, 2008). However, not all phages are lytic; filamentous phages replicate and escape from the bacteria cell without killing the organism (Russel, 1995).

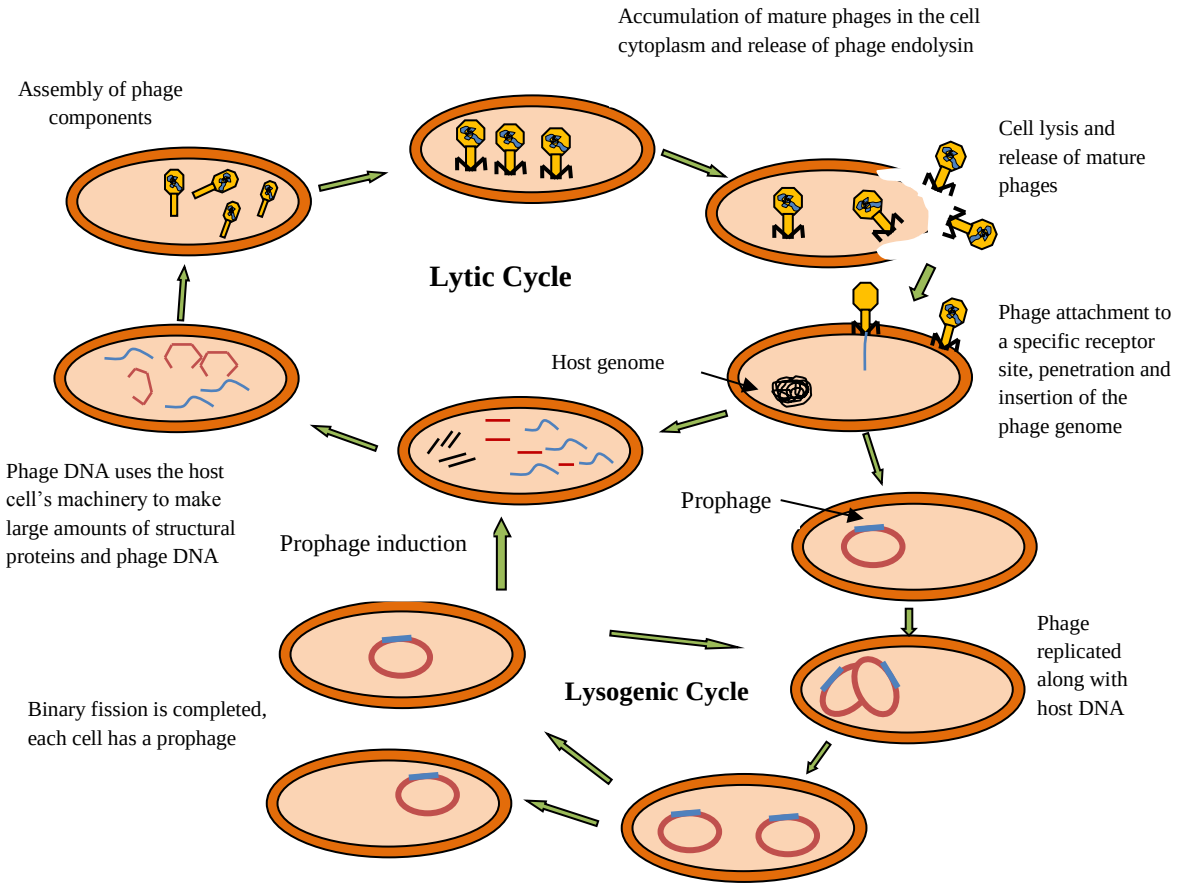


Figure 1.6. Bacteriophage life cycle (lytic and lysogenic cycles). Image is modified from Sulakvelidze *et al*. (2001)

1.10 Phage therapy

1.10.1 Phage discovery and early therapeutic approaches

The concept of using phages as antibacterial agents is not new. In 1896, a British bacteriologist, Ernest Hankin, reported that the water of the Ganges and Jumna rivers in India seemed to possess antibacterial capabilities against gastrointestinal infections, particularly cholera. He described a substance which was capable of passing through a fine porcelain filter and was heat labile, as an undefined substance (Sulakvelidze and Morris, 2001). Two years later, Gamaleya confirmed Hankin's observation with *Bacillus subtilis*. In 1915, Frederick Twort isolated filterable entities that produced small cleared areas on bacterial lawns due to the destruction of bacterial cells (micrococci and intestinal bacilli). Furthermore, he found the same phenomenon with *Staphylococcus aureus*, which led him to suggest that this antibacterial activity could be due to the actions of a virus. However, Twort did not explore his findings further (Sulakvelidze and Morris, 2001).

In 1917, Felix d'Herelle named these viruses bacteriophages (bacteria eaters) (Hanlon, 2007) and realised their potential for the treatment of bacterial infections following an outbreak of severe haemorrhagic dysentery amongst troops stationed in Paris. At that time, he prepared filtrates from patients' stool samples and mixed and incubated these filtrates with *Shigella* strains isolated from the same patients. A portion of the mixtures was tested on experimental animals and a portion was spread on agar plates. The appearance of minor but clear lysis of bacterial cells was seen by d'Herelle on these plates and he later called these areas taches, and then plaques (Summers, 2001).

The first phage therapy experiments performed by D'Herelle, using a phage preparation against dysentery, gave extremely promising results with the elimination of infection (Sulakvelidze *et al.,* 2001). However, the results of these studies were not published immediately, and so the use of bacteriophages to treat staphylococcal skin disease in 1921, by Richard Bruynoghe and Joseph Maisin, was the first reported application of phages to treat infectious diseases in humans (Sulakvelidze *et al.,* 2001).

However, failures in phage therapy occurred in later years, which could be attributed to a variety of factors such as low quality of phage preparations, a lack of detailed understanding of phage biology and poor experimental methods (O'Flaherty *et al.,* 2009). The commercialisation of antibiotics in the 1940s led to decline in the use of phages as therapeutic agent in Western civilisation; however, in the East, they continued to use phages as therapeutic agent either alone or in combination with antibiotics (O'Flaherty *et al.,* 2009).

1.10.2 Reappearance of interest in phage therapy in the west

Bacteriophage therapy returned to the West in 1989 after the collapse of the Soviet Union. Despite some initial barriers, such as language problems and a lack of information in this field, the problems associated with increased levels of antibiotic resistance and advances in molecular techniques and purification methods encouraged scientists and clinicians in Europe and America to undertake clinical trials on phage therapy (Jonathan, 2009).

One example was the treatment of infected leg ulcers by using a cocktail of eight bacteriophages: five against *Pseudomonas aeruginosa*, two against *Staphylococcus aureus* and one against *E. coli*. This was carried out in Texas, USA, in 2006 (Marza *et al.*, 2006). Another example is the use of T4 bacteriophage by Nestlé (the Swiss multinational food corporation) in Bangladesh to treat *E. coli*-mediated diarrhea (Bruttin and Brüssow, 2005). Furthermore, a cocktail of six bacteriophages was used in 2007 at the Royal National Throat, Nose and Ear Hospital in London as a treatment for chronic inner ear infections caused by *Pseudomonas aeruginosa* with the phage cocktail being administered by aerosol (Jonathan, 2009). The use of a bacteriophage to treat cystic fibrosis patients infected with *Burkholderia cepacia* has recently been described (Golshahi *et al.*, 2008). Bacteriophage therapies have been used against a range of infections such as otitis media and urinary tract infections, and against bacteria such as *Pseudomonas aeruginosa*, *Klebsiella* spp., *Burkholderia*, Group A *Streptococcus* and *Bacillus anthracis* (Jonathan, 2009). Phages have also been used in food decontamination; for instance, in 2006, the FDA approved the use of a cocktail of six bacteriophages in order to eradicate strains of *L. monocytogenes* from ready-to-eat meat and poultry products (Macgregor, 2006). While phages have a number of unique advantages over antibiotics, the approach is not without its disadvantages. Details of some of the possible advantages and disadvantages are given in table 1.1.

Table 1.1. Advantages and disadvantages of phage therapy

Advantages	Disadvantages
• Specific for a particular bacterial target, so phages will not harm the beneficial bacteria that are present in the body (Sulakvelidze *et al.*, 2001).	• Phages have a limited host range due to the specificity of phage attachment. Therefore, the disease-causing organism has to be identified prior to the commencement of phage therapy (O'Flaherty *et al.*, 2009).
• Phage therapy can be delivered via numerous methods such as injection, nasal sprays, etc. (Abhilash *et al.*, 2009).	• As a result of bacterial cell mutation, resistance can be developed against phages (Skurnik and Strauch, 2006).
• Relatively simple and inexpensive to propagate compared with antibiotics manufacturing; therefore, treatment costs will be reduced (O'Flaherty *et al.*, 2009).	• The effectiveness of phage therapy against intracellular pathogens (e.g., *Salmonella* spp.) is unclear (Sulakvelidze *et al.*, 2001).
• Safe therapy; no serious adverse effects have been reported (Sulakvelidze and Morris, 2001).	• Phages may swap genes with each other and other organisms, so there is a chance of spreading toxins or antibiotic resistance genes (Parisien *et al.*, 2008).
• Self-replicating at the site of infection, so the number will increase as long as there are bacteria present. There is often no need to carry out repeat dosing and phages will disappear once the target is destroyed (Alisky *et al.*, 1998).	• Lack of pharmacokinetic data, making phage therapy risky (Parisien *et al.*, 2008).

1.11 Prophage induction

Treatments that damage cellular DNA or inhibit DNA replication can switch the prophage from a highly stable lysogenic state to a lytic state; this is called prophage induction (Janion, 2001).

1.11.1 The cellular response

The SOS (save our souls) response can be triggered by ultraviolet (UV) light, mitomycin C or another inducing agent. This response is designed to cope with any DNA damage, including that which happens during prophage induction, in order to repair the DNA and restore normal replication (Little, 1991). Prophage induction can be highly efficient at generating a high level of DNA damage, which leads to switching on the SOS response in all cells in a culture (Janion, 2001). However, switching takes place at a low rate in the absence of the DNA damage, which usually happens with spontaneous induction (Little, 1991).

The SOS regulatory system is controlled by two important proteins: LexA and RecA. The LexA protein acts as a repressor of SOS genes (there are about 40 of these genes in *E.coli*) during normal cell growth. Inducing treatments activate RecA, the major bacterial recombinase (Little, 2005). Little (1991) stated that inducing agents damage dsDNA, leading to the production of ssDNA, which works as an inducing signal for RecA, which binds to ssDNA. Binding leads to activation of RecA as a co-protease which then cleaves and inactivates LexA.

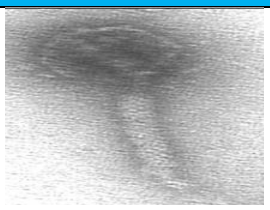
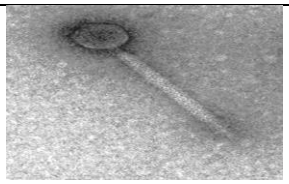
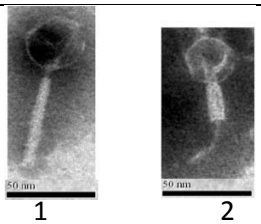
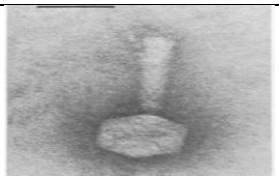
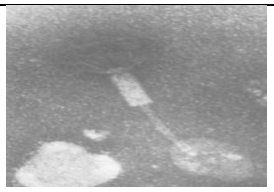
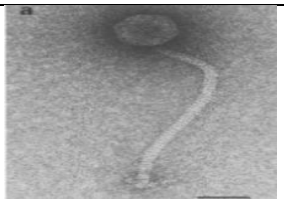
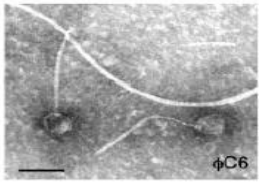
In the induced state, activated RecA continues to mediate LexA cleavage, and SOS genes are also expressed at a high level. If the prophage is present, the phage repressor gene is also cleaved, resulting in phage induction in a form that catalyses the specific proteolytic cleavage of LexA, which results in de-repression of SOS genes due to inactivation of the LexA repressor function. This mechanism usually happens in the λ phage and other related phages (Little, 2005). In addition, prophage 186 is known to be induced by treatment with UV light. In this phage, the *tum* gene is under the control of the LexA repressor. After the induction of the SOS system and inactivation of LexA, prophage induction is initiated by permitting expression of the *tum* protein product. This protein acts as an antirepressor, antagonising the repressor of 185 cI, which will result in phage release (Lamont *et al.,* 1989; Little, 2005). Other alternative mechanisms are also used, although some of the induction pathways are not well-understood (Little, 2005).

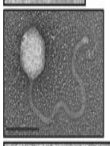
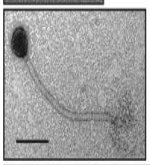
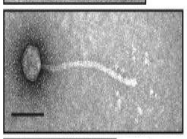
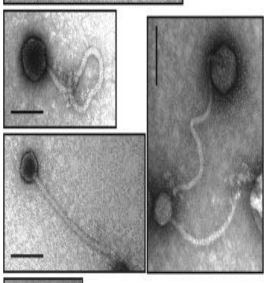
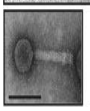
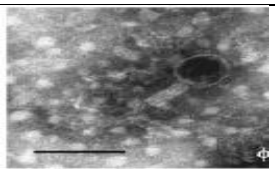
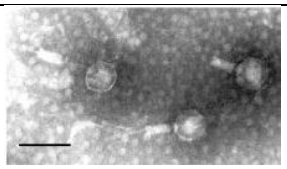
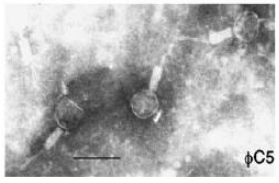
1.12 *C. difficile* phages identified to date

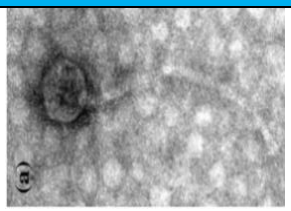
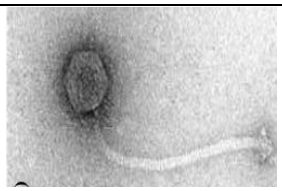
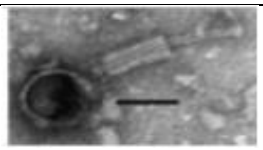
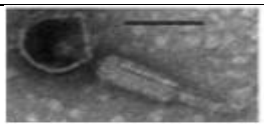
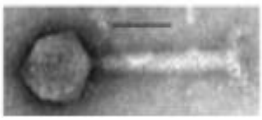
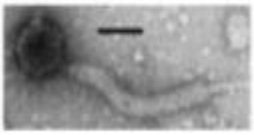
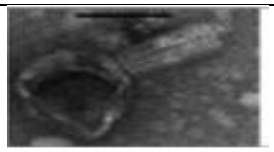
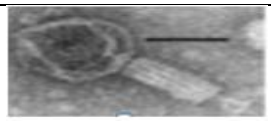
A number of bacteriophages have been used successfully in the typing of *C. difficile* strains (Sell *et al.,* 1983). Recently, there have been successful isolations of whole phages and purified phage components, with a number of temperate *C. difficile* phages having been isolated and observed by the use of electron microscopy (EM). These phages are classified into two families. Those of the *Myoviridae* family have an isometric capsid and a contractile tail, for example φCD5, φCD52, φCD119, φCD2, φC2, φCD27, φCD8, φCD630-1 and φCD630-2. In addition, prophages belonging to the *Siphoviridae* family have been isolated, for example φCD8-1, φCD8-2, φCD24, φCD38-1, φCD38-2, φCD6356 and φCD6365. This family has an isometric capsid and a long non-contractile tail (Table 1.2).

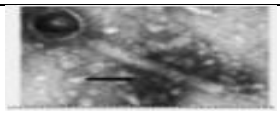
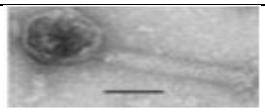
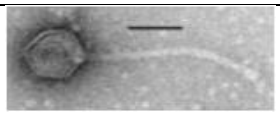
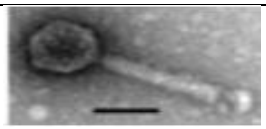
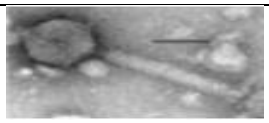
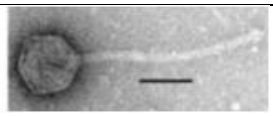
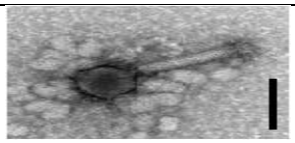
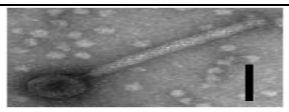
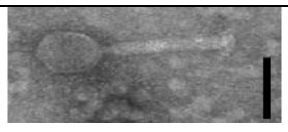
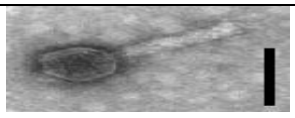
However, only seven phages have been fully sequenced and annotated, these being φCD119, φCD2 , φCD27 , φMMP03 and φMMP04 from the *Myoviridae* family (Govind *et al.,* 2006; Goh *et al.,* 2007; Mayer *et al.,* 2008; Meessen-Pinard *et al.,* 2012) and two phages from the *Siphoviridae* family, namely φCD6356 (Horgan *et al*., 2010) and φCD38-2 (Sekulovic, *et al.,* 2011).

Table 1.2. *C. difficile* phages identified so far.

Phage	Host strain	Indicator strain(s)	Family	EM image	Reference
φCD119	CD119	*C. difficile* strains 602, 660 and 460	*Myoviridae*		(Govind *et al.*, 2006; Govind *et al.*, 2009)
φCD27	NCTC 12727	NCTC 11204, 11205, 11207 and 11209	*Myoviridae*		(Mayer *et al.*, 2008)
φCD630 1 & 2	CD630		*Myoviridae*	 1 2	(Goh *et al.*, 2007)
φ56		79	*Myoviridae*		(Mahony *et al.*, 1985)
φCld 1			*Myoviridae*		(Sell *et al.*, 1983)
φ41		36	*Siphoviridae*		(Mahony *et al.*, 1985)
φC6	CD371	CD60	*Siphoviridae*		(Goh *et al.*, 2005b)

Phage	Host strain	Indicator strain(s)	Family	EM image	Reference
φCD5	CD5		*Myoviridae*		
φCD8-1	CD8		*Siphoviridae*		(Fortier and Moineau, 2007)
φCD8-2	CD8		*Siphoviridae*		
φCD24	CD24		*Siphoviridae*		
φCD38-1	CD38		*Siphoviridae*		(Sekulovic *et al.*, 2011)
φCD38-2	CD38	CD274	*Siphoviridae*		
φCD52	CD52		*Myoviridae*		
φC8	CD371	CD843	*Myoviridae*		(Goh *et al.*, 2005b)
φC2	CD242	CD062	*Myoviridae*		(Goh *et al.*, 2005b)
φC5	CD578	CD062	*Myoviridae*		(Goh *et al.*, 2005b)

Phage	Host strain	Indicator strain(s)	Family	EM image	Reference
φCD6356	DPC6356	DPC6365	*Siphoviridae*		(Horgan *et al*., 2010)
φCD6365	DPC6365	DPC6356	*Siphoviridae*		(Horgan *et al*., 2010)
φ001			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ002			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ013			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ014			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ015			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ020			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ027-1			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)

Phage	Host strain	Indicator strain(s)	Family	EM image	Reference
φ076			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ078			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ087-1			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ087-2			*Siphoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ107			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ220-1			*Myoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φ220-1			*Siphoviridae*		(Nale *et al*., 2012 ; Shan *et al*.,2012)
φMMP01	CD407	CD19	*Myoviridae*		(Meessen-Pinard *et al*., 2012)
φMMP02	CD408	CD117	*Myoviridae*		(Meessen-Pinard *et al*., 2012)
φMMP03	CD411 CD368	CD117	*Myoviridae*		(Meessen-Pinard *et al*., 2012)
φMMP04	CD412 CD380	CD73	*Myoviridae*		(Meessen-Pinard *et al*., 2012)

1.13 Phages and their relation to hosts

Phages may play an important role in infectious diseases. As a consequence of lysogenic infection, phages may add new properties to lysogen, especially with phages that carry a toxin gene or a drug resistance gene which may convert normally harmless bacteria into pathogens or make sensitive bacteria resistant (Grabow, 2001). Examples include pyrogenic exotoxin A production in group A *streptococci* (Johnson *et al.,* 1986), cholera toxin production in *Vibrio cholerae* (Waldor and Mekalanos, 1996) and shiga toxin-converting phages involved in the pathogenicity of enterohaemorrhagic *Escherichia coli* (Muniesa and Jofre, 1998).

In terms of *C. difficile,* temperate phages have been induced, isolated and characterised from different *C. difficile* isolates with different ribotypes, including the hypervirulent strain 027 (Fortier and Moineau, 2007; Nale *et al*., 2012; Shan *et al.,* 2012). Furthermore, the complete genome sequence of *C. difficile* strain CD630 has revealed that up to about 11% of the genome is composed of mobile genetic elements including two highly related prophages (Sebaihia *et al.,* 2006; Goh *et al.,* 2007). These results have revealed that *C. difficile* strains often harbour lysogenic phages in their genome (Govind *et al.,* 2009). However, to date, production of toxins in *C. difficile* has not been shown to be encoded or regulated by temperate phages. Furthermore, according to several studies, temperate phages have no ability to convert non-toxigenic strains to those capable of toxin production (Mahony *et al.,* 1985; Goh *et al.,* 2005a).

However, preliminary results have shown that infection of toxigenic strains by some phages could modulate toxin A and/or toxin B production (Goh *et al.,* 2005a; Govind *et al.,* 2009).

1.14 Bacteriophage endolysins

1.14.1 Introduction

To date, lytic phages have not been isolated from *C. difficile*, although temperate phages have been isolated and characterised as described before (Table 1.2) .One of the major challenges faced by researchers seeking to isolate temperate phages is highlighted by the narrow host range of the phages which have been isolated from *C. difficile*, possibly due to high-frequency carriage of prophages (Mayer *et al.,* 2008; Fortier and Moineau, 2007; Horgan *et al*., 2010; Nale *et al*., 2012) which can be attributed to the role of the system of clustered regularly interspaced short palindromic repeats (CRISPR) within these hosts (Deveau *et al.,* 2008), which may be problematic for therapeutic applications (Mayer *et al.,* 2008)

This highlights the need to investigate the potential of bacteriophage endolysins as antimicrobial agents. This has been done by Mayer and his group, who managed to sequence the φCD27 genome, identify the gene code for its endolysin and sub-clone this gene into *E. coli.* This endolysin was found to be active against 30 different *C. difficile* strains, including the major epidemic ribotype 027(B1, NAP1). It has a broader host range than the phage itself (Mayer *et al.,* 2008).

Endolysins are dsDNA bacteriophage-encoded enzymes that are designed to attack the peptidoglycan bonds in the bacterial cell wall. They accumulate in the bacterial cytoplasm during a lytic infection cycle. At the terminal stage of the phage life cycle, lysins undergo holin-mediated translocation, which results in the cleavage of one of the

five major peptidoglycan bonds, thereby allowing progeny phage release after lysis of the cell wall (Fischetti, 2008). Muralytic or mureolytic enzymes, lysins, lysozymes or virolysins are the alternative names of these enzymes (Loessner, 2005; Courchesne *et al.*, 2009). Their lytic action is regulated by small hydrophobic proteins called holins. These proteins which are encoded by the phage target the cytoplasmic membrane and produce patches, allowing the endolysins to pass through these patches to access the peptidoglycan (Young and Bläsi, 1995).

The most important thing that makes endolysins promising as potential antibacterial agents is their ability to act as exolysins that target and damage the cell wall rapidly when added exogenously (Borysowski *et al.,* 2006). Current data shows that these kinds of enzyme only work with Gram-positive bacteria *in vitro*, since the outer membrane does not exist, as it does in Gram-negative bacteria, thus enabling them to directly target cell wall compounds. However, in Gram-negative bacteria, when the outer membrane is disrupted by, for example, detergents, cells immediately become sensitive to endolysins (Loessner, 2005).

1.14.2 Structure

Endolysins from Gram-positive bacteria bacteriophages usually have a classically modular structure (250–400 amino acids, 25-40 kDa) with an N-terminal enzymatic domain which encodes for peptidoglycan hydrolase activity and C-terminal binding responsible for the species specificity of the molecule, separated by a short linker (Fischetti, 2008; Loessner, 2005) (Figure 1.7).

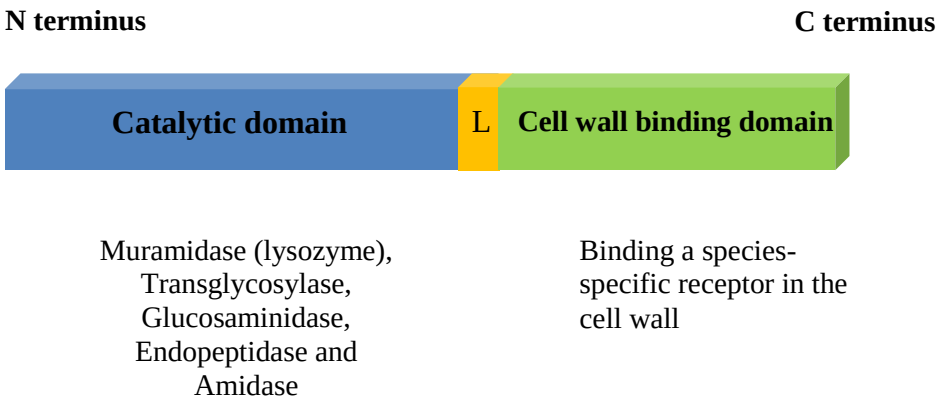


Figure 1.7. Modular structure of phage lytic enzymes. Image modified from Fischetti (2008).

The catalytic domain encodes the catalytic activity of the enzyme that can cleave one of the five major bonds in the bacterial cell wall peptidoglycan. These catalytic domains fall into one of five classes: Muramidase (lysozyme), Transglycosylase, Glucosaminidase, Endopeptidase and Amidase. For all endolysins reported to date, the majority appear to be amidases or muramidases (Fischetti, 2005). However, in rare cases, endolysins might have two or three different catalytic domains that link to one cell binding domain (Navarre *et al.*, 1999).

The cell wall binding domain at the C-terminus is responsible for binding species-specific carbohydrate epitopes in the cell wall, which offers a high degree of specificity. Catalytic domains in the lysin can be swapped, resulting in new activity against new targets (Fischetti, 2005). However, not all Gram-positive endolysins conform to this standard structure; for example, Gram-positive endolysins have been identified with only an N-terminal enzymatic domain, such as T4 lysozyme (Matthews & Remington, 1974).

Some possess two enzymatic domains, such as a pneumococcal phage Cpl-1 lytic enzyme which contains both muramidase and endopeptidase activity (Hermoso *et al*., 2003). Furthermore, an endolysin with three enzymatic domains has been described, staphylococcal phage187 lysin (Ply187), which contains an endopeptidase, an amidase and a glucosaminidase module (Cheng *et al*., 2005).

In contrast Gram-negative endolysins are smaller as they only possess the N-terminal enzymatic domain. However, a few endolysins with a modular structure have been identified and all have inverted molecular orientations compared to Gram-positive endolysins, such as the KZ144 endolysin of the *Pseudomonas aeruginosa* bacteriophage phiKZ which has an enzymatic domain at the C-terminal and the binding domain at the N-terminal (Briers *et al*., 2007). Another type of inverted endolysin with two C-terminal binding domains has been identified, i.e. OBPgp279 endolysin of *Pseudomonas putida* phage OBP (Cornelissen *et al*., 2011).

1.14.3 Mode of action

The bacterial cell wall is essential to bacterial viability since it protects the cell from mechanical damage and lysis. In Gram-negative bacteria cell wall is relatively thin about 10 nm and is composed of a single layer of peptidoglycan covered from the outside by the outer membrane. In contrast, Gram-positive bacteria, the cell wall is thick about 15-80 nm and consist of several layers of peptidoglycan (Winn *et al*., 2006 Borysowski *et al.,* 2006) (Figure 1.8).

The main building block in the cell wall is peptidoglycan, which is formed from linear chains of two alternating amino sugars, namely N-acetylglucosamine (GlcNAc or NAG) and N-acetylmuramic acid (MurNAc or NAM). These amino sugars are connected alternately by a β-(1,4)-glycosidic bond. Each N-acetylmuramic acid unit is attached to a short amino acid chain (four to five residues), usually consisting of L-alanine, D-glutamic acid, either meso-diaminopimelic acid in Gram-negative bacteria or L-lysine in Gram-positive bacteria, and D-alanine (Barer, 2007). Adjacent tetrapeptides are cross-linked by interpeptide bridges in Gram-positive bacteria or by interpeptide bonds in Gram-negative bacteria (Figure 1.9).

Holins target the cytoplasm by making pores that allow endolysin to pass through peptidoglycan (Loessner, 2005). According to the catalytic domain structure, endolysin will target a specific bond in the cell wall. As can be seen in Figure 1.9, N-acetyl-β-D-muramidase (or lysozyme) targets the β-1,4-O-glycosidic bond between C1 of N-acetylmuramic acid (MurNAc) and C4 of N-acetylglucosamine (GlcNAc), whereas N-acetyl-β-D-glucosaminidase cleaves the β-1,4-O-glycosidic bond C1 of GlcNAc and C4 of MurNAc. N-acetylmuramoyl-L-alanine amidases cleave the amine bond between L-alanine in the peptide chain and MurNAc. Endopeptidases cleave the peptide moiety (Borysowski *et al.,* 2006).

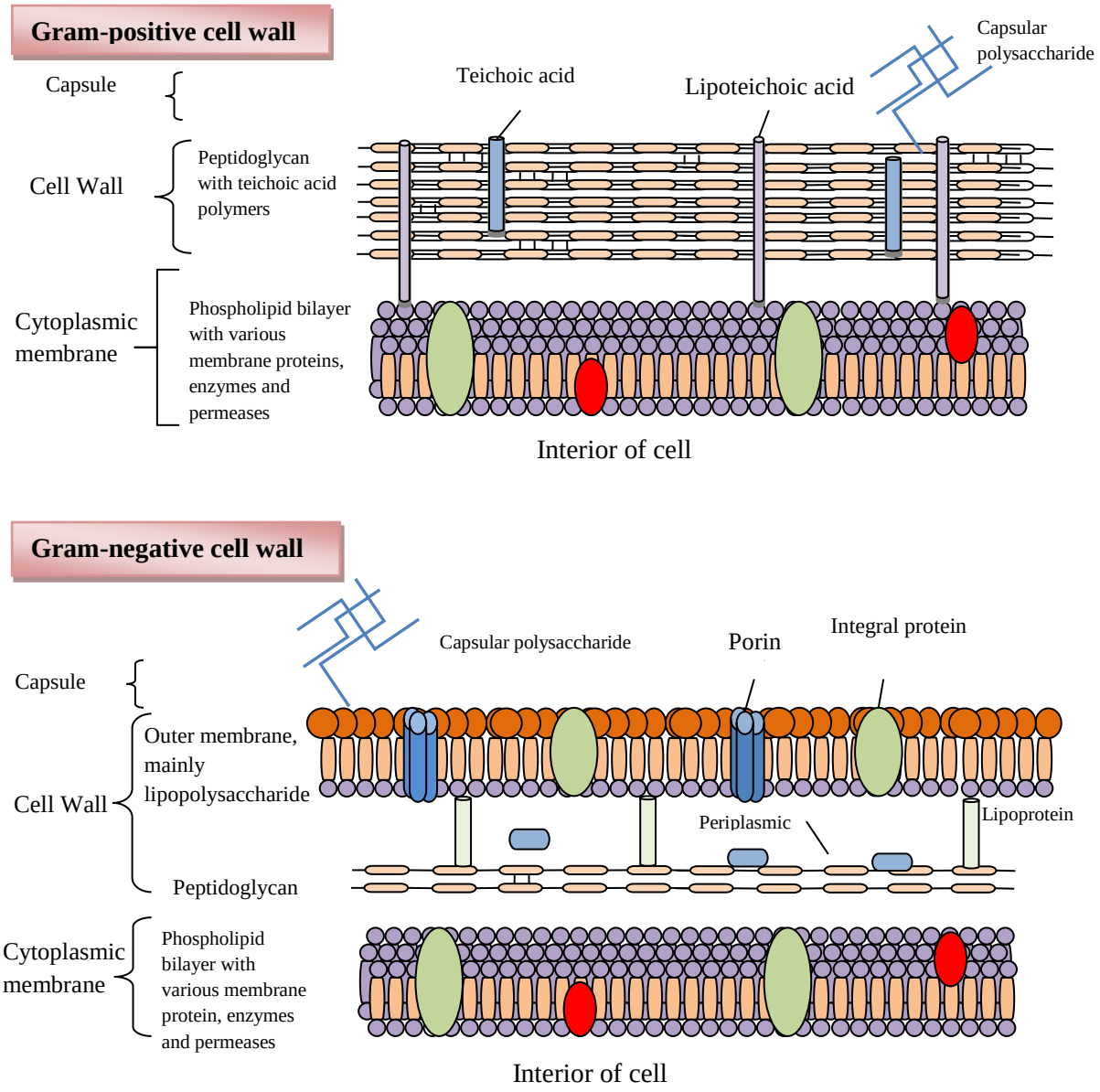


Figure 1.8. Schematic representation of the cell wall of Gram-positive and Gram-negative bacteria. Picture modified from Barer (2007).

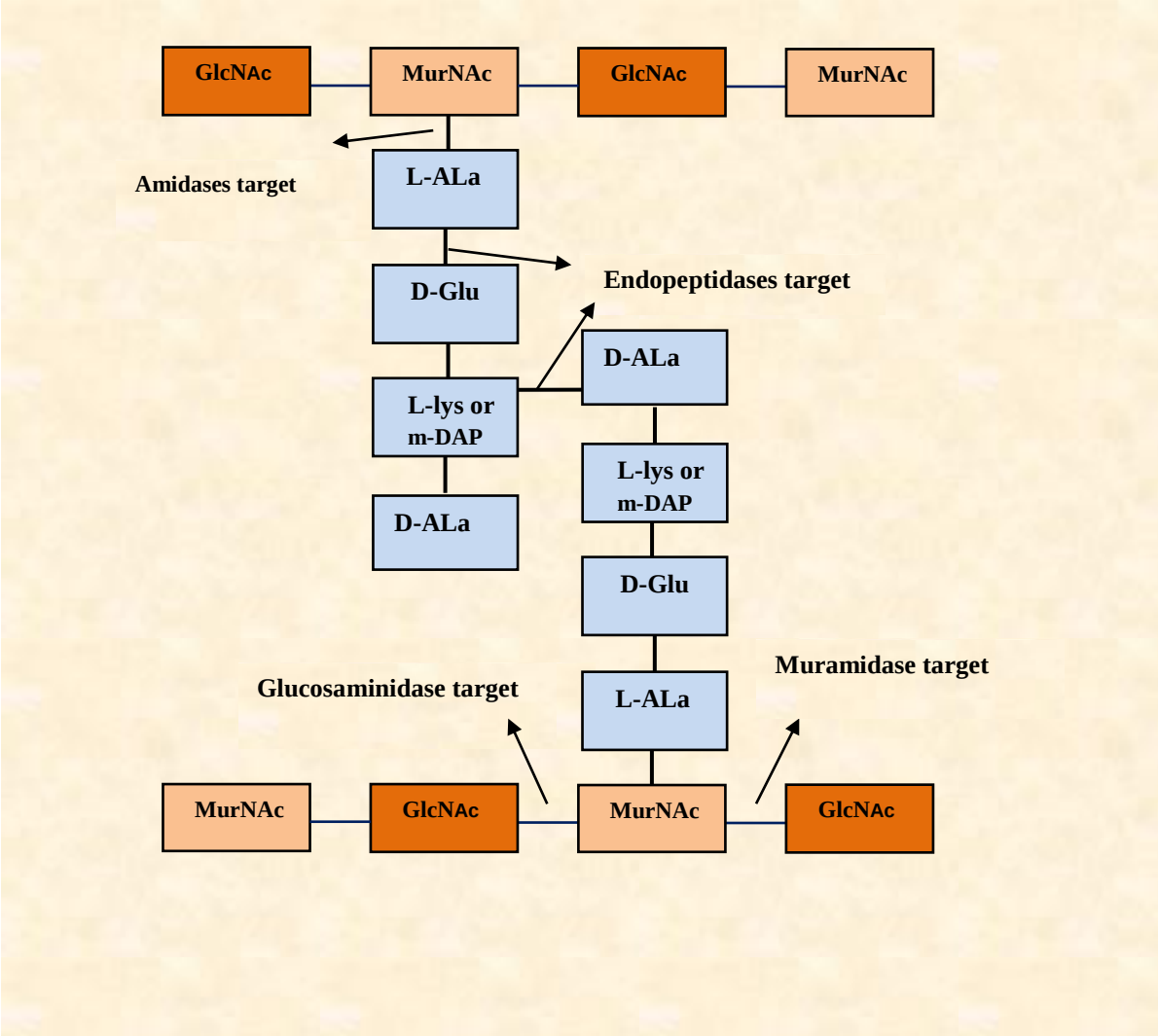


Figure 1.9. Bacterial cell wall peptidoglycan structure and sites of cleavage by various endolysins: N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc), L-alanine (L-Ala), D-glutamic acid (D-Glu), meso-diaminopimelic acid (m-DAP), L-lysine (L-Lys) and D-alanine (D-Ala). Picture is modified from Borysowski *et al*. (2006)

1.14.4 Therapeutic applications

Examples of the basic applications of endolysins as therapeutic agents include the elimination of bacterial colonisation of mucous membranes, for example the elimination of upper respiratory colonisation in mice by group *A. streptococci* (Nelson *et al.*, 2001) or *S. agalactiae*, which colonises the human genitourinary tract and lower gastrointestinal tract (Cheng *et al.*, 2005; Borysowski *et al.*, 2006).

In terms of the treatment of bacterial infections, the first attempt was made by using *B. anthracis* phage γPlyG amidases which were found to lyse *B. anthracis*, and a single strain of *B. cereus* that is closely related to *B. anthracis* (Schuch *et al.*, 2002). Another example of a medical lysin application is the treatment of pneumococcal bacteraemia in mice by Cpl-1 and Pal endolysins (Jado, 2003). In addition, a recent study showed that the lysin of bacteriophage CD27 has the ability to lyse *C. difficile* cells in vitro, also, the host range of the endolysins is significantly broader than that of the phage itself (Mayer *et al.*, 2008). Furthermore, lysins can be used for biocontrol of bacteria in food and feed, as well as in the protection of plants against phytopathogenic bacteria (Loessner, 2005).

There are several unique features that make endolysins very promising antibacterial agents, including their specificity for pathogens without affecting normal microbiota, relatively easy manufacturing and modification by genetic engineering, the low chance of developing resistance and the fact that they can be identified and isolated from both lytic and temperate phages (Borysowski *et al.*, 2006). However, endolysin therapy is not without disadvantages, especially since their actions have only been described against Gram-positive bacteria. Also, since they are proteins, they have short half-lives and are

susceptible to inactivation (Fischetti, 2008). Research into lysin dosage will be required in the future because they are not self-replicating (O'Flaherty *et al.,* 2009). Furthermore, it is unlikely that they can be used against intracellular bacteria that require entrance to the cells, since they are proteins (Fischetti, 2008).

1.15 Holins

In phage λ, which presents the best characterised example of a holin-endolysin system, the holin gene (*S*) and the endolysin gene (*R)* are located at the beginning of the late transcriptional unit (Young *et al.*, 2000). During late gene expression, holins are believed to accumulate in the cytoplasmic membrane as oligomers until a trigging event occurs that results in membrane lesions (Gründling *et al.*, 2001; Wang *et al.,* 2003).

Holins are small phage proteins. Most holin sequences are short (about 160-145 amino acid residues) with a highly hydrophilic carboxyl terminal domain (Young *et al.*, 2000). Holins are classified into three classes based on their membrane topology. Class I holins, such as the holin S of λ phage, (about 95 residues or longer) have three potential transmembrane domains (TMDs), while class II holins are usually small (about 65-95 residues) with two TMDs and class III holins have only one TMD (Young, 2002).

Holins can be regulated by inhibitors, known as antiholin, which usually transcribed from the same reading frame as holin (Loessner, 2005). After lysis, antiholin which exist in the

cytoplasm (as in T4) or bound to the inner membrane (as with λ) or); binds and inactivates the holin, producing an inactive holin-antiholin dimer (Young *et al.*, 2000).

In all phages tested to date, endolysin needs holin, except in some phages of Gram-positive bacteria where the endolysins use the host *sec* system to achieve their goals. For example, the endolysin of fOg44, a phage of the Gram-positive bacteria *Oenococcus oeni*, causes lysis of the host without a holin when endolysin undergoes leader peptidase cleavage of its signal sequence (Sao-Jose *et al.*, 2000). The endolysin of *Lactobacillus plantarum* phage φg1e (Kakikawa *et al.*, 2002) works in a similar fashion. Furthermore, some phage sequences have revealed that no holin encoding genes are located upstream of endolysin genes; therefore, these lysins may pass through the cytoplasmic membrane via a holin-less mechanism, such as in the *B. cereus* phages Bastille, TP21 and 12826 (Loessner *et al.*, 1997 ; Borysowski *et al*., 2006) and the *Listeria monocytogenes* phage A511 (Loessner *et al.*, 1995 ; Borysowski *et al*., 2006).

1.6 The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) System

1.16.1 Introduction

Various strategies have been developed by different microbial species to tackle exposure to different viruses and other foreign DNA and to allow cells to differentiate between the host genome and foreign genomes. These systems maintain genetic integrity and allow foreign DNA to be taken and used in ways that will help the organism to adapt to the environment and gain resistance to these elements (Horvath and Barrangou, 2010). Strategies such as prevention of adsorption, blocking of injection and aborting infection are effective against viruses, and restriction modification systems (R-M) and the use of sugar non-specific nucleases are also effective defence systems against invading nucleic acids (Karginov and Hannon, 2010).

An adaptive microbial immune system has recently been described that provides acquired immunity against viruses and plasmids. This system is known as the clustered regularly interspaced short palindromic repeats (CRISPR) system (Barrangou *et al.,* 2007). These repeats are a diverse family of DNA repeat sequences found in the DNA of many Bacteria (about 40%) and Archaea (about 90%) (Kunin *et al.,* 2007). The first description was provided in 1987 by Ishino and colleagues, who found 14 repeats of 29 base pairs (bp), interspersed by 32-33 bp non-repeating spacer sequences, in *E. coli*. Subsequently, similar CRISPR arrays have been found in *Mycobacterium tuberculosis* (Hermans *et al.,* 1991) and other bacteria. Mojica and colleagues were the first group to

describe the presence of short regularly spaced repeats in prokaryotes (Mojica *et al.*, 2000), which was subsequently incorporated into the CRISPR name (Jansen *et al.*, 2002).

1.16.2 Structure

The CRISPR system is formed from CRISPR arrays and CRISPR-associated sequences (CAS). As can be seen in figure 1.10, the CRISPR system consists of repeats, spacers, a leader and CAS genes.

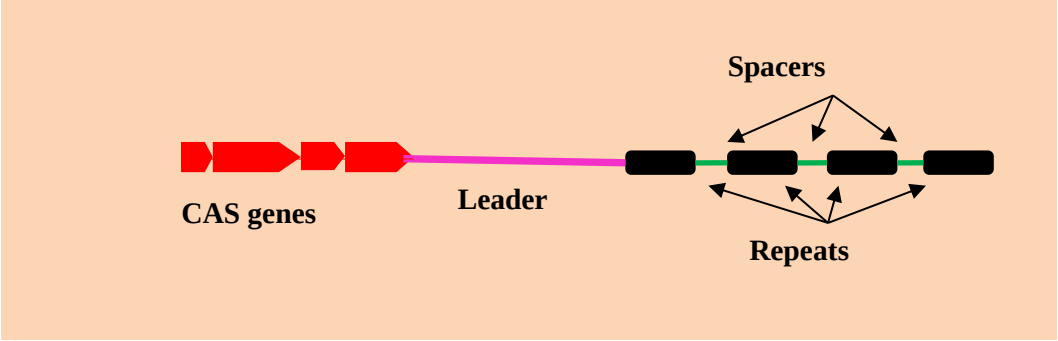


Figure 1.10. Typical structure of a CRISPR locus. Image modified from Sorek *et al.* (2008).

Repeats

The size of a repeat can vary between 24 and 47 bp. The number of repeats per array can vary from 2 to 249. Within a single array, the repeat sequences are highly conserved in terms of size and sequence (Jansen *et al.*, 2002).

Spacers

Spacers usually have a size between 26 and 72 bp. Mojica *et al.* (2005) showed that CRISPR spacer sequences derive either from mobile genetic elements such as conjugative plasmids and bacteriophages or chromosomal sequence

Leader

The leader is the sequence that is located between the repeats and CAS genes. This sequence is usually up to 550 bp long (Jansen *et al.,* 2002). The leader sequence seems to be conserved between several CRISPR loci in the same chromosome; however, it is not conserved between species (Sorek *et al.,* 2008). The leader could function as a recognition sequence since a new repeat-spacer unit is always added between the leader and the previous unit in a CRISPR array (Sorek *et al.,* 2008) and likely acts as the promoter for the transcription of CRISPER array since it is found directly upstream of the first repeat (Lillestøl *et al.,* 2006; van der Oost *et al.,* 2009).

CRISPR-Associated Sequence (CAS) Genes

These are a large set of gene families that are always found adjacent to the leader sequence. Six "core" CAS genes have been identified, including cas1, which is regarded as a universal marker of the CRISPR/CAS system. Furthermore, six or fewer CAS genes are usually present in a CRISPR system, but some may have up to 20 CAS genes (Haft *et al.,* 2005; Sorek *et al.,* 2008). CAS genes encode for proteins that carry nucleic acid-related domains such as nucleases and helicases, and are involved in nucleotide binding (Barrangou and Horvath, 2009).

In general, single CRISPR loci are usually present in one genome; however, some organisms may have more than one, such as *Methanococcus jannaschii* which has 18 loci (Bult *et al.,* 1996; Sorek *et al.,* 2008).

Mobile genetic elements, such as plasmids, *skin* mobile elements and prophages, have been found to carry CRISPR and CAS elements, which may suggest the role of these elements in the distribution of this mechanism (Godde and Bickerton, 2006). Regarding the roles of CRISPRs, several hypotheses have been suggested, such as the role of CRISPRs in DNA repair, since DNA manipulating domains are found in many CAS gene contents (Makarova *et al.,* 2002). It has been suggested that the CRISPR-CAS system mediates immunity against foreign genomes by means of an RNA interference-like mechanism, since there are many similarities between the spacer and extrachromosomal DNA, such as phage or plasmid DNA (Mojica *et al.,* 2005; Pourcel *et al.,* 2005 ; Bolotin *et al.,* 2005; Sorek *et al.,* 2008).

Barrangou and colleagues (2007) recently showed that CRISPRs provide acquired resistance against viruses in prokaryotes, after showing that new spacers that were acquired after phage infection give resistance against the phage (Barrangou *et al.,* 2007; Kunin *et al.,* 2007). Furthermore, in 2008, Marraffini and Sontheimer showed a CRISPR system that prevented conjugation and plasmid transformation in *S. spidermidis* (Marraffini and Sontheimer, 2008).

A CRISPR-CAS mechanism that provides immunity against foreign genetic elements after acquisition of the novel spacer that is derived from these elements is not fully characterised. As can be seen in figure 1.11, in a sensitive cell following plasmid

transformation or infection by a phage, exogenous DNA proliferates in the cell and either new phage particles are produced, leading to the death of the bacterial cell, or there is production of a new cell that carries the plasmid DNA. However, bacterial cells become resistant after integration of a novel spacer unit, from the exogenous DNA, at the leader end of the CRISPR locus. In this immunisation process, the repeat-spacer array is transcribed into a poly-spacer precursor crRNA (pre-crRNA), then pre-crRNA sequences are recognised and processed by the CAS complex to produce mono-spacer crRNAs which contain a spacer and two half repeats. Mature crRNA is used by the CAS complex to base pair with invading nucleic acids, leading to their inactivation and degradation (Sorek *et al.*, 2008; Horvath and Barrangou, 2010).

It has been reported recently that the proto-spacer adjacent motif (PAM) sequence plays a role in CRISPR-encoded resistance. Previously, this motif was proposed as playing a role in the recognition or cleavage of a phage sequence by the CRISPR system (Sorek *et al.*, 2008). This motif sequence is a short sequence found in the phage genome located downstream of the proto-spacer. A mutation in either the proto-spacer or the adjacent motif allows a phage to bypass the CRISPR-CAS defence mechanism, which strongly indicates the involvement of the motif in resistance (Deveau *et al.*, 2008).

CRISPR loci provide useful applications that include strain typing, such as the Spoligotyping method (Kamerbeek *et al.*, 1997), and engineering of industrial bacteria for resistance to phage infection, by inserting a known sequence from a phage into a CRISPR array as a spacer (Sturino and Klaenhammer, 2006; Sorek *et al.*, 2008). However, phages can circumvent CRISPR-encoded immunity and continue to infect host

cells. This happens when the phage develops mutations either in the proto-spacer (mutation, insertion or deletion) or in the CRISPR motif (point mutation), or there is large-scale deletion of both the proto-spacer and CRISPR motif (Barrangou and Horvath, 2009).

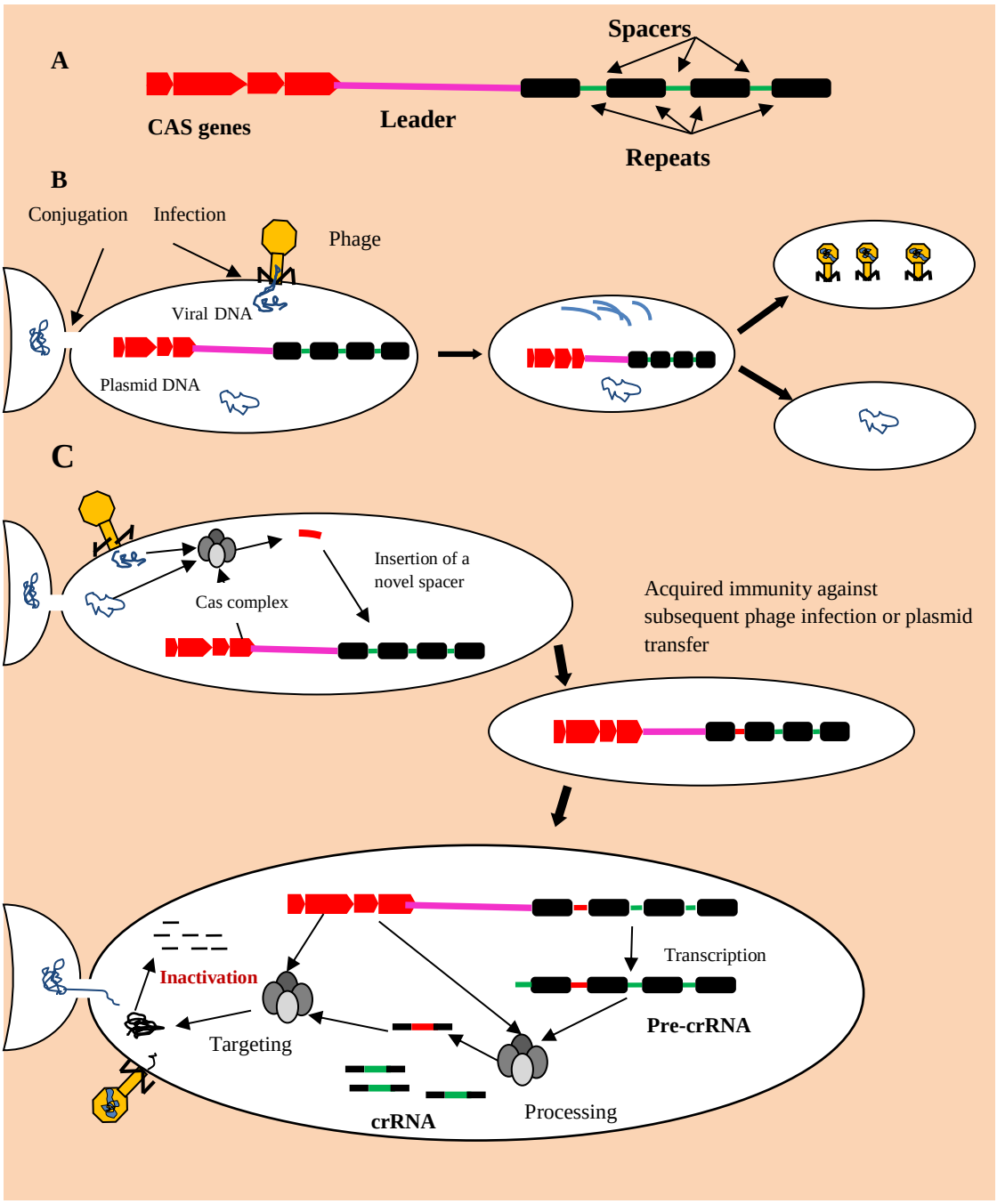


Figure 1.11. Putative model for CRISPR action ; CRISPR structure (A), normal mechanism without the CRISPR system (B) and an overview of the CRISPR/CAS mechanism of action (C). Diagram is modified from Sorek *et al.*, (2008) and Horvath and Barrangou (2010).

1.16.3 CRISPR sequences in the *C. difficile*

Sebaihia *et al.* (2006) identified a number of CRISPR sequences in the *C. difficile* 630 genome. Table 1.3 presents these sequences and their positions. Furthermore, some of these sequences were also found within a *C. difficile* 630 prophage, which may explain the role of phages in using CRISPR to prevent super-infection by other phages (Sebaihia *et al.*, 2006). In addition, this may explain the rarity of *C. difficile* as a strain that is sensitive to the temperate phages that have been isolated so far (Mayer *et al.*, 2008). Interestingly, the sequences from different CRISPR regions have homology with some sequences from *C. difficile* φC2, φCD119 and φCD27, as shown in table 1.3.

Table 1.3. The different CRISPR sequences in *C. difficile* 630

CRISPR	Coordinates	Number of repeats	Genomic location	*Homology with *C. difficile* phages
CRISPR1	1124036-1125688	9	Prophage 1	φC2 & ΦCD119
CRISPR2	1433859-1435639	14	Mobile element	φCD119
CRISPR3	1644671-1645952	12	Mobile element	φC2
CRISPR4	1756116-1756833	9	Chromosome	φCD119
CRISPR5	1935320-1936495	16	Chromosome	φCD27
CRISPR6	2298609-2298929	3	Chromosome	
CRISPR7	2662916-2665492	15	Mobile element	φC2, φCD119 & φCD27
CRISPR8	2907464-2908427	7	Mobile element	φC2 & φCD119
CRISPR9	3396953-3398605	10	Prophage 2	
CRISPR10	3455612-3456944	18	Chromosome	φCD27

*all the information in the table was taken from Sebaihia *et al.* (2006), except for this column

1.17 Aims and objectives

The aims of this research are as follows:

- To isolate lytic phages which could be used for the treatment of *C. difficile*
- To isolate phage lysins which could be used for the treatment of *C. difficile*
- To understand the role of phages in the evolution of *C. difficile*
- To isolate the spore lysins that can target *C. difficile* spores

Chapter Two

General materials and methods

2.1 Microbiological materials and methods

2.1.1 Growth and sporulation media

Cycloserine-Cefoxitin-Fructose Agar (CCFA), *Clostridium difficile* Moxalactam Norfloxacin (CDMN) agar and Brain Heart Infusion (BHI) media were purchased from Oxoid (Basingstoke, UK). Luria-Bertani (LB) agar and broth were purchased from Fisher Scientific (Leicestershire, UK), while Listeria selective agar was obtained from Sigma-Aldrich (Gillingham, UK). Culture media were prepared according to the manufacturer's instructions and sterilised by autoclaving at 121°C for 15 minutes (Prestige Medical, UK). To make up CCFA and CDMN agar plates, the media was cooled to 50°C prior to the addition of *Clostridium difficile* supplement antibiotic (Oxoid, Basingstoke, UK) and 7% (v/v) Defibrinated Horse Blood (Oxoid, Basingstoke, UK). Prior to culture, agar plates and broth were incubated at 37°C for 24 hours in an anaerobic cabinet (Ruskinn Technology Ltd) containing an anaerobic environment, consisting of 85% N_2, 10% CO_2, and 5% H_2 to facilitate the removal of dissolved oxygen. Inoculated cultures were incubated at 37°C for 48 hours in an anaerobic workstation (Ruskinn Technology Ltd).

2.1.2 Bacterial strains

Ninety four isolates of *Clostridium difficile*, eleven isolates of *Clostridium perfringens* and individual strains of diverse clostridia species and aerobic strains were kindly donated by the Anaerobe Reference Laboratory, University Hospital of Wales, Cardiff.

National Collection of Type Culture (NCTC) strains were purchased from the Health Protection Agency Culture Collection (Health Protection Agency, Colindale). *E. coli* recombinant protein host strains M15[pREP4] and SG13009[pREP4] were supplied by QIAexpress Type IV Kit (Qiagen Ltd, Crawley, UK). Details of these strains are noted in table 2.1 and 2.2.

Table 2.1: The *C. difficile* isolates used in this study (Appendix 1)

C. difficile ribotypes	Strain
001	DS1759 - DS1747- R8652- DS1750- R17752
027	DS1813- DS1801- R20291 -DS1807-R24626
106	R10459 -DS1798 -DS1787- DS1771 -R28614
014	DS1742 -R30061- R22537 -R31755-R31757
002	DS1748 -R25577 -R27039 -R31760-R31763
005	DS1721 -R27038 -R25028- R31762- R31778
012	DS1752-R31771-R31080-R30904-NCTC 13307 (CD630)
078	DS1723-R19058-R31710-R31708-R31777
020	DS1724-R31774-R31550-R31464-R31502
010	DS1684-R31750-R31722-R30762-R30327
023	DS1665- R15552-R31377-R31148-R31314
017	R9557 -R13695 –R18091- R2139- R19222
047	R10542- R18045-R25961-R17732-R10543
056	R24498- DS2008 -R26796-R31312-R31229
110	R7771- R17978-R30967-R18040-R17985
046	R19168-R31583-R31263-R30869-R30870
045	R20408-R30776-R28450-R30072-R27086
NCTC (001)	NCTC 11204 - NCTC 12727- NCTC 11209

Table 2.2: The bacterial isolates used in this study

Bacterial species	Strain
C. perfringens	**R31707,R31837,R31599,R31737,R31745,R31885,R31973,R31782,R 31936, R31972- NCTC 8359 (ATCC 12915)**
Other *Clostridium* strains	***Clostridium sordellii* R20453, *Clostridium septicum* R22030, *Clostridium novyi* R14479**
Aerobic species	***Listeria monocytogenes* (NCTC 5412), *Micrococcus luteus* NCTC 2665 (ATCC 4698), *Bacillus subtilis* NCTC 10400 (ATCC 6633), *Staphylococcus aureus* NCIMB 9518 (NCTC 10788) , *E. coli* strains M15(pREP4) and SG13009 (pREP4)**

2.1.3 Culture and growth conditions of anaerobic bacterial strains

Individual *C. difficile* isolates were recovered from spore stocks and introduced to pre-reduced BHI and CDMN agar plates using the direct streak method, in which 10 µl of spore suspension was streaked across the surface of an agar plate using a sterile loop (Fisher Scientific, Leicestershire, UK). The inoculated plate was incubated anaerobically for 48 hours at 37°C.

Freeze-dried NCTC bacterial strains were recovered using the method recommended by NCTC (How to Handle Bacteria and Fungi on Receipt at http://hpacultures.org.uk/technical/Howto Handle Bacteria.jsp; accessed on 05-2009). Between 0.3 and 0.4 ml of pre-reduced BHI broth was added aseptically to the freeze-dried material, using a Pasteur pipette, and mixed. The mixture was transferred to a 5 ml pre-reduced BHI broth and an aliquot was subsequently spread across a BHI agar plate. All cultures were incubated at 37°C anaerobically for 48 hours.

To recover cultures of non-*difficile* anaerobic strains, protect beads were passed over the surface of a pre-reduced BHI agar (enriched with 5% w/v Defibrinated horse blood). Plates were incubated anaerobically for 24-48h at 37°C.

All anaerobic cultures were performed in a gloveport Bugbox Plus anaerobic workstation (Ruskinn Technology Ltd). The anaerobic environment consisted of 85% N_2, 10% CO_2, and 5% H_2. Disposable anaerobic indicator strips (Oxoid, UK) were replaced daily to ensure that a strict anaerobic environment was maintained. The temperature of the workstation was set at 37°C.

2.1.4 Culture and growth conditions of aerobic bacterial strains

E. coli host strains M15 and SG13009 were grown on Luria-Bertani (LB) agar supplemented with ampicillin (100 μg/ml) at 37°C for 24h; subsequently distinct colonies were re-suspended and stored frozen at -80°C using Protect™ vials (Technical Service Consultants, Heywood, UK)

Transformed *E. coli* M15 and SG3009, which was used throughout as a cloning host, were routinely cultured using (LB) broth or agar, as appropriate, at 37°C, with 300 rpm shaking for liquid cultures. Media was supplemented as appropriate with kanamycin (25 μg/ml) and ampicillin (100 μg/ml) (Sigma-Aldrich, Gillingham, UK). *Listeria monocytogenes* **(**NCTC 5412) was cultured aerobically on Listeria selective agar at 37°C for 48h. *Micrococcus luteus* (NCTC 2665), *Staphylococcus aureus* (NCTC 10788) and *Bacillus subtilis* (NCTC 10400) were grown in BHI agar aerobically at 37°C for 24 h.

2.1.5 Freezer cultures

Freezer cultures of all strains were prepared as follows: 850 µl of a 48 h. broth culture was added to 1.5 ml cryogenic vials (Fisher Scientific, Leicestershire, UK) along with 150 µl of glycerol. The vials were vortexed and stored at -80°C.

Freezer cultures of all strains were also prepared by using Protect™ vials containing cryopreservative fluid and porous beads. One ml of the 48 h broth culture was added to each of the vials, which were vigorously shaken; the residual liquid was removed using a sterile Pasteur pipette, after which the vials were stored at -80°C.

2.1.6 *C. difficile* spore production

A single colony from a 48 h BHI agar plate culture was sub-cultured into 45 ml of pre-reduced BHI broth (50ml Falcon tube) and incubated anaerobically for 10 days at 37°C. At the end of this period, spores were extracted using two modified methods.

In the first method, the broth was centrifuged at 4,000xg for 15 minutes at room temperature (Heraeus Primo, Thermo Scientific, UK). The supernatant was discarded and the pellet was washed twice with 10 ml of chilled sterile deionised water (SDW). The washed pellet was re-suspended in 1 ml of SDW and heated for 10 minutes at 80°C in a water bath (Fisher Scientific, Leicestershire, UK) to inactivate any vegetative cells. The spore suspension was then centrifuged at 4,000xg for a further 5 minutes at room temperature and finally re-suspended in 1 ml of SDW and stored at 4°C until needed.

To confirm purity, spore suspensions were cultured on the pre-reduced BHI plate to assess purity. This method was modified from that described by Perez *et al.* (2005).

The second spore preparation method is based on a technique for purifying *C. difficile* 630 spores that was originally described by Lawley *et al*. (2009). The spores were harvested by centrifugation at 5,000xg for 15 minutes (Heraeus Primo, Thermo Scientific, UK) and were re-suspended in 10 ml of sterile water. The samples were then washed four to six times in water, during which the supernatant became clear. The samples were then pelleted and re-suspended in 30 ml of PBS (Fisher Scientific, Leicestershire, UK) prior to sonication (four times) with second 30 bursts at an amplitude of 12 using a Soniprep 150 ultrasonic disintegrator (Sanyo, Leicester, UK). Following sonication, 3 ml of 10% Sarkosyl (Fisher Scientific, Leicestershire, UK) was added to the sample, which was then incubated at 37°C for 1 hour. The samples were then pelleted by centrifugation at 4,000xg for 10 minutes, after which the pellet was re-suspended in 10 ml of PBS containing 0.125 M Tris buffer pH 8 (Fisher Scientific, Leicestershire, UK) and 10 mg/ml lysozyme (Roche, Germany) before being incubated overnight at 37°C. The samples were then sonicated as described above, and 3ml of 1% Sarkosyl was added to the sample prior to incubation for 1 hour at 37°C. The samples were then layered on to 50% sucrose (Fisher Scientific, Leicestershire, UK) and centrifuged for 20 minutes at 4,000xg. The pellets were re-suspended in 2 ml of PBS containing 200mM EDTA, 300 ng/ml Proteinase K (recombinant PCR grade; Roche, Germany), and 3 ml of 1% Sarkosyl, before being incubated for 1 to 2 hours at 37°C until the samples cleared. The samples were then layered on to sucrose as described above.

The pellets were re-suspended in water and were washed twice prior to re-suspension in sterile water. The resulting spore suspensions were plated on the pre-reduced BHI plate to assess purity, and then stored at 4°C until needed.

2.1.7 Viable count for spores and bacterial cells

The spore/bacterial cell suspensions were diluted by adding 100 µl of the stock suspension to 900 µl of BHI broth, creating a 10^{-1} dilution. Five further 10 – fold serial dilutions were made i.e. to a 10^{-6} dilution. The dilutions were then used to perform viable counts using a drop count method (Miles and Misra, 1938), in which three 10 µl drops of each dilution were applied to pre-reduced BHI agars, three plates being prepared for each dilution. The plates were left to dry and then inverted and incubated anaerobically at 37°C for 48 hours. The colony forming units (CFUs) present in dilution, which contained between 3 and 30 CFUs, were counted after 48 hours and used to calculate the number of CFU/ml, using the following formulae:

Mean number of colonies x dilution factor = CFU/10µl

CFU/10µl x 100 = CFU/ml

This technique was validated by performing 10 separate dilution series, as described above, and statistically analysing the results using a one-way ANOVA in the Minitab Statistical Software Package (release 14 software, Minitab Inc., PA, USA) to ensure that there were no statistically significant differences in the results. Validation was deemed successful when the *p-value* ≥ 0.05.

2.1.8 Identification of *C.difficle*

2.1.8.1 Colony morphology

On media colonies are typically cream in colour, irregular, lobate or rhizoid edges, a fried egg appearance can be seen due to spore formation (Figure 2.1). Furthermore, the horse manure odour that is associated with *C. difficile* colonies is very distinctive and may give a quick and easy aid to identification. In addition, the ability of colonies to give a golden-yellow fluorescence under long-wave ultraviolet light is a very useful and recommended approach for *C. difficile* detection (Aspinall and Hutchinson, 1992 ; Brazier and Borriello, 2000).

Figure 2.1 shows the colony morphology of *C.difficile* NCTC 12727 on BHI agar and CDMN agar plates after 48 hours incubation

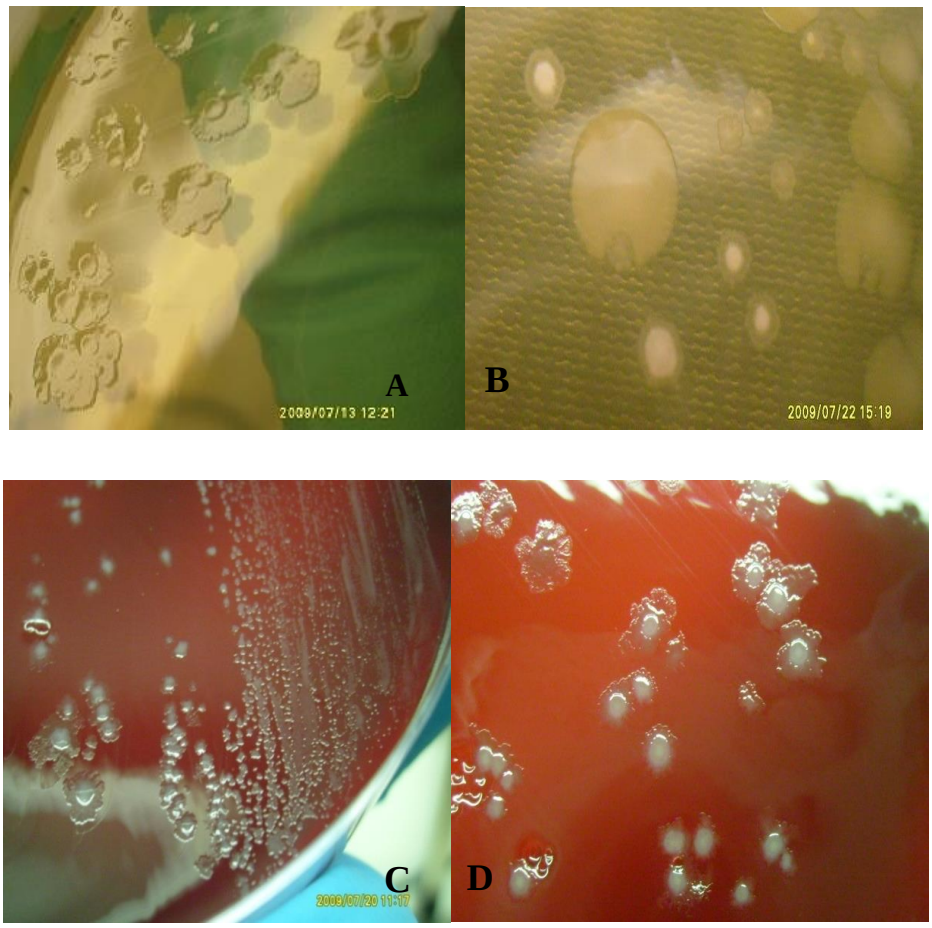


Figure 2.1. Colony morphology of *C.difficile* NCTC 12727 on BHI agar (A,B) and CDMN (C,D) agar plates, incubated anaerobically for 48 hours at 37°C

2.1.8 .2 Gram stain

All Gram stain dyes were supplied by Fisher Scientific, Leicestershire, UK.

The Gram stain technique, which was developed by Hans Christian Gram in 1884, is used to classify bacteria on the basis of their forms, sizes, cellular morphologies, and Gram reactions (Winn *et al.*, 2006).

A drop of suspended culture was transferred on to a glass slide using an inoculation loop. If the culture was from a Petri dish, a minute amount of a colony was smeared with a drop of water and allowed to air dry; the slide was then flooded with crystal violet for 1 minute and washed with distilled water. Gram's iodine was then added for 1 minute and the slide was washed again with distilled water. The slide was then decolorised with 95% ethanol for 10 seconds and washed with distilled water. Finally the slide was counterstained with safranin for 1 minute and rinsed with water (Winn *et al.*, 2006). To remove excess fluid, the slides were blotted with a disposable towel and examined under the 100x oil immersion objective of an Olympus DX-50 microscope, the images being captured on an Olympus DP10 camera.

C. difficile is a Gram-positive bacillus with rounded or sometimes pointed ends, as can be seen in figure 2.2.

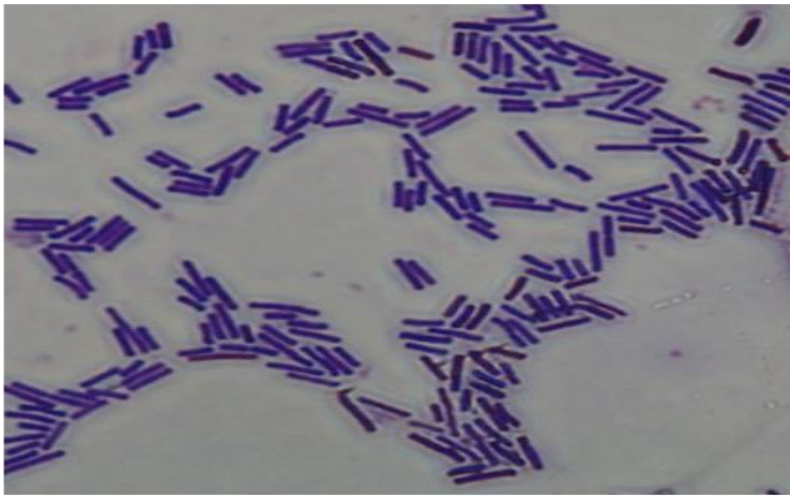


Figure 2.2. Gram stain of *C. difficile* showing Gram-positive bacilli often arranged in pairs or short chains

2.1.8 .3 Spore stain

Spore stain solutions were supplied by Fisher Scientific, Leicestershire, UK.

A one colony (~ 2-3mm) was smeared onto a microscope slide with a drop of water and allowed to air dry, and the slides were then flooded with malachite green. The slides were heated from below with a Bunsen flame until steam was seen to rise from the stain, then the slides were left for 1 minute before being washed with distilled water and counterstained with safranin dye for 15-30 seconds (Schaeffer and Fulton, 1933) and washed with distilled water. The slides were blotted dry and examined under the 100x oil immersion objective of an Olympus DX-50 microscope, with the images being captured on an Olympus DP10 camera.

The Spore stain of *C. difficile* illustrates the brownish red vegetative cells with a few interspersed green spores (Figure 2.4). Spores are usually detected after 48 hours of incubation.

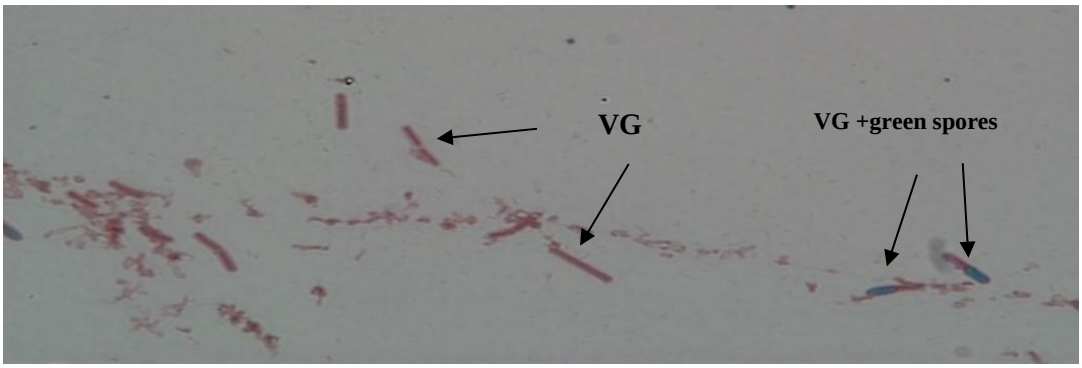


Figure 2.3. Spore stain of *C. difficile* illustrating the brownish red vegetative cells with a few interspersed green spores

2.1.8.5 Latex agglutination

The latex agglutination test is rapid and easy to perform, but the test detects only bacterial antigens such as glutamate dehydrogenase (GDH) by using latex particles coated with antibody. However, this test detects both toxigenic and nontoxigenic strains of *C. difficile* (DiPersio *et al.*, 1991). In this study latex agglutination test (Oxoid, Hampshire, UK) were used.

2.2 Molecular biological materials and methods

2.2.1 Identification of restriction enzyme sites in primary gene sequence

The NEBcutter v2.0 program (http://tools.neb.com/NEBcutter2) was used to determine the presence of restriction enzyme sites in a particular nucleotide sequence. The gene sequences of all of the genes which were cloned and expressed as recombinant proteins were analysed against all commercially available restriction enzymes (single, double and triple cutter enzymes). Restriction enzyme sites which were also present in the cloning vector pQE30 were identified and removed. To facilitate cloning into the vector, BamH1 (GGATCC) and SalI (GTCGAC) sites were incorporated into the 5' and 3' primer ends of the gene insert

2.2.2 Codon optimisation

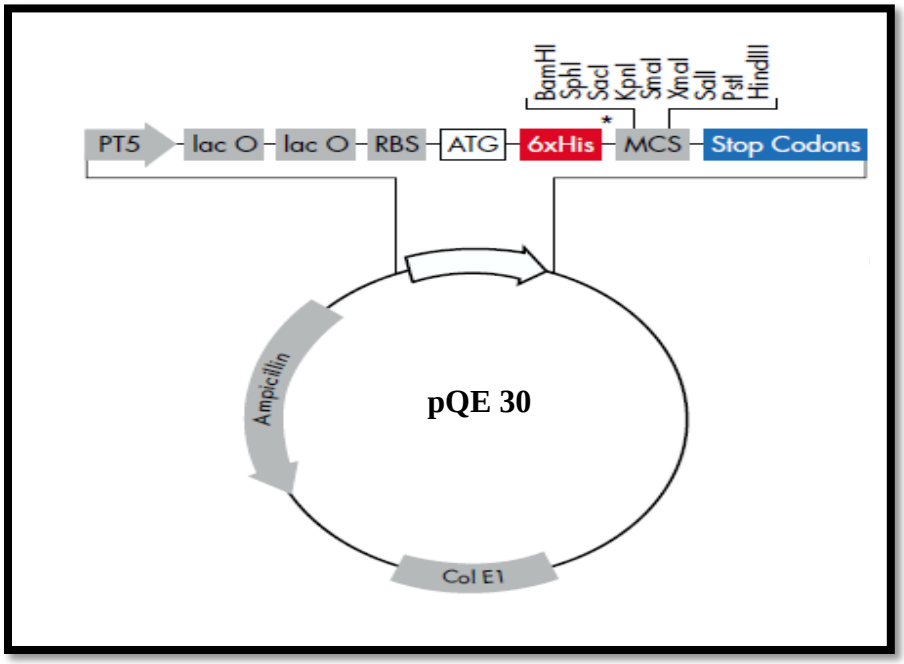
The genetic code employs 61 nucleotide triplets (codons) to represent 20 amino acids and three codons for translation termination, meaning that a single amino acid can be encoded by more than one codon, with alternate codons for the same amino acid being referred to as synonyms or synonymous (Swartz, 2001; Gustafsson *et al.*, 2004). Synonymous codons frequently vary at the third codon position by one nucleotide, and a subset of the cell's transfer ribonucleotide acid molecules (tRNA) are able to recognise each of these codons.

Indeed, according to Zdanovsky and Zdanovskaia (2000), many clostridial proteins are poorly expressed in *E. coli* because the codons that are commonly used in clostridial coding sequences are rarely used in *E. coli,* and the quantities of the corresponding tRNAs in *E. coli* are not sufficient to support efficient translation.

Fortunately, with modern molecular biology methods it is now possible to modify the codon composition of a gene so that it more closely matches that of the host bacteria, resulting in an increased level of expression of the protein of interest (Makoff *et al*., 1989).

To enhance expression from *E. coli* the nucleotide sequence of each endolysin codon was optimised using a commercial algorithm (Genscript Inc., USA). Optimised sequences were then chemically synthesised and cloned into the vector pUC57, after which they were sub-cloned into the QIAgen high level inducible protein expression system (pQE 30) using the BamHI and Sa1I sites so that expression of the insert was placed under the

control of the T5 promoter and so that each endolysin sequence was fused to an N-terminal histidine tag (Figure 2.4). The correct orientation and fidelity of each coding sequence following insertion was confirmed by DNA sequencing.



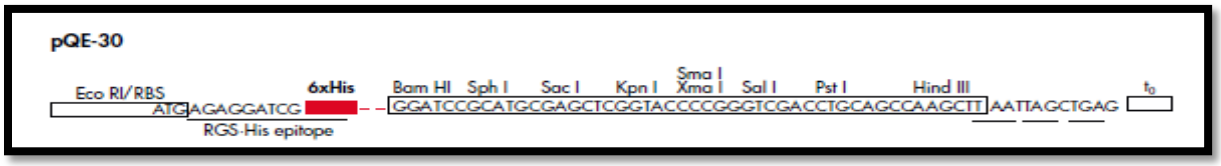


Figure 2.4. pQE 30 vectors for N-terminal 6xHis tag constructs

PT5: T5 promoter - Optimised promoter/operator element consists of the phage T5 promoter and two lac operator sequences, which increase the probability of lac repressor binding and ensure efficient repression of the powerful T5 promoter; **RBS**: Ribosome-binding site - Synthetic ribosomal binding site RBSII for efficient translation; **Stop Codons**: Translational stop codons in all three reading frames for convenient preparation of expression constructs; **Ampicillin resistance**: beta-lactamase gene (bla) confers ampicillin resistance; **6xHis** tag sequence: 6xHis-tag coding sequence either 5' or 3' to the poly-linker cloning region; **Col E1**: Col E1 origin of replication from pBR322 - Two strong transcriptional terminators from phage lambda, and T1from the rrnB operon of *E. coli*, to prevent read-through transcription and ensure stability of the expression construct; **lac O**: lac operator; **ATG**: Start codon; **MCS**: Multiple cloning site (Qiagen Ltd, Crawley, UK).

An aliquot of 4 μg of each construct was provided by Genscript in a lyophilised form which was stored at -20°C. To reconstitute this sample, the tube was centrifuged at 6,000xg for 1 minute (Eppendorf 5417R centrifuge, fixed angle rotor F45-30-11, UK). A 20 μl aliquot of sterilised distilled water was added and the vials were vortex mixed for 1 minute.

2.2.3 Preparation of competent cells

The QIA *express* system provides two strains of *E. coli*: M15, which supports high-level expression and is easy to handle, and SG13009 (Gottesman *et al.,* 1981), which can be used to express proteins that are poorly expressed in M15.

Both the M15 and SG13009 strains are derived from *E. coli* K12 (DH5α strain) and have the phenotype NaIS, StrS, RifS, Thi–, Lac–, Ara+, Gal+, Mtl–, F–, RecA+, Uvr+, Lon+. The *E. coli* host strains M15and SG13009 were grown on LB agar supplemented with ampicillin (100 μg/ml) at 37°C for 24h; subsequently, distinct colonies were re-suspended and stored frozen at -80°C in the Protect™ microbial preservation kit.

An aliquot of M15 or SG13009 cells from a Protect™ vial was taken with a sterile toothpick or inoculating loop and streaked out on an LB agar plate containing 25 μg/ml kanamycin and incubated at 37°C overnight, aerobically. A single colony of M15 or SG13009 was subsequently inoculated into 10 ml LB broth containing kanamycin (25 μg/ml) and grown overnight at 37°C and 300 rpm (Max 4450, Thermo Scientific, UK). One ml of this culture was then added to 100 ml of pre-warmed LB-kanamycin and incubated at 37°C and 300 rpm until an absorbance of OD_{600} 0.5 (Ultra spec 3100 pro, GE Healthcare) was achieved.

The culture was cooled on ice for 5 minutes prior to centrifugation at 4,000xg for 5 minutes at 4°C (Heraus Primo R, Thermo Scientific, UK). The cells were gently re-suspended in cold TFB1 buffer (100 mM RbCl, 50 mM $MnCl_2$, 30 mM potassium acetate, 10 mM $CaCl_2$, 15% glycerol, pH 5.8, sterile filter), and placed on ice for 90 minutes. The cells were centrifuged as above, the supernatant was discarded, and then the pellet was gently re-suspended in ice-cold TFB2 (10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$, 15% glycerol, adjusted to pH 6.8 with KOH, sterile filter) buffer. Aliquots of 100–200 µl were prepared in sterile microcentrifuge tubes, frozen in liquid nitrogen and stored at -80°C.

2.2.4 Transformation of competent *E. coli* cells

Chemically competent cells of *E. coli* M15 and SG13009 were thawed on ice for approximately 5 min, after which 100 µl of cells were gently transferred into a cold microcentrifuge tube. An aliquot containing 10 ng of the pQE30 plasmid construct was then added to the cells and mixed, after which the tubes were incubated on ice for 20 minutes.

Following incubation, cells were heat shocked for 90 s in a water bath (Fisher Scientific, Leicestershire, UK) at exactly 42°C, after which they were placed on ice for 2–3 minutes. Five hundred microliters of ice-cold SOC medium (Super Optimal broth with Catabolite repression) (2% tryptone, 0.5% yeast extract, 0.05 % NACL, 250mM KCl, 2M $MgCl_2$, and 1M sterile glucose) was added to the tube, which was then incubated at 37°C for 90 min, with shaking at 300 rpm (Max 4450, Thermo Scientific, UK). Following incubation,

5, 10, 50 and 100 µl aliquots of the culture were plated out on LB-agar plates containing 25 µg/ml kanamycin and 100 µg/ml ampicillin. Following inoculation, the plates were incubated overnight at 37°C, after which the number of positive transformations were enumerated and individual colonies were picked off and stored at −80°C using Protect™ storage vials.

Transformation efficiency was compared to that achieved when competent cells were transformed with 1 ng of pQE-40 (Qiagen Ltd, Crawley, UK) control plasmid in 20 µl of Tris-EDTA buffer (pH 8.0) (Sigma-Aldrich, Gillingham, UK). Transformation efficiency was determined using the following calculation:

Transformation efficiency = Total number of cells growing on the agar plate /
Amount of DNA spread on the agar plate (in µg).

2.2.5 Recombinant protein expression

An individual protect bead containing a specific *E. coli* recombinant expression system was recovered from -80°C storage and inoculated into 10ml of LB broth containing 25 µg/ml kanamycin and 100 µg/ml ampicillin, and was incubated for 14 h in an orbital shaker at 37°C and 300 rpm (Max 4450, Thermo Scientific, UK). Following incubation, 5ml of this culture was transferred into a further 100 ml of pre-warmed antibiotic supplemented LB broth (100 µg/ml ampicillin and 25 µg/ml kanamycin). The resulting culture was incubated at 37°C and 300 rpm until it achieved an OD_{600} of 0.6. At this point, recombinant protein expression was induced by the addition of 100 µl of 1mM

Isopropyl β-D-1-thiogalactopyranoside (IPTG) (Invitrogen, Paisley, UK). Following the addition of IPTG, cultures were incubated with shaking (300 rpm) for a further 4 h at 27°C to reduce the formation of inclusion bodies.

To optimise recombinant protein induction conditions for each recombinant protein, 1 ml aliquots of cultures were collected prior to induction, at the commencement of induction and each hour post-induction until the end of the induction period (4 h). As a control, samples were collected from cultures to which IPTG had not been added. Samples were harvested by centrifugation at 15,000 g for 3 min at 4°C (Eppendorf 5417R centrifuge, fixed angle rotor F45-30-11, UK), after which recombinant protein was visualised on SDS-PAGE using a ChemiDoc XRS equipped with a 12-bit CCD (charge-coupled device) camera (Bio-Rad).

Once the optimal induction conditions had been identified, the bacteria were harvested by centrifugation at 5,000 xg for 20 min at 4°C (Heraeus Primo, Thermo Scientific, UK) and the cell pellet was retained and stored at -20°C prior to purification.

2.2.6 Recombinant protein purification

Recombinant proteins were extracted from the induced bacterial cell pellet using the procedure described in the QIA*expressionist* kit instructions (Qiagen, Crawley, UK). Frozen cell pellets were thawed for 15 min on ice and resuspended in a lysis solution containing 10ml 1X Ni-NTA Bind Buffer (50 mM NaH2PO4, pH 8.0; 300 mM NaCl; 10 mM imidazole) (Ni-NTA buffer kit; Novagen, USA), lysozyme (1 mg/ml) (Roche,

Germany) RNase A (10 μg/ml) (Sigma-Aldrich, Gillingham, UK) and DNase I (5 μg/ml) (Sigma-Aldrich, Gillingham, UK), which was then incubated on ice for 30 min. The cell suspension was then subjected to sonication (five 40 second bursts at 14 am with a 30 second cooling period between each burst) using a Soniprep 150 ultrasonic disintegrator (Sanyo, Leicester, UK).

The resulting lysates were then centrifuged at 10,000g for 30 min at 4°C (Beckman Coulter Optima L-100K centrifuge, rotor 14) to pellet cellular debris. The His-tagged recombinant proteins were recovered from the clear lysate under native conditions by affinity chromatography using a Nickel-nitrilotriacetic acid (Ni-NTA) Buffer Kit (Novagen, USA), in accordance with the manufacturer's instructions.

Briefly, 1–2 ml of 50% Ni-NTA Agarose (Qiagen, Crawley, UK) was added to the cleared lysate and gently mixed by shaking at 30 rpm on a Stuart SSL 3 Gyro Rocker (Sigma-Aldrich, Gillingham, UK) at room temperature for 60 min. Following mixing, the lysate–Ni-NTA mixture was loaded into a 10 ml polypropylene disposable column (Fisher Scientific, Leicestershire, UK) and washed twice with 4 ml 1X Ni-NTA Wash Buffer (50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 20 mM imidazole). The bound recombinant protein was eluted from the column using 0.5 ml 1X Ni-NTA Elution Buffer (50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 250 mM imidazole). This elution step was repeated four times. Eluted proteins were immediately buffer-exchanged into 20 mM sodium phosphate buffer (pH 6) using a 10 ml Zeba Spin Desalting Column (Thermo Scientific, UK). Purified recombinant protein was stored at 4°C until required.

To monitor the efficiency of the purification process 50 µl aliquots were collected at the following stages during the process: culture lysate supernatant, column wash flow-through (wash fractions 1 and 2) and elution fractions 1, 2, 3 and 4. Samples were analysed by SDS-PAGE.

2.2.7 Protein quantification

Protein quantification of purification fractions was determined using the Bicinchoninic acid (BCA) assay using the method described in Pierce BCA protein assay kit instructions (Thermo Scientific, at http://www.piercenet.com/instructions/2161296.pdf . accessed on 5-2011) with the following modifications.

A series of bovine serum albumin (BSA, Fisher Scientific, Leicestershire, UK) standards ranging from 0.05 to 1.0 mg/ml were prepared in PBS in order to construct a standard curve. To determine the protein concentration of a sample, 10 µl of the sample was mixed with 190 µl of PBS buffer and all test samples were prepared in triplicate. The BCA working solution was prepared by mixing 50 parts BCA (Sigma-Aldrich, Gillingham, UK) solution with 1 part 4% (w/v) copper sulphate ($CuSO_4$) (Fisher Scientific, Leicestershire, UK). One ml of this BCA solution was added to each standard and protein sample and vortex mixed, after which they were incubated at 37°C for 15 minutes. At the end of this period, samples were immediately placed on ice and the absorbance of each sample was determined at 562nm (Ultra space 3100pro, GE Healthcare), using a sample comprising PBS as the blank.

Using the results, a Beer's law plot was constructed by plotting the mean absorbance at 562 nm of the protein standards against their concentration to produce a straight line, from which the protein concentration of the unknown samples could be determined. The final concentration of the sample was corrected for the dilution factor by multiplying by 20.

2.2.8 Sodium dodecyl sulfate polyacrylamide gel electrophoresis and protein gel staining

Proteins were visualised by SDS-PAGE using 4–20% pre-cast gels Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad). Samples were resuspended in 50 µl of 2x SDS-PAGE sample buffer (0.09M Tris-Cl pH 6.8, 20% (w/v) glycerol, 2 % (w/v) SDS, 0.02% (w/v) bromophenol blue, 0.1 M DTT) and stored at -20°C prior to analysis. Following removal from storage, samples were denatured using a dry block heater (Techne DB-3D, Bibby Scientific Ltd, Staffordshire, UK) at 95°C for 5 minutes, after which 10 µl volumes were loaded onto the PAGE gel. To determine the migration distance of recombinant proteins, a Molecular weight standard (Sigma Marker wide Range: 6,500-200,000 Da, Sigma-Aldrich, Gillingham, UK) was run on each gel. Following sample loading, gels were placed in a tank (Bio-Rad) and overlaid with 1X SDS-PAGE running buffer (10x running buffer contains: 250 mM Tris base, 1.92 M glycine, 1% SDS, pH 8.3; Fisher Scientific, Leicestershire, UK). Gels were then subjected to electrophoresis for 45 minutes at 150 V (400 mA).

On completion of electrophoresis, the SDS-PAGE gel was then washed twice, for five minutes each time, in sterile distilled water. The gel was then fixed in 100 ml of fixing solution (40% (v/v) methanol, 10% (v/v) glacial acetic acid in diH_2O, 10 ml methanol, 7 ml glacial acetic acid and 83 ml water) for 1 hour. A 1x working solution of Brilliant Blue G-Colloidal concentrate (Sigma-Aldrich, Gillingham, UK) solution was prepared and stored at 4°C. Prior to staining, 4 parts of working solution were mixed with 1 part methanol. Staining was carried out for 2 hours and the gel was de-stained with 25% methanol, continuously changing the methanol until the stain had been removed from the gel and the gel was ready for imaging. On completion of staining, gels were imaged on the ChemiDoc XRS equipped with a 12-bit CCD (charge-coupled device) camera (Bio-Rad).

2.2.9 Western blot

Prior to antibody probing, proteins were separated by SDS-PAGE as described above. A 6x His Protein Ladder (Qiagen Ltd, Crawley, UK) was included as a positive control with a culture of *E. coli* M15 serving as a negative control. All incubations were carried out at room temperature on a Stuart SSL 3 Gyro Rocker. Upon completion of electrophoresis, proteins were transferred from the polyacrylamide gel to a nitrocellulose membrane (0.2 μM pore size, Bio-Rad) using the Bio-Rad Trans-Blot® Turbo™ System at 25 volts, 1.0A for 30 h.

Detection of histidine-tagged proteins was achieved using a histidine-specific murine antibody, as described in the QIA*express*® Detection and Assay Handbook. Tween 20 was purchased from Sigma-Aldrich and all other reagents were purchased from Fisher Scientific. All steps were carried out at room temperature on a Stuart SSL 3 Gyro Rocker. The membrane was washed twice for 10 minutes with PBS and 0.1 % (v/v) Tween 20 (PBST) and thereafter incubated in blocking buffer (3% non-fat dried milk (w/v) in PBST buffer) for 1 hour. Membranes were then washed twice for 10 minutes in PBST. The detection antibody, Penta-His Antibody (Qiagen Ltd, Crawley, UK) was diluted 1:10,000 in blocking buffer and then incubated with the membrane for 1 h.

Prior to the addition of the secondary antibody, the membranes were washed four times for 10 min with PBST. The secondary anti–mouse IgG/HRP conjugate antibody (Thermo Scientific, UK) was diluted 1:10,000 in blocking buffer and then incubated with the membrane for 1h, after which the membrane was washed four times for 10 min with PBST. Immobilon Western chemiluminescent HRP substrate (Millipore; Billerica, MA, USA) was used to detect the presence of bound antibody complexes which we subsequently visualised using a ChemiDoc XRS with a CCD camera (Bio-Rad) detection system.

2.3 Statistical analysis

The results were compared using one-way ANOVA in the Minitab Statistical Software Package (release 14 software, Minitab Inc., PA, USA) and means results were compared using the TukeyHSD method. Two-group comparisons were analyzed by Student's t test with the statistical function of excel (Microsoft Office 2010)

2.4 Bioinformatics analysis

Nucleotide sequences were analysed using the Basic Local Alignment Search Tool (BLASTn) to identify homologous regions to sequences deposited in the NCBI database (http://www.ncbi.nlm.nih.gov). The presence of restriction enzyme sites in a nucleotide sequence was determined using NEBcutter v2.0 program (http://tools.neb.com/NEBcutter2/).

BLASTp (Altschul *et al*., 1997) was used to compare protein sequences to the non-redundant protein sequence database in the NCBI database. Multiple sequence alignments were performed using either Clustal W2 (available at http://www.ebi.ac.uk/Tools/msa/clustalw2/) or T-Coffee (Notredame *et al*., 2000), and the creation of phylogenetic trees using the average distance method was performed using Clustal W2.

The prediction of protein functional domains was made using Pfam (http://pfam.janelia.org/search). The isoelectric point (pI) and molecular weight (Mw) of protein sequences were calculated using Compute pI/Mw tool in ExPASy

(http://www.expasy.org/tools/pi_tool.html). Signal peptide predictions were performed using SignalP 4.0 Server (Petersen *et al*., 2011; Gram-positive organism group), while the solubility of recombinant proteins in *E. coli* were predicted using Recombinant Protein Solubility Prediction at http://www.biotech.ou.edu/.

Chapter Three

Isolation of bacteriophages infecting *Clostridium difficile*

3.1 Introduction

3.1.1 Isolation of temperate bacteriophages

The induction of phages from *C. difficile* by treatment with mitomycin C (MMC) has previously been described (Sell *et al.,* 1983; Nagy *et al.,* 1991; Goh *et al.,* 2005b, 2007; Govind *et al.,* 2006; Fortier and Moineau, 2007; Mayer *et al.,* 2008; Horgan *et al*., 2010; Meader *et al*., 2010; Sekulovic, *et al*., 2011; Meessen-Pinard *et al*., 2012; Nale *et al*., 2012; Shan *et al.,* 2012). Recently, there have been successful isolations of whole phages and purified phage components, with a number of temperate *C. difficile* phages having been isolated and observed by the use of electron microscopy (EM). These phages have been classified into two families. The first family are the *Myoviridae* which have an isometric capsid and a contractile tail and comprise of φMMP01,φMMP02,φMMP03,φMMP04, φC2, φCD27, φCD8, φCD630-1 and φCD630-2;the second family consists of a number of prophages belonging to the *Siphoviridae* family, such as φ 41, φCD8-2, φC6, φCD38-1, φCD38-2, φCD6356 and φCD6365. This family has an isometric capsid and a long non-contractile tail (Chapter 1, table 1.2).

Mitomycin C is an alkylating agent which has been used as an anti-tumour antibiotic since the 1950s. Mitomycin C is a natural product isolated from *Streptomyces caespitosus* and was discovered by a Japanese microbiologist in the 1950s (Tomasz, 1995). Its mode of action involves cross-linking DNA bases, which leads to cell death (Tomasz, 1995).

The DNA damage to the bacterial cell appears to be the trigger that stimulates the production of RecA via the SOS repair system, which induces prophages in the lytic cycle (Kuzminov, 1999).

The antibiotic norfloxacin has also been used as a phage induction agent; it is a specific inhibitor of the A subunit of DNA gyrase which inhibits DNA synthesis, resulting in the accumulation of single-stranded DNA fragments. This results in the activation of the RecA protein, facilitating the cleavage of a phage repressor, thus allowing prophage induction (Matsushiro *et al.*, 1999). Norfloxacin has been used successfully in the induction of phages from *E. coli* O157:H7 and Shiga toxin (Stx)-producing *E. coli* (STEC) strains (Matsushiro *et al*., 1999; McDonald *et al*., 2010). In addition, it has also been used to induce the expression of prophages from examples of the *C. difficile* 027 ribotype (Nale *et al*., 2012; Shan *et al.*, 2012). Moreover other quinolones such as Ciprofloxacin, Moxifloxacin, and Levofloxacin have recently been used at sub-MIC concentrations to stimulate the expression of *C. difficile* prophages (Meessen-Pinard *et al.*, 2012).

Other induction methods such as UV light treatment have been used to induce the expression of *C. difficile* prophages (Mahony *et al.*, 1985; Sekulovic, *et al*., 2011).

It is interesting to note that *C. difficile* prophages can enter the lytic cycle in the absence of inducing agents by a process known as spontaneous phage induction (Little, 2005; Govind *et al.*, 2006). The *C. difficile* temperate phages φ41 and φ56 were spontaneously released and detected using a direct agar plating technique similar to the method employed in this chapter (Mahony *et al.*, 1985, 1991; Ramesh *et al.*, 1999; Govind *et al.*,

2006). Spontaneous *C. difficile* temperate phages φMMP03 and φMMP04 have also been identified using a method known as host adsorption, which is also employed in this chapter in an attempt to isolate phages from environmental samples (Grabow, 2001; Meessen-Pinard *et al*., 2012).

3.1.2 Isolation of virulent bacteriophages

Bacteriophages are widely distributed in the environment and can be isolated from seawater, soil, fresh water and sewage ecosystems (Jensen *et al*., 1998). Indeed, sewage represents a particularly rich source of phages as it contains a large diversity of coliforms due to faecal contamination (Sundar *et al.*, 2009).

Bacteriophages have been isolated from drinking water (Grabow and Coubrough, 1986), fresh water (Goyal, 1987; Grabow, 2001) and seawater (Tartera and Jofre, 1987). Wiebe and Liston (1968) isolated a bacteriophage that was active against a marine *Aeromonas* species from surface sediments of the North Pacific Ocean. Baross *et al*. (1978) reported the isolation of bacteriophages active against *Vibrio* from marine water and sediments.

Phage recovery methods should ideally be rapid, cheap and easy to perform. They should be applicable to a wide range of samples and should give high efficiency in recovery (Grabow, 2001). Several methods have been used to isolate phages from environmental samples. These consist of direct culture methods, enrichment methods, adsorption-elution

methods using negatively or positively charged filters or glass powder, precipitation using polyethylene glycol (PEG), and ultrafiltration (Goyal, 1987; Grabow, 2001). However, in this study, direct culture methods and host adsorption methods were used, since these methods are easy to use, inexpensive, suitable for a large amount of samples and have been used before for *C. difficile* lytic phage isolation (Seeley and Primrose, 1982; Grabow, 2001; Goh *et al*., 2005b).

3.1.3 Aim of this study

The aims of this chapter are:

- To isolate phages from twenty-three clinical isolates of *C. difficile*, representing a range of ribotypes using mitomycin C and norfloxacin.
- To isolate phages from environmental samples those are lytic for *C. difficile*.
- To determine the physical characteristics of phages isolated during this study.

3.2 Materials and Methods

3.2.1 Isolation of temperate bacteriophages from clinical isolates

Unless indicated otherwise, all centrifugations were performed using a Heraeus Primo centrifuge (Thermo Scientific, UK) and all anaerobic work was performed in a gloveport Bugbox Plus anaerobic workstation (Ruskinn Technology Ltd). The Optical density (OD_{600}) of bacterial cultures was determined using an Ultra spec 3100 pro spectrophotometer (GE Healthcare, Piscataway, NJ, USA).

3.2.2 Strains of *C. difficile* used to isolate and propagate bacteriophages

The ribotypes of the twenty clinical isolates and three NCTC reference strains of *C. difficile* used in this phage study are shown in table 3.1. *C.difficile* NCTC 12727 is of particular note, as it has been shown to contain a phage whose expression is inducible by treatment with mitomycin C and which is able to infect *C.difficile* NCTC 11204 (Mayer *et al.,* 2008).

Table 3.1: *C. difficile* strains that were used in phage induction studies described in this Chapter (Appendix 1)

Ribotype	Strain
001	R8562 –DS1747-DS1750-DS1759, NCTC 12727, NCTC 11209, NCTC 11204
023	DS1665
027	DS1801-DS1807 -R20291-DS1813
106	R10459-DS1798- DS1771
002	DS1748
078	DS1723
012	DS1752
014	DS1742
005	DS1721
010	DS1684
017	R22660
020	DS1724

For this reason, these two strains, NCTC 12727 and NCTC 11204, were used to optimise culture conditions, to establish the most appropriate plaque assay and to identify the optimal phage induction conditions for mitomycin C and norfloxacin.

3.2.3 Optimisation of *C. difficile* culture conditions prior to phage induction.

The manner in which a bacterial culture is grown and its stage of growth is known to influence the efficacy of phage induction (Gill and Hyman, 2010). To identify the optimal culture conditions for our collection of *C. difficile* isolates, the following methods were assessed: static tube culture and Bioscreen microtiter plate.

3.2.3.1 Tube method

Individual isolates were cultured on pre-reduced BHI agar plates for 48 hours at 37ºC under anaerobic conditions, after which a single colony was harvested and used to inoculate 5 ml of pre-reduced BHI broth; this was subsequently incubated for 24 hours at 37ºC. At the end of this period, 1.25 ml (adjusted volume) of cell suspension was transferred into 8.75 ml of fresh pre-reduced BHI broth and incubated for a further 48 hours in an anaerobic environment. Aliquots (1ml) were removed at different time points and the optical density (OD_{600}) was determined.

3.2.3.2 Bioscreen microtiter plate method

Individual isolates were cultured on pre-reduced BHI agar plates for 48 hours at 37ºC under anaerobic conditions, after which a single colony was harvested and used to inoculate 5 ml of pre-reduced BHI broth; this was subsequently incubated for 24 hours at 37ºC. Following this period, 50 µl of cell suspension was transferred into a microtiter plate well, where it was mixed with 350 µl of the pre-reduced BHI broth. For each strain, a total of four wells were inoculated on each plate. Negative control wells contained culture media only. To maintain anaerobic conditions, individual wells were sealed with either silicone glue, sterile immersion oil or adhesive tape. The inoculated plate was then loaded on to the Bioscreen-C system (Lab Systems, Helsinki, Finland) and incubated at 37ºC for 48 h. The optical density (OD_{600}) of each well was measured every 15 minutes. The resulting data was analysed using spread sheet software (Excel 97) and the mean of the quadruplicate values for each isolate were calculated and used to generate a growth curve.

3.2.4 Phage induction using mitomycin C and norfloxacin

The ability of different concentrations of mitomycin C and norfloxacin (3 or 5 µg/ml) to induce the production of lytic phages by bacterial cultures at different stages of physiological development (mid log and stationary phase) was determined using the methods described elsewhere (Sell *et al.,* 1983; Nagy *et al.,* 1991; Fortier and Moineau, 2007; 2012; Nale *et al*., 2012; Shan *et al.,* 2012) with following modifications.

The test strain *C.difficile* NCTC 12727 was initially cultured on degassed BHI agar for 48 hours at 37°C under anaerobic conditions. A single colony was harvested and transferred into 5 ml of BHI broth and grown at 37°C under anaerobic conditions for 24 h. A 5 ml inoculum from this culture was then transferred into 45 ml of fresh degassed BHI medium, and growth was monitored until the optical density (OD_{600}) reached the required values. In the case of mid log phase cultures this was OD_{600}= 0.4-0.6, while for stationary phase cultures the values were OD_{600} =1.0-1.2.

Once the desired OD_{600} was achieved, 10 ml aliquots were pipetted into individual 50 ml Falcon tubes where the cultures were mixed with either mitomycin C (Arcos Organics, Geel, Belgium) or norfloxacin (Sigma-Aldrich, UK) to achieve a final concentration of 3 or 5 µg/ml , reference control cultures were left untreated. Following treatment, cultures were incubated at 37°C under anaerobic conditions and OD_{600} was monitored. A significant decrease (>60% of initial OD) in cell density compared to the untreated control at 3 to 5 hours post-induction was taken to indicate the induction of phage expression (Fortier and Moineau, 2007) . Induced broth cultures were centrifuged at 5,000xg for 10 minutes; the supernatant was collected and filtered through a 0.45 µm

membrane filter (Millipore, Massachusetts, US) to remove bacterial debris. The subsequent filtrates were stored at 4°C prior to further analysis.

3.2.5 Plaque assay development: Optimisation of *C. difficile* inoculum

Two assay formats were assessed for their suitability as a means of determining the phage susceptibility of *C. difficile* isolates; direct plating (Mazzocco *et al.,* 2009) and agar overlay (Adams, 1959; Sell *et al.,* 1983; Mahony *et al.,* 1985 ; Govind *et al.,* 2006 ; Fortier and Moineau, 2007, Fortier and Moineau, 2009; Sekulovic, *et al.,* 2011) . The optimum bacterial inoculum was determined for each assay.

3.2.5.1 Direct plating plaque assay

A 1ml culture of NCTC 11204 in BHI broth, which was incubated at 37ºC until the culture reached an OD_{600} of 0.4 to 0.6 (mid log phase), was spread over the surface of a pre-reduced BHI agar plate. The number of viable bacteria in the culture was also determined using the method described in section 2.1.7. Excess liquid was removed by Pasteur pipettes, the plate was allowed to dry and was then incubated anaerobically at 37ºC for 24 and 48 hours, in order to assess the method's ability to produce a confluent lawn of bacteria growth.

3.2.5.2 Agar overlay plaque assay

A single colony of a 48 h culture of *C.difficile* NCTC 11204 on BHI agar was harvested and inoculated into 5 ml of pre-reduced BHI broth and incubated at 37ºC for 24 h. Following incubation, 1.25 ml of cell suspension was added to 8.75 ml of pre-reduced

BHI broth and the mixture was subsequently incubated until the culture reached an OD_{600} 0.4 to 0.6 (mid log phase). Volumes of 25 µl, 100 µl, 200 µl, 300 µl, 400 µl, 500 µl and 1 ml of this culture were seeded into molten agar (0.75% BHI agar -cooled to 45°C) supplemented with $MgCl_2$ (0.1- 0.4 M) and $CaCl_2$ (0.1- 0.4 M) as the overlay. Once inoculated the agar/ bacterial mixture was spread over the surface of a BHI agar plate and allowed to solidify. The plates were then incubated at 37ºC for 48 h under anaerobic conditions and the subsequent degree of growth was assessed. In the two methods, samples were assayed for the presence of phage by spotting 10 µl of filtered phage extract onto the surface of the solidified agar.

3.2.6 Concentration and purification of phage

Broth cultures were induced and centrifuged in a 45 ml volume, as previously described in prophage induction (section 3.2.4) .The phage concentration and purification method was performed based on the method described elsewhere (Fortier and Moineau, 2007; Nale *et al*., 2012) with following modifications The supernatants were transferred to a sterile conical flask (500ml) that contains 1M sodium chloride (NaCl) (Fisher Scientific, UK). The suspensions were mixed and incubated on ice for 1 h with continuous stirring, then 10% (w/v) of polyethylene glycol (PEG) 8000 (Fisher Scientific, UK) was slowly added with continuous mixing until completely dissolved. The mixtures were left overnight at 4°C with constant mixing. The solutions were then centrifuged at 11,000xg at 4°C for 30 minutes (Beckman Coulter Optima L-100K centrifuge, rotor 14). The supernatants were discarded and the pellets re-suspended in 3 ml of SM buffer (50 mM

Tris-HCl, pH 7.5, 100 mM NaCl, 8 mM $MgSO_4$, 0.002% gelatine, Fisher Scientific, Leicestershire, UK). After this, each mixture was vortexed and left at room temperature for one hour. The solution was then centrifuged at 5,000xg at 4°C for 20 minutes and the supernatant containing the phage was filtered through 0.22 mm filters and stored at 4ºC.

3.2.7 Transmission Electron Microscopy (TEM)

All TEM materials and chemicals were purchased from Agar Scientific (Essex, UK) unless otherwise stated in the text. A 5 µl drop of purified concentrated phage suspension was added to carbon coated Formvar grids (400 mesh), which were dried at room temperature for 2 minutes, with any excess fluid being removed using filter paper. The samples were subsequently negatively stained with 5 µl of 2% uranyl acetate and dried at room temperature for 2 minutes. The samples were imaged using a Phillips CM12 transmission electron microscope with an accelerating voltage of 80 kilovolts (kV); magnifications between 25,000x and 63,000x were used. Digital images were captured with a Megaview III digital TEM camera and iTEM software (Soft Imaging System GmbH, Münster, Germany). Representative images were taken using 30 to 40 fields of vision. Morphology was assigned by using figure 1.8 (Chapter 1) after measuring the dimensions of the phages' capsids and tails (Murphy *et al*., 1995).

3.2.8 Isolation of virulent bacteriophages from environmental samples

Eighteen water samples were collected in 1 litre Sterile Duran bottles from the following locations around Cardiff: three from the River Taff, three from the River Ely, five from Cardiff Bay, and seven from Roath Park Lake. Each sample was centrifuged at 2100xg for 30 minutes at 4°C and then filtered through a 0.22 μm-pore-size filters (Millipore, Massachusetts, US) in order to remove bacterial and other suspended material.

3.2.8.1 Direct culture method

The direct culture method includes direct plating and agar overlay technique as follows:

Isolation of bacteriophages using the direct plating plaque assay

The test strains of *C. difficile* were cultured in pre-reduced BHI agar for 48 hours under anaerobic conditions at 37°C. Subsequently, the direct plating plaque technique was performed as described before in section 3.2.5.1**.** A 10 μl aliquot of the filtrated environmental sample was then spotted on top of the streak and left to dry. The plates were transferred to an anaerobic cabinet, incubated for 48 hours at 37°C and examined for the presence of plaques.

Isolation of bacteriophages using the agar overlay plaque assay

Cultures of *C. difficile* were prepared as described above. A total of 3ml of molten BHI overlay agar (0.75% supplemented with 0.4 M $MgCl_2$) was seeded with 1ml of culture (OD_{600}, 0.4 to 0.6) and allowed to set. Samples were assayed for the presence of phage by spotting 10 µl of filtered environmental sample onto the surface of the solidified agar. The plates were left to dry and then incubated overnight anaerobically and examined for the presence of plaques.

3.2.8.2 Host adsorption method (enrichment method)

A single colony of a 48 h culture on BHI agar was harvested and inoculated into 5 ml of pre-reduced BHI broth, before being incubated at 37ºC for 24 h. Following incubation, the entire culture was added to 15 ml of pre-reduced BHI broth in a Falcon tube (50 ml). The tube was incubated at 37ºC until the culture reached an OD_{600} of 0.4 to 0.6, at which point 20 ml of filtered environmental sample was added. The mixture was incubated overnight anaerobically at 37°C and centrifuged at 2100xg for 30 minutes, after which the culture supernatant was passed through a 0.45µm filter. To screen for the presence of lytic phages, 10 µl aliquots were assayed using the agar overlay assay. This method was modified and simplified from methods described by Purdy *et al*. (1984) and Grabow (2001).

3.3 Results

3.3.1 Optimisation of *C. difficile* broth growth conditions

The Bioscreen microtiter plate system allows a large number of samples to be screened over time and thus represents an ideal system with which to screen the growth characteristics of multiple bacterial isolates. We attempted to develop an approach which would allow us to monitor the growth of *C. difficile*. Unfortunately, the system lacks an anaerobic culture capacity and none of the sealing methods employed were able to achieve an effective oxygen seal.

In contrast, the tube method was able to support anaerobic growth, as can be seen for the growth curves of *C. difficile* NCTC 11204 and 12727 presented in figure 3.1. The growth kinetics of each strain was similar, with a peak in cell density at 24 hours. As one would expect, the production of spores was detected at 48 hours. For this reason, the tube method was employed to support phage induction studies.

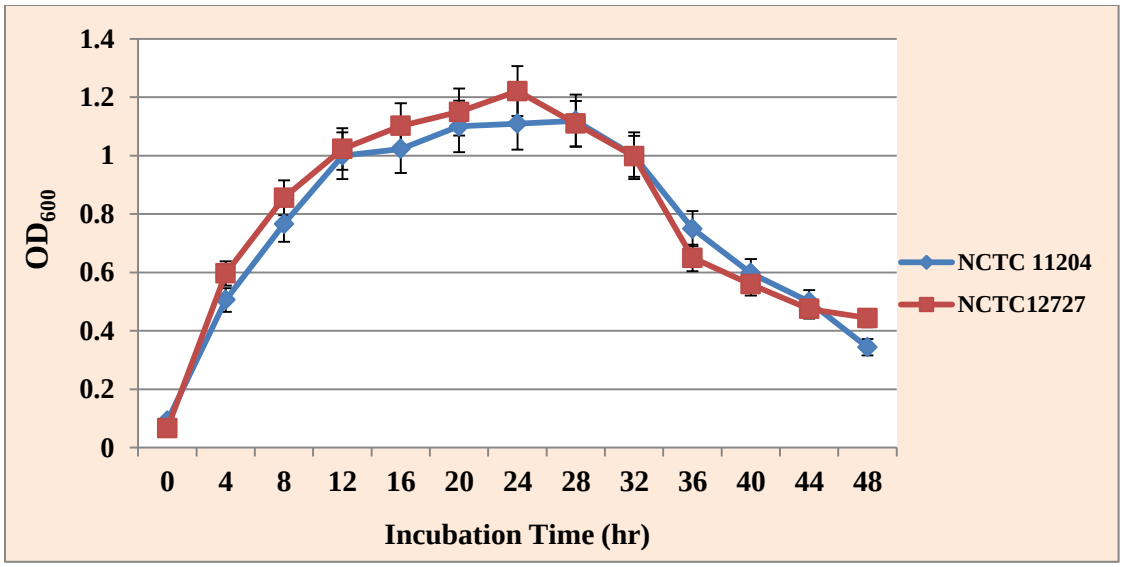


Figure 3.1: Growth curves of *C. difficile* strains NCTC 11204 and 12727. Strains were grown anaerobically in BHI medium and aliquots were removed for OD_{600} readings. Data represent mean ± S.D from triplicate assays

3.3.2 Optimisation of *C. difficile* agar culture conditions

The method we employed to plate *C. difficile* directly onto the surface of BHI agar produced a thick, non-confluent growth and bacteria colonies are discrete on the plate when repeated on four separate occasions. In contrast, the agar overlay method using 400 µl of culture at an OD_{600} (0.4-0.6) in 3 ml of overlay agar was reproducible (repeated 4 times) and gave confluent growth of *C.difficile* NCTC 11204. Therefore, agar overlay method was used for further phage propagation methods. Furthermore, due to a lack of positive phage controls that can be used to optimise $MgCl_2$ and $CaCl_2$ concentrations, we decided to use 0.4 M $MgCl_2$, since this concentration had been previously proven to be the optimum concentration for propagation of *C. difficile* phages (Mahony *et al.*, 1985).

3.3.3 Phage induction

We determined the optimal phage induction conditions using *C.difficile* NCTC 12727. Mid-log phase and stationary phase cultures of the organism were exposed to 3 or 5 μg/ml or mitomycin C and norfloxacin. As determined by a reduction in culture optical density, we failed to observe any evidence of phage induction in stationary phase cultures of *C.difficile* NCTC 12727 following treatment with either concentration of mitomycin C or norfloxacin. These results suggested that there was no cell lysis and no release of viral particles, a hypothesis which was confirmed by TEM.

However, mid-log phase cultures of *C.difficile* NCTC 12727 (OD_{600}~0.4-0.6) with 3 and 5 μg/ml of mitomycin C resulted in a marked reduction in optical density suggestive of phage mediated cell lysis. The presence of phages in the culture supernatant following induction was confirmed by TEM. However, prophage expression was not induced from this strain of *C. difficile* by treatment with norfloxacin (Figure 3.2).

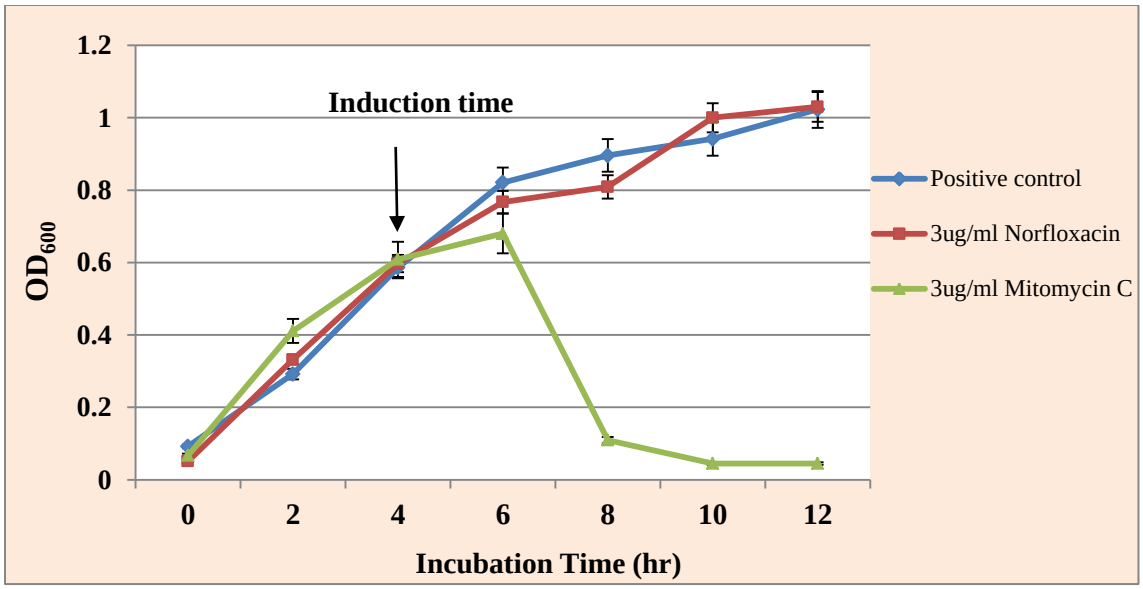


Figure 3.2. Growth curves of the induced *C. difficile* strain NCTC12727. Strain was grown anaerobically in BHI medium and induced at the mid-log phase (4h) using mitomycin C and norfloxacin (3 μg/ml each); aliquots were removed for OD_{600} readings every 2 h. Results are the means ±standard deviations (SD) from triplicate assay.

3.3.4 Phage propagation

The purified and concentrated phage preparation isolated from *C.difficile* NCTC 12727 following mitomycin C induction produced a hazy plaque when plated on a lawn of *C.difficile* NCTC 11204 on agar supplemented with 0.4 M $MgCl_2$ following 24 h of incubation (Figure 3.3). However, it failed to produce plaques upon a second passage.

We tried to identify a phage indicator strain that could be used to propagate newly induced prophages from our collection of clinical isolates. Unfortunately, none of the 23 *C. difficile* isolates examined proved to be susceptible to any of the phage extracts produced in this study, including that of *C.difficile* NCTC 12727. This failure could reflect our failure to efficiently induce the expression of lytic phages from these isolates

or may reflect the inherent resistance of *C. difficile* isolates to phage infection. The difficulty in finding suitable sensitive strains has been reported before (Fortier and Moineau, 2007; Nale *et al*., 2012).

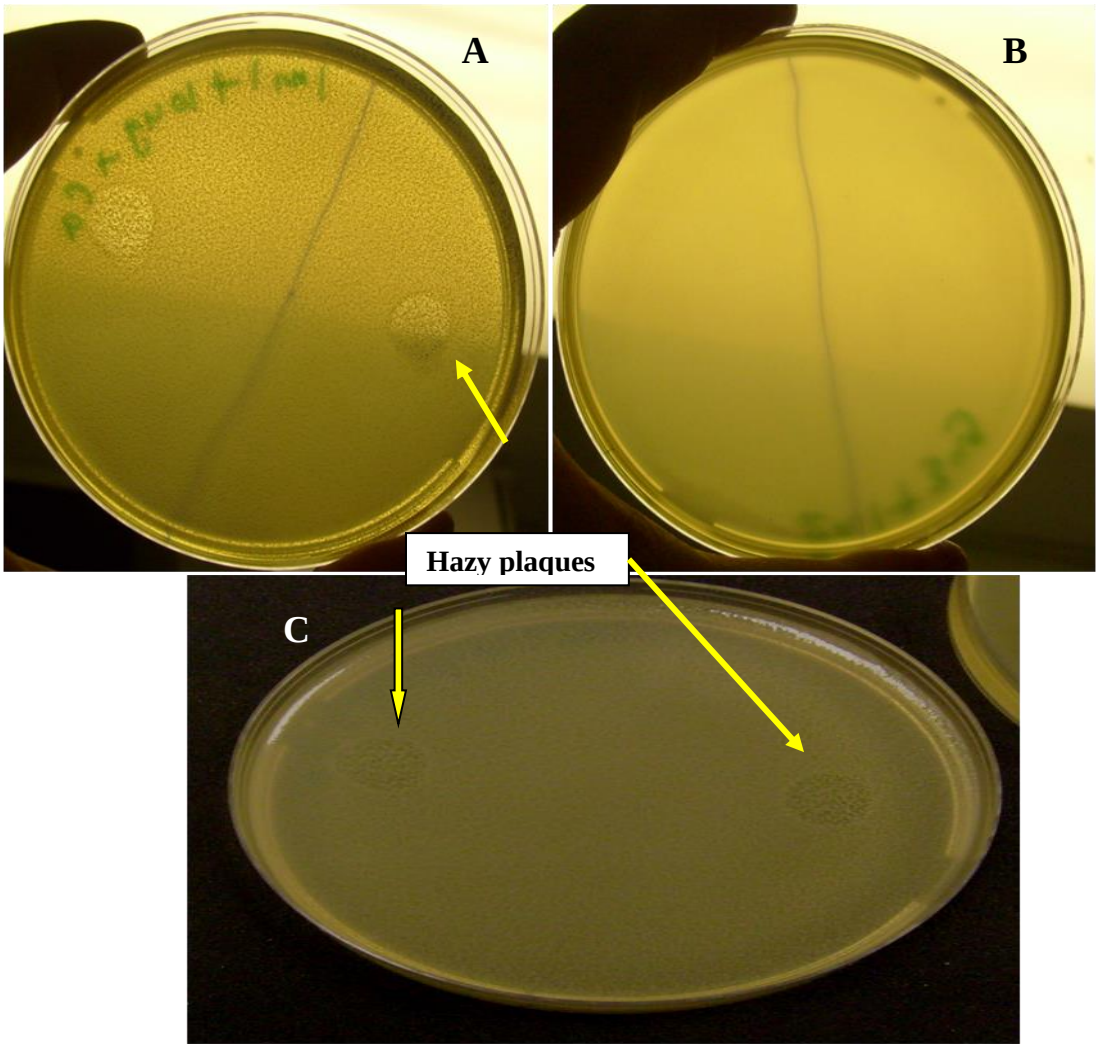


Figure 3.3. A & C) Plaque assay of phages isolated from *C. difficile* NCTC 12727 following mitomycin C induction on a culture of *C. difficile* NCTC 11204. B) An inoculated control culture of *C. difficile* NCTC 11204.

3.3.5 Electron microscopy studies of mitomycin C treated cultures

The presence of viral particles in culture supernatant of induced cultures was confirmed by electron microscopy (Figure 3.4).

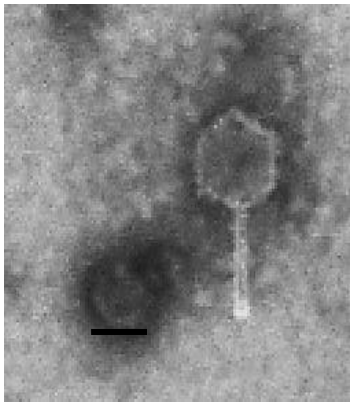


Figure 3.4. TEM results of phage induction method of *C. difficile* NCTC 12727, showing that the phage belongs to the *Myoviridae* family. Bar represents 50nm; the size of the phage was determined by measuring 3 phages. Representation from 30 fields of view.

Using the optimised phage induction method at the mid log phase, we were able to induce the expression of a phage from *C. difficile* NCTC 12727, which showed a clear resemblance by electron microscopy to the φCD27 bacteriophage isolated from *C. difficile* NCTC 12727 by Mayer and colleagues using mitomycin C (Mayer *et al.*, 2008). However, the phage was induced at a low level, since a low number of phages was seen per grid (one in 30 fields).

Of the twenty-two clinical isolates of *C. difficile* exposed to mitomycin C, twelve strains, representing nine different ribotypes, underwent lysis when exposed to mitomycin C.

Electron microscopy studies of the culture supernatant of these induced cultures revealed the presence of twelve phages, eleven of which were morphologically identical with an icosahedral head measuring ~35-75 nm and contractile tails between ~90-170 nm in length (Table 3.2). Measurement was estimated by measuring 3 phages in each sample. The presence of a tail sheath suggests that these phages belong to the family *Myoviridae.* These phages show a clear resemblance to other *C. difficile* bacteriophages (Chapter 1, table 1.2). Defective phage particles were also seen in some of the samples. However, these phages were also induced at low levels, since a low number of phages were seen per grid (one in 40 fields).

Table 3.2. TEM results of phages induced using mitomycin C. All phages belong to the *Myoviridae* family. Bar represents 50 nm. Measurements were estimated by measuring 3 phages of each sample .Representation from 40 fields of view**.**

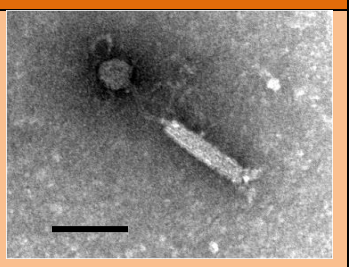
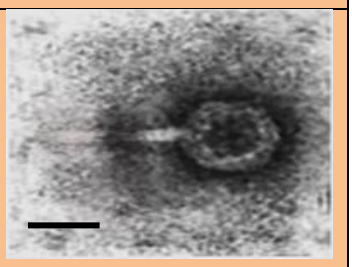
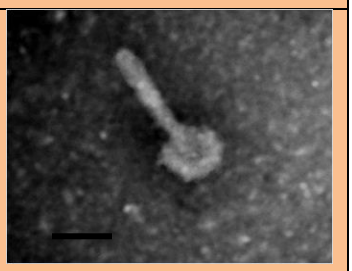
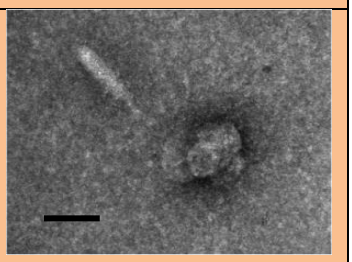
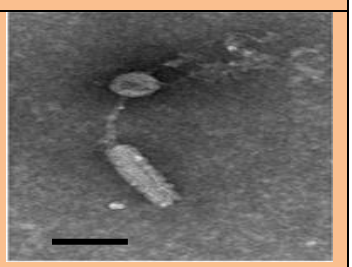
Ribotype	Strain	TEM
001	**DS1750**	
001	**DS1759**	
001	**R8562**	
001	**DS1747**	
014	**DS1742**	

Table 3.2 (continued)

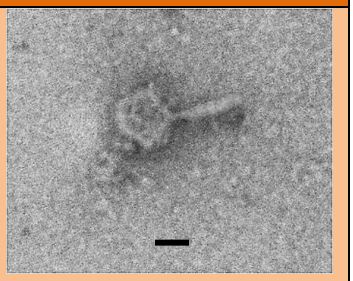
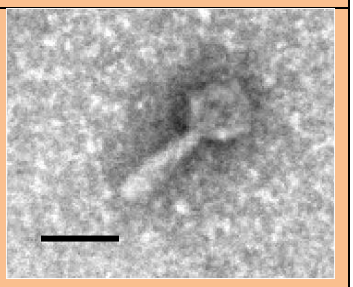
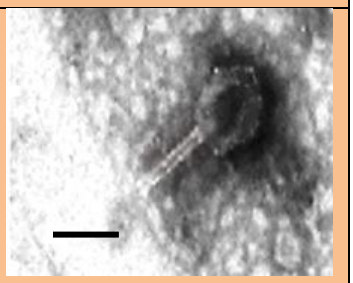
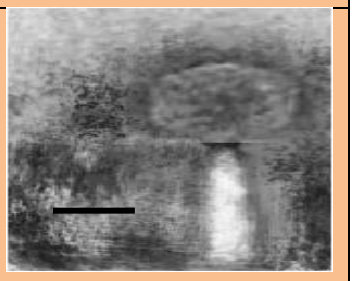
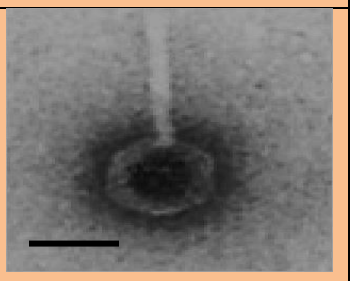
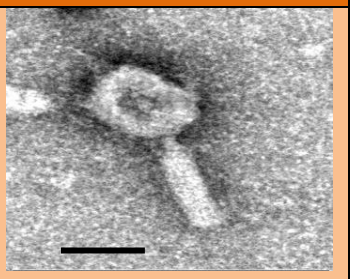
Ribotype	Strain	TEM
002	DS1748	
106	DS1771	
027	DS20291	
010	DS1665	
023	DS1684	

Table 3.2 (continued)

Ribotype	Strain	TEM
005	DS1721	

Interestingly, one of the induced phages (Figure 3.5) induced from *C. difficile* R22660 (ribotype 017) belonged to the *Siphoviridae* family, with an isometric capsid (~92 nm) and a long non-contractile tail (~211 nm).

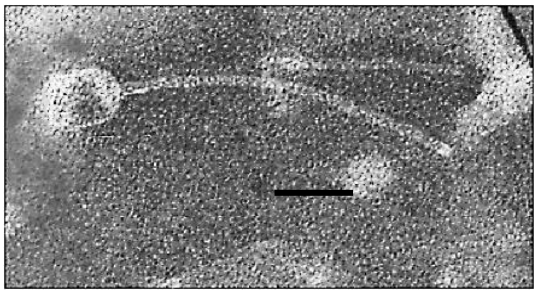


Figure 3.5. TEM results of phages induced from *C. difficile* R22660 using mitomycin C. The morphology of the phages shows that they belong to the *Siphoviridae* family. Bar represents 50nm; Measurements were estimated by measuring 3 phages. Representation from 40 fields of view.

The presence of phage tail-like particles (PT-LPs) were observed in some of our induced lysates (strains DS1750, DS1665, DS1747, NCTC12727 and DS1684) and these particles were morphologically similar to those previously described (Nagy and Foidas, 1991; Fortier and Moineau, 2007; Nale *et al*., 2012) (Figure 3.6). These can be classified as R-type bacteriocins, which resemble the structures of *Myoviridae* phage tails (contractile, nonflexible tails) (Gebhart *et al*., 2012). Similar PT-LPs were also observed in *Vibrio*, *Budvicia aquatica* and *Pragia fontium* cultures (Smarda and Benada, 2005; Gnezda-Meijer *et al*., 2006).

Interestingly, some of these PT-LPs possess flower-like appendages similar to those observed before (Fortier and Moineau, 2007; Nale *et al*., 2012; Gebhart *et al*., 2012), which were proven experimentally to have bactericidal activity against strains of *C. difficile* (Gebhart *et al*., 2012).

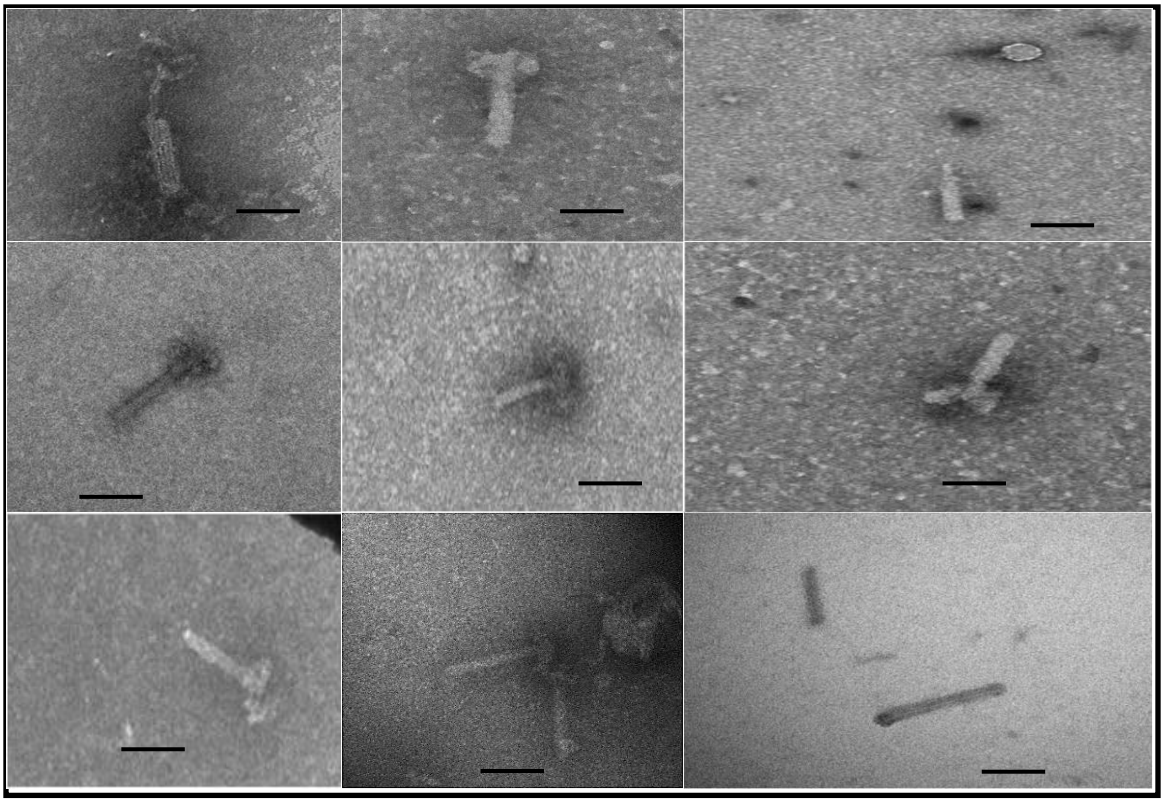


Figure 3.6. Phage tail-like particles (PT-LPs) induced in the culture supernatant of mitomycin C treated cultures of *C. difficile* DS1747, 1665, 1750, NCTC 12727 and DS1684; bar represents 50 nm. Representation from 30 fields of view.

3.3.6 Isolation of bacteriophages from environmental samples

Eighteen environmental samples (water samples) were collected from areas around Cardiff. These samples were centrifuged at 2100xg for 30 minutes at 4°C and filtered through 0.22 μm pore size filters. Two methods, direct culture (with and without agar overlay) and host adsorption were used in an attempt to isolate virulent phages from these samples.

We failed to isolate any environmental phages with activity against the 23 *C. difficile* isolates included in this study. Although this method is easy and fast, the volume of the test sample is generally limited and is only suitable when the phages are present in high numbers, which is unlikely in most situations (Goyal, 1987). Also, it may be the case that the strains used in this study are inherently resistant to phage infection or that there were no phages present in the original samples.

The enrichment method principle is to increase the number of phages that are present in the environmental sample by mixing it with an appropriate host. This method is sensitive, simple and easy to perform; however, it is not suitable for the detection of a wide spectrum of phages and the correct host has to be chosen (Goyal, 1987). Again, we failed to isolate any lytic phages.

3.4 Discussion

3.4.1 Optimisation of *C. difficile* growth conditions

C. difficile is an obligate anaerobic organism, which means that the total exclusion of oxygen is an important element of any attempt to culture the organism. Our attempts to employ a screening method based on the Bioscreen system were unsuccessful due to the fact that we were unable to house the system inside the anaerobic cabinet for technical reasons. For this reason, three different methods were investigated as means of achieving a suitable anaerobic environment, namely oil, silicon glue and adhesive tape. None of these approaches supported active growth, and for this reason the tube method was adopted. This method has been used successfully by other groups to induce the expression of significant quantities of lysogenic phages (Brüssow *et al.,* 2004).

In addition to optimising culture conditions, we also tried to develop a plaque assay to enable us to identify the presence of induced phages. In our hands, the double-agar-layer plaque technique proved to be the most reproducible approach (Santos *et al*., 2009). Phages can diffuse through the agar gel to form plaques . The concentration of the gel has an effect on the size of the plaque; if the concentration of gel increases, the plaque size will decrease, and so on (Douglas, 1975). However, the recommended concentration is 0.75% of agar (Sell *et al.,* 1983). Standardisation of inoculum is of prime importance in ensuring reproducible and accurate results in plaque assay. Despite confluent growth being difficult to assess, 400 µl of the sensitive bacterial strain which was grown until the mid-log phase (OD_{600} ~0.4-0.6), provides the ideal confluent growth in 3 ml of overlay agar.

3.4.2 Phage induction

Examination of the mitomycin C induced supernatant from 22 strains of *C. difficile* by electron microscopy identified the presence of twelve distinct phages from different *C. difficile* ribotypes.

These results were encouraging given that attempts to isolate *C. difficile* specific phages with mitomycin C has previously met with little success (Sell *et al*., 1983; Goh *et al*., 2005b; Mayer *et al*., 2008; Horgan *et al*., 2010). The major difference between our studies and these earlier attempts to recover phage was the fact that the mitogen was added to actively growing cells rather than to cells which had reached the stationary phase (Sell *et al*., 1983).

Phage induction in the early (Fortier and Moineau, 2007), mid-log (Shan *et al*., 2012) and stationary (Sell *et al.,* 1983; Goh *et al*., 2005b; Mayer *et al.,* 2008; Horgan *et al.,* 2010) phases suggest that there is no one period of growth when phage induction is likely to be more successful. Furthermore, some induced *C. difficile* strains were found to release phages only following sporulation or after 72 h of incubation (Nale *et al*., 2012).

Interestingly, we were unable to isolate prophages from any of our *C. difficile* strains using norfloxacin, even though this approach has been successfully employed by others (Nale *et al*., 2012). It could be the case that our cultures were static rather than moving, as was the case in the studies by Nale *et al.* (2012). The fact that we saw phage induction with mitomycin C under the same conditions would suggest that it is a more efficient

mitogen, possibly due to its different mechanism of action as described in the introduction of this chapter (Ingrey *et al*., 2003; Shan *et al.,* 2012; Meessen-Pinard *et al.,* 2012).

Our ability to detect the presence of phages was greatly hindered by our inability to identify a suitably susceptible indicator strain that could be used to propagate newly induced prophages. We attempted to employ *C.difficile* NCTC 11204, a susceptible strain previously identified by Meyer and colleagues, but met with little success (Meyer *et al*., 2008). This failure could be due to the fact that insufficient quantities of phages were released following mitomycin C induction, as shown by our EM studies.

The same was true for all of the twenty-two *C. difficile* isolates included in this study; as a consequence we were compelled to employ electron microscopy as a means of screening for the presence of phages. Perhaps we should not have been surprised as the difficulty of finding suitable sensitive strains has been reported by others (Fortier and Moineau, 2007; Nale *et al*., 2012). One possible explanation could be that the induced phages had undergone proteolytic degradation during storage, as was evidenced by the presence of defective phage particles in some of the samples. Alternatively, these phages may have been released at a low frequency due to inefficient induction and as a consequence we were unable to detect their presence due to the way in which the plaque assay had been configured (Hayes, 1968 ; Grabow, 2001; Clokie *et al.,* 2011).

Another possible contributory factor is the presence of the CRISPR system (as described in the introduction, Chapter 1, section 1.16). This system is thought to confer immunity to infection with phages which the bacterial cell has encountered in the past (Barrangou *et al.,* 2007; Deveau *et al.,* 2008) . To date, a number of CRISPR sequences have been identified in the *C. difficile* 630 genome and some of these sequences have also been found within a *C. difficile* 630 prophage (Sebaihia *et al.,* 2006). This could have resulted in our inability to detect an indicator strain or isolate lytic bacteriophages.

The phages isolated in this study showed a clear resemblance to previously isolated *C. difficile* bacteriophages (Chapter 1, table 1.2). Our results supported by previous results of phage isolation (Sell *et al.,* 1983; Mahony *et al.,* 1985 ; Nagy *et al.,* 1991; Goh *et al.,* 2005b, 2007; Govind *et al.,* 2006; Fortier and Moineau, 2007; Mayer *et al.,* 2008; Horgan *et al*., 2010; Meader *et al*., 2010; Sekulovic, *et al*., 2011; Meessen-Pinard *et al*., 2012; Nale *et al*., 2012; Shan *et al.,* 2012) seem to suggest that particular phage families (*Myoviridae* and *Siphoviridae*) are associated with *C. difficile* strains.

3.4.3 Isolation of phages from environmental samples

We were unable to detect the presence of lytic phages with activity against *C. difficile* in any of our environmental samples. This is consistent with other studies attempts to detect and isolate lytic phage against *C. difficile* phages (Goh *et al*., 2005b; Horgan *et al*., 2010) and to date, no lytic *C. difficile* phages have been isolated. It may be the case that *C. difficile* is a strict anaerobe that exists only as spores in the environment and these spores lack receptors which are essential in phage infection; this may explain the absence or rarity of *C. difficile*-specific lytic phages in the environment (Goh *et al*., 2005b). A lysogenic state for phages of spore-forming anaerobic bacteria is convenient, as in this case phages are not affected by the availability of metabolically active anaerobic cells (Stewart and Levin, 1984).

3.5 Conclusions

We were able to induce the expression of *C. difficile* specific lysogenic phages using mitomycin C, but only at low levels. The phages which were isolated were morphologically similar to those reported by others. Our inability to efficiently propagate phages represents a major obstacle to our attempts to develop lytic phages as a potential therapeutic. For this reason, we adopted an alternative strategy, which is described in the next chapter.

Chapter Four

The cloning and expression of the lysin genes of *Clostridium difficile* and its bacteriophages

4.1 Introduction

4.1.1 Endolysin

Endolysins (Lysins for short) are bacteriophage-encoded lytic enzymes that are designed to break down the peptidoglycan bonds in the bacterial cell wall during the terminal stage of the phage life cycle. These enzymes attack the cell wall that is needed to maintain the internal cytoplasmic pressure, which results in the lysis of cell, thereby allowing progeny phage release (Fischetti, 2008).

Endolysins are also capable of causing rapid cell lysis when applied externally to the bacterial cell wall as purified recombinant proteins. This unique ability, along with their species specificity, makes them interesting antibacterial agents, particularly with increasing bacterial drug resistance (Borysowski *et al*., 2006).

In terms of structure, endolysins from Gram-positive bacteria classically have a modular structure (250–400 amino acids) with an N-terminal domain which encodes peptidoglycan hydrolase activity and C-terminal binding responsible for the species specificity of the molecule (Fischetti, 2008). Gram-negative lysins are smaller as they only have the N-terminal enzymatic domain. However, not all lysins conform to this standard structure; for example Gram-positive lysins have been identified which comprise multiple enzymatic domains (Baker *et al*., 2006), while Gram-negative lysins with an N-terminal binding domain have been discovered (Briers *et al*., 2007; Schmitz *et al*., 2010).

Interestingly, although the C-terminal domain is responsible for recognising specific cell wall components, is not always essential for the activity of endolysin. This property was shown in a number of bacterial endolysins including *C. difficile* φCD27 endolysin (Borysowski *et al.,* 2006; Mayer *et al.,* 2011).

To date, several temperate bacteriophages with activity against *C. difficile* have been identified and described in detail (Chapter1, table 1.2). However, only seven phages have been studied at the molecular level, these being φCD119, φCD2 ,φCD27, φMMP03 and φMMP04 from the *Myoviridae* family (Govind *et al.,* 2006; Goh *et al.,* 2007; Mayer *et al.,* 2008; Meessen-Pinard *et al*., 2012) and two phages from the *Siphoviridae* family, namely φCD6356 (Horgan *et al*., 2010) and φCD38-2 (Sekulovic, *et al*., 2011). In addition, most *C. difficile* strains sequenced to date were found to carry a prophage in their genome. In recent studies, prophages have been identified in a large number of clinical isolates of *C. difficile* (Fortier and Moineau, 2007; He *et al*., 2010; Nale *et al*., 2012; Shan *et al.,* 2012)

To date, a single enzymatic class of phage lysin for *C. difficile* has been subject to recombinant expression and analysis. Mayer *et al*. (2008) induced and sequenced φCD27 from *C. difficile* NCTC12727. The gene encoding the lytic enzyme (*cd27l*) was subsequently identified by sequence homology and PCR-cloned into *E. coli* and *Lactococcus lactis* expression plasmids (Mayer *et al.,* 2008). The authors demonstrated that LysCD27 was capable of lysing viable suspensions of all of the *C. difficile* strains tested, with a significantly broader host range than the phage itself, and had no activity against clostridium-like members that are part of the gastrointestinal (GI) tract of organisms.

The majority of phage lysins characterised to date have been identified using the method adopted by Meyer and colleagues (2008). While it is an effective approach, it is time-consuming and dependent on the isolation of sufficient quantities of phage particles to enable subsequent sequencing (Schmitz *et al*., 2010).

Due to our failure to induce and propagate sufficient quantities of *C. difficile* phage, as described in Chapter 3, we were compelled to adopt an alternative bioinformatics-based approach in which we identified potential lysin genes by screening published *C. difficile* genome and bacteriophage sequences for homologues to known lysins. Using this approach we identified, cloned and expressed a 789bp open reading frame (ORF) encoding a putative phage endolysin from the genome of *C. difficile* 630, based on its similarity to previously characterised *C. difficile* phage lysins and its presence in all *C. difficile* strains that have been tested by PCR (Chapter 5) and furthermore five endolysin genes from the following published bacteriophage nucleotide sequences: φCD27, φC2, φCD119, φCD6356 and φCD38-2.

Furthermore, we report the cloning and expression of mutated endolysin sequences, namely LysCD227, which was generated from φCD27 endolysin sequences, in order to investigate the effect of sequence modification on the activity of endolysin.

4.1.2 Aim of this study

The aims of this chapter are:

- To clone and express endolysins from *C. difficile* phage and to determine their ability to lyse vegetative *C. difficile*.
- To clone and express potential lysins from the *C. difficile* genome and to determine their ability to lyse vegetative *C. difficile* and assess the utility of bacterial genomic sequencing in the identification of non-phage origin endolysins.
- To investigate the effect of cell wall binding domain-based modifications on the activity and specificity of endolysin.

4.2 Materials and methods

This project was approved as a Containment Level 1 study by Cardiff University's Genetic Modification Safety Committee (GMSC) and the Health and Safety Executive (HSE).

4.2.1 Bioinformatics methods

4.2.1.1 Bioinformatic analysis of *C. difficile* phage-encoded endolysin sequences

The nucleotide sequences of the endolysin-encoding genes were identified from the National Center for Biotechnology Information (NCBI) database-by employing several database search strategies (Table 4.1). The protein sequences of each endolysin were subject to multiple sequence alignment using T-Coffee (Notredame *et al*., 2000), and functional domains of each endolysin were searched against the Pfam database (http://pfam.sanger.ac.uk/search) (Finn *et al*., 2010). Signal peptide predictions were made using SignalP v4.0 (Petersen *et al*., 2011; Gram-positive organism group) and SOSUIsignal**,** (Gomi *et al*., 2004). The predicted molecular weight (MW) of each endolysin protein was determined using ExPASy tools (http://web.expasy.org/compute_pi/) (Gasteiger *et al*., 2003) and protein solubility was estimated according to Wilkinson-Harrison (Wilkinson and Harrison, 1991; Davis *et al*., 1999; Harrison, 2000) using their web-server (http://www.biotech.ou.edu/).

Bacteriophage	Endolysin	Locus	Gene length (bp)	References
φCD119	**LysCD119**	AAZ32275	816	(Govind *et al.*, 2006)
φC2	**LysC2**	ABE99499	813	(Goh *et al.*, 2005)
φCD27	**LysCD27**	ACH91325	812	(Mayer *et al.*, 2008)
φCD6356	**LysCD6356**	ADK37890	807	(Horgan *et al.*, 2010)
φCD38-2	**LysCD38-2**	AEF56898	807	(Sekulovic *et al.*, 2011)

Table 4.1. The endolysin gene nucleotide sequences used in this study

4.2.1.2 Bioinformatic analysis of *C. difficile* genome-encoded lysin sequences

In the current study, the genomes of the following sequenced strains of *C. difficile* were analysed for the presence of endolysins:

- 630 (#NC_009089)
- NAP08 (#NZ_GG770710)
- QCD-97b34 (#NZ_ABHF02000060)
- QCD76w55 (#NZ_CM000661)
- CIP107932 (#NZ_CM000659)
- QCD-66c26 (#NZ_CM000441)
- ATCC43255 (#NZ_CM000604)
- QCD-23m63 (#NZ_CM000660)
- NAP07 (#NZ_GG770749
- QCD-63q42 (#NZ_ABHD02000060)
- 002-P50-2011 (#NZ_JH379598)
- QCD37x79(#NZ_ABHG02000042)
- 70-100-2010 (#AGAC01000002)
- CD196 (#NC_013315)
- R20291 (#NC_013316)
- BI1 (#NC_017179)
- QCD-32g58 (#NZ_CM000287)

The protein sequences were chosen and analysed using the following criteria:

- The proteins that were designated as endolysin or lysin by the GenBank database (Benson *et al*., 2012). Following identification, each sequence was run against BLASTp (Altschul *et al*., 1990) to identify further potential lysin homologues.
- The proteins that were designated as lysin for germination or cell surface protein were eliminated.
- All protein sequences were subject to Pfam domain analysis (Finn *et al*., 2010).
- Signal peptides at the N-terminal region of each lysin were predicted using SignalP v4.0 (Petersen *et al*., 2011; Gram-positive organism group) and SOSUIsignal**,** (Gomi *et al*., 2004).
- Any newly identified homologues were added to the original list until September 2012.
- Phylogenetic analyses of lysin amino acid sequences was performed using Molecular Evolutionary Genetics Analysis (MEGA) software MEGA5 (Tamura *et al*.,2011) utilising the ClustalW2 program from EBI (http://www.ebi.ac.uk/Tools/msa/clustalw2; Larkin *et al*., 2007; Goujon *et al*., 2010).

4.2.1.3 Bioinformatic analysis of lysin sequences derived from mutated CD27 endolysin

A mutated version of LysCD27 called LysCD227 was generated by introducing a single point mutation to the nucleotide sequence of the C-terminal domain at nucleotide 768 where an aspartic acid was substituted for a glutamic acid. The LysCD27 endolysin was chosen to perform this study on the basis that its activity has been well characterised (Mayer *et al.*, 2008).

To model the potential effect of this mutation on the tertiary structure of the protein, we constructed an in *silico* predictive structural model using Molecular Operating Environment (MOE) (Chemical Computing Group Inc., Montreal, Canada) software (www.chemcomp.com). The models were built by using homology modelling using the structures of the φCD27 endolysin (Protein Data Bank (PDB) code 3QAY (at http://www.pdb.org) (10.1128/JB.00439-11) and *Listeria monocytogenes* phage PSA endolysin (PlyPSA)(PDB code: 1XOV) (10.1016/j.jmb.2006.08.069) as templates. The sequences were aligned and the φCD27 endolysin was used to model the catalytic domain, while the PlyPSA was only used to model the binding domain. The final model was then energy minimized using the Amber 99 force field, until a RMSD gradient of 0.05 Kcal mol^{-1} was reached.

4.2.2 Molecular biological materials and methods

Individual endolysin sequences used in this study were codon optimised to enhance expression from *E. coli* using a commercial algorithm (Genscript Inc.). Sequences were then chemically synthesised and cloned into pUC57, after which they were sub-cloned into pQE30 (QIAgen Inc.) to generate N terminal histidine-tagged fusion proteins.

The constructed plasmids were subsequently transformed into chemically competent *E. coli* M15 [pREP4] and SG13009 [pREP4] as described in Chapter 2 (2.2.4)

Recombinant proteins were then expressed, analysed, and purified under native conditions, as described in Chapter 2 (2.2.5 and 2.2.6).

4.2.3 Microbiological materials and methods

4.2.3.1 Bacterial strain and growth conditions

The bacterial strains employed in this study were described in Chapter 2 (Table 2.1 and table 2.2).

4.2.3.2 Biochemical properties of endolysins

To characterise the activity of each recombinant lysin, a range of biochemical properties were determined using the Bioscreen method, which included the effect of suspension media, pH, heat, concentration and thermostability on activity and stability.

To determine the effect of bacterial suspension media on the enzymatic activity, cells of *C. difficile* were suspended in PBS, BHI broth and 20mM sodium phosphate buffer and treated with 7.0μg/ml of recombinant lysin at 37°C for 1 h.

To determine the effect of different concentrations of purified lysin on cell lysis, the following concentrations were tested: 0.875, 1.75, 3.5 and 7.0 μg/ml.

The activity profile under different pH conditions was examined in PBS buffer adjusted to pH between 4 and 9 and suspensions were treated with 7.0μg/ml protein concentration at 37°C.The sensitivity to heat of each lysin was also assessed by pre-incubation of the enzyme for 30 min at a range of temperatures (37,45,50,60,70°C) prior to mixing with a cell suspension of *C. difficile* cells in 20mM sodium phosphate at pH7.

The thermostability of the recombinant lysins stored at 37°C and room temperature for a period of 2 weeks was also determined.

4.2.3.3 Biological activity of endolysins

The lytic activity of each endolysin was determined by measuring the decrease in optical density of a *C. difficile* cell suspension following the addition of the endolysin using a method based on that previously described by Mayer and colleagues (2008). Briefly, cultures of selected isolates of *C. difficile* and other Gram positive strains were grown anaerobically-aerobically as appropriate at 37°C to mid-log phase (OD_{600} 0.4-0.6) in 45 ml of BHI broth and were then harvested by centrifugation (5,000xg for 5 min); pellets were then resuspended in 20mM sodium phosphate buffer (pH 7) to an OD_{600} of 0.8–1.2. A 270 µL aliquot of bacterial cell suspension was then added to the well of a Bioscreen plate (Honeycomb Bioscreen plates: Thermo Scientific, UK) and mixed with 20 µL of recombinant lysin (~7 µg). Negative controls comprised of cell suspensions mixed with 20 µL of buffer. Following inoculation, the Bioscreen plates were incubated aerobically at 37°C with shaking for 2 sec on a Bioscreen C (Labsystems, Helsinki, Finland), and monitored for changes in Optical density OD_{600} at 2 min intervals over a 1 h period.

Lytic activity was calculated using the following equation: 100 x (OD600nm of control without lysin - OD600nm of reaction mixture)/OD600nm of control without enzyme.

The activity of lysins against *C. difficile* cells on CDMN plates were also determined, using a method described previously by Simmons and co-workers (2010), with slight modifications. Plates were prepared as follows: 1 ml of *C. difficile* BHI broth, which was left to grow to the mid-log phase (OD_{600} 0.4 to 0.6), was poured directly on to the pre-reduced CDMN agar. Excess liquid was removed and the plate was allowed to dry; then, 20 μL of recombinant lysin (7 μg) was added. Subsequently, plates were incubated anaerobically at 37°C for 24 h and then checked for lysis zones. PBS buffer was also spotted onto plates as a negative control.

4.3 Results

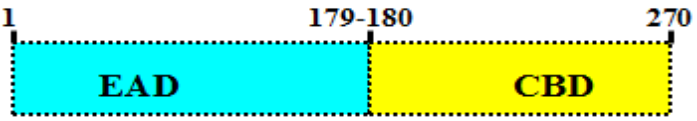
4.3.1 Bioinformatic analysis of *C. difficile* phage-encoded endolysin sequences

The amino acid sequences of the endolysins from the published *C. difficile*-specific phage φCD119, φC2, φCD27, φCD6356 and φCD38-2 were downloaded from the NCBI database and subjected to bioinformatic analysis.

All had a similar modular structure composed of a single enzymatically-active domain (EAD) and the cell-wall binding domain (CBD). Structural analysis using Pfam domain analysis showed that all of the phage endolysins belonged to the amidase-3 (*N*-acetylmuramoyl-L-alanine amidase) family. Surprisingly, comparison of the amino acid sequence of the N -terminal regions found that rather than being highly conserved, as one might expect given that they all belonged to the same functional group, they showed a greater degree of diversity than the C terminal domain (Figure 4.1).

Interestingly, although CBD is designed to act on the same bacterial cell wall, each endolysin appears to carry a distinct CBD, suggesting that several different ligands of that cell compartment might be targeted by the endolysins. However, Pfam analysis failed to recognise any known CBD in their C-terminal regions (Figure 4.1).

A



B

```
LysCD27         1 MKICITVGHSILKSGACTSADGVVNEYQYNKSLAPVLADTFRKEGHKVDVIICPEKQFKTKNEEKSYKI   69
LysC2           1 MKICITVGHSILKSGACTSADGVVNEYQYNKSLAPVLADTFRKEGHKVDVIICPEKQFKTKNEEKSYKI   69
LysCD119        1 MKIGVNCGHTK--TGAGSGAIGKINESIETRNVGYKVIDKLKTLGNNV--VDCTIDKASTQSECLSKIA   65
LysCD6356       1 MKVVLTAGHTL--TGKGTGATGYINEGKENRILMDLVVKWLKKGGATV--YSGKVDKS---NNYLSEQC   62
LysCD38-2       1 MRVALTAGHTL--TGKGTGATGYINEGTENRILMDLVVKWLKKGGATV--YSGKVDKS---NNYLAEQC   62
                  *:: :. **:   :*  :.* * :**   .: :   : . ::. * .*       .:    .:  :

LysCD27        70 PRVNSGGYDLLIELHLNASNGQ--GKGSEVLYYSNKGLEYATRICDKLGTV-FKNRGAKLDK-RLYILN  134
LysC2          70 PRVNSGGYDLLIELHLNASNGQ--GKGSEVLYYSNKGLEYATRICDKLGTV-FKNRGAKLDK-RLYILN  134
LysCD119       66 TQANRQDLDWFISIHFNAGK----GRGCEVYTYKGKQYQDAIDVCKKISDLGFTNRGVKDGS-GLYVVK  129
LysCD6356      63 QIANKRNVDLAVQIHFNANKTTINPMGTETIYKTNNGKVYAERVNEKLATV-FENRGAKSDARGLYWLR  130
LysCD38-2      63 QIANKQNVDVAVQIHFNADHTTLNVMGTETIYKTNNGKVYAERVNEKLATI-FKNRGAKSDARGLYWLS  130
                    .*  . *  :.:*:**.:      * *.   ..:    *  : .*:. : * ***.* .   ** :

LysCD27       135 SSKPTAVLIESFFCDNKEDYDKAKKLGHEGIAKLIVEGVLNKNINN-EG---VKQMYKHTIVYDGEVDK  199
LysC2         135 SSKPTAVLIESFFCDNKEDYDKAKKLGHEGIAKLIVEGVLNKNINN-EG---VKQMYKHTIVYDGEVDK  199
LysCD119      130 KTKAKSMLIEVCFVDSEDA-NKYLSLGADKLATAIVEAIT-KHISSAEEN--NYNRYKHTIVYSGD-DK  193
LysCD6356     131 HTKAPAILIEVCFVDSKADTDYY-IRHKDIVAKLIAEGILNKNIDN-KENGEDKKMYKHTIVYDGEVDK  197
LysCD38-2     131 HTKAPAILIEVCFVDSKADTDYY-IRHKDIVAKLIAEGILNKTIDN-KENSEDKKMYKHTIVYDGEVDK  197
                   :*. ::***  * *.:   :       : :*. *.*.:  * *.. :      : *******.*: **

LysCD27       200 ISATVVGWGYNDGK--ILICDIKDYVPGQTQNLYVVGGGACEKISSI---TKEKFIMIKGNDRFDTLYK  263
LysC2         200 ILANVLSWGYSPSK--VLVCDIKDYVPGQTENLYVVGGGACEKISSI---TKEKFIMIKGNDRFDTLYK  263
LysCD119      194 VSADILGLYYKRKKESYLVTDIKDYKPHRTQNLYVIGGVTCNKMKEMSKTTGEKFTQLYGNDVWSTMDK  262
LysCD6356     198 ILATVLGWGYSSSK--VLVCDIKDYIPGQTQNLYVVGGGACEKIGSI---TKEHYTTIKGNDRFDTLHK  261
LysCD38-2     198 IPATVVGWGYNDGK--ILICDIKDYVPGQTQNLYVVGGGACEKIGSM---TKEKFTMIKGNDRFDTLYK  261
                  : * ::.  *.  *   *: ***** * :*:****:** :*:*: .:   * *::  : *** :.*: *

LysCD27       264 ALDFIN--R  270
LysC2         264 ALDFIN--R  270
LysCD119      263 AIEFVKEKL  271
LysCD6356     262 VLDFIK--K  268
LysCD38-2     262 ALDFID--R  268
                  .::*:.
```

Figure 4.1. Domain architecture of endolysins (A) Schematic representation of endolysin domain organisation. The blue rectangle delimits the enzymatically-active domain (EAD) of N-acetylmuramoyl-L-alanine amidase (amidase3) and the yellow rectangle denotes the cell-wall binding domain (CBD), using previously characterised endolysin (Mayer *et al*., 2011). (B)Alignment of the endolysin sequences taken from NCBI was conducted using T-Coffee. All endolysins carry N-acetylmuramoyl-L-alanine amidase (amidase3) at their N-terminal region. Regions of sequence homology are indicated by (*) and similar amino acids (:).

4.3.2 Bioinformatic analysis of *C. difficile* genome-encoded lysin sequences

Putative lysin sequences located within the genome sequences of the *C. difficile* isolates deposited in GenBank were screened for the presence of putative proteins with the following properties:

(A) A N-terminal enzymatic domain and a C-terminal binding region with the absence of additional domains which may encode non-lysis function (Schmitz *et al*., 2010).

(B) A N-terminal sequence lacking a signal peptide which is often present in cell autolysins (Blackman *et al.,* 1998; Vollmer *et al*., 2008; Schmitz *et al*., 2010).

In total, 87 proteins were identified from 17 *C.difficile* genomes which fit the above criteria (Table 4.2). Of this total, 52 were predicted to possess an N-acetylmuramoyl-L alanine amidase at the N-terminal. This enzyme cleaves the amine bond between the L-alanine of the peptide chain and N-Acetylmuramic acid (MurNAc). The majority (51) belonged to the N-acetylmuramoyl-L-alanine amidase family 3 (PF01520), while the remaining protein was from the N-acetylmuramoyl-L-alanine amidase family 2 (PF01510).

The remaining 35 proteins were predicted to carry an Mannosyl-glycoprotein endo-beta-N-acetylglucosaminidases (Glucosaminidase) (PF01832), which hydrolyses peptidoglycan bonds between *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc). Interestingly, Pfam failed to identify any conserved domains in the C terminal region of all of the proteins.

Source	Gene accession no	Domains
BI1	YP_006198983	Amidase_3
	YP_006199265	Amidase_3
	YP_006198983	Amidase_3
	YP_006198687	Amidase_3
	YP_006198091	Glucosaminidase
QCD-97b34	ZP_05386798	Amidase_3
	ZP_05384969	Amidase_3
	ZP_05385260	Amidase_3
	ZP_05384675	Amidase_3
	ZP_05385264	Glucosaminidase
R20291	YP_003218580	Glucosaminidase
	YP_003218307	Amidase_3
	YP_003218576	Amidase_3
	YP_003217960	Amidase_3
	YP_003217390	Glucosaminidase
CD196	YP_003215071	Glucosaminidase
	YP_003214798	Amidase_3
	YP_003215067	Amidase_3
	YP_003214516	Amidase_3
	YP_003213943	Glucosaminidase
050-P50-2011	ZP_17072773	Amidase_3
	EHJ24718	Amidase_3
	EHJ26281	Amidase_3
	EHJ25719	Glucosaminidase
	EHJ26800	Glucosaminidase
	EHJ31012	Glucosaminidase
	EHJ31282	Glucosaminidase
	ZP_17076819	Amidase_3
	ZP_17075489	Amidase_3
	ZP_17074811	Amidase_3
	EHJ29023	Amidase_3
	EHJ25043	Amidase_3
	EHJ27323	Amidase_3
QCD-63q42	ZP_05332329	Amidase_3
	ZP_05331981	Amidase_3
	ZP_05331962	Amidase_3
	ZP_05331929	Amidase_3
	ZP_05331399	Amidase_3
	ZP_05330908	Amidase_3
	ZP_05330302	Glucosaminidase
	ZP_05329048	Glucosaminidase
	ZP_05330299	Amidase_3
	ZP_05330002	Amidase_3
QCD-76w55	ZP_05355902	Amidase_3
	ZP_05356497	Glucosaminidase
	ZP_05355293	Glucosaminidase

Source	Gene accession no	Domains
QCD-37x79	ZP_05399195	Amidase_3
	ZP_05399316	Amidase_3
	ZP_05399324	Amidase_3
ATCC43255	ZP_05351369	Amidase_3
	ZP_05350750	Amidase_3
	ZP_05350132	Glucosaminidase
	ZP_05352351	Amidase_2
	ZP_05351060	Amidase_3
	ZP_05350856	Amidase_3
	ZP_05351372	Glucosaminidase
NAP08	ZP_06892601	Amidase_3
	ZP_06892148	Glucosaminidase
	EFH07492	Glucosaminidase
	ZP_06891530	Glucosaminidase
	EFH08309	Glucosaminidase
NAP07	ZP_06903479	Glucosaminidase
	ZP_06901942	Glucosaminidase
	EFH15484	Glucosaminidase
	EFH16819	Glucosaminidase
	ZP_06903476	Amidase_3
70-100-2010	EHJ38445	Glucosaminidase
	EHJ40888	Glucosaminidase
	ZP_17077665	Amidase_3
	EHJ39933	Amidase_3
630	YP_001087453	Amidase_3
	YP_001089407	Amidase_3
	YP_001088405	Amidase_3
	CAJ67875	Glucosaminidase
	CAJ69071	Glucosaminidase
CIP107932	ZP_05322060	Amidase_3
	ZP_05322650	Glucosaminidase
	ZP_05321453	Glucosaminidase
QCD-66c26	ZP_05271664	Amidase_3
	ZP_05271057	Glucosaminidase
	ZP_05272257	Glucosaminidase
QCD-23m63	ZP_05402087	Amidase_3
	ZP_05400420	Glucosaminidase
	ZP_05401584	Glucosaminidase
QCD-32g58	ZP_07406968	Glucosaminidase
	ZP_07406964	Amidase_3
	ZP_07406453	Amidase_3
	ZP_07406726	Amidase_3

Table 4.2. Domain structure of 87 endolysin and lysin-like proteins from the 17 *C. difficile* strains. The 87 protein sequences were identified representing two classes of endolysin: Amidase and Glucosaminidase.

Subsequently, we decided to focus on protein sequences which carried amidase lysin class 3 domains, since all of the *C. difficile* phages isolated to date also carried this type of endolysin at their N-terminal region.

The evolutionary analysis was conducted by MEGA5 and constructed using the Neighbor-Joining method (Figure 4.2) and Maximum Likelihood method (Appendix 2) .The evolutionary history showed that all *C. difficile* amidase lysin amino acid sequences clustered in two groups: group 1, which contains endolysin sequences from *Siphoviridae* phages, including those from *C. perfringens* phages φ39-O, φCP13O, φCP26F, φCP34O and φCP9O, and from a *C. sporogenes* phage φ8074-B1 (Oakley *et al*., 2011; Mayer *et al*., 2012). Interestingly, the endolysin sequence of *C. difficile* phage CD119, which belongs to the *Myoviridae* group, is also present in this group.

Group 2 comprises sequences derived from phages belonging to the *Myoviridae* family.

Many of the proteins identified by this process were encoded in regions of the genome thought to correspond to prophages or prophage remnants. In addition, some of these sequences showed considerable homology to previously described phage endolysins. For example, the genome of *C. difficile* QCD-63q42 contains two protein sequences (ZP_05331399 and ZP_05331929) which share 100% amino acid identity with the endolysins φCD27 and φCD38-2, respectively. Furthermore, the lysin sequence from *C. difficile* QCD-37x79 (ZP_05399316) shares 100% amino acid identity with the φCD38-2 endolysin sequence (Figure 4.2). Interestingly, these genes are located in a region of the genome which does not contain other putative phage-derived genes.

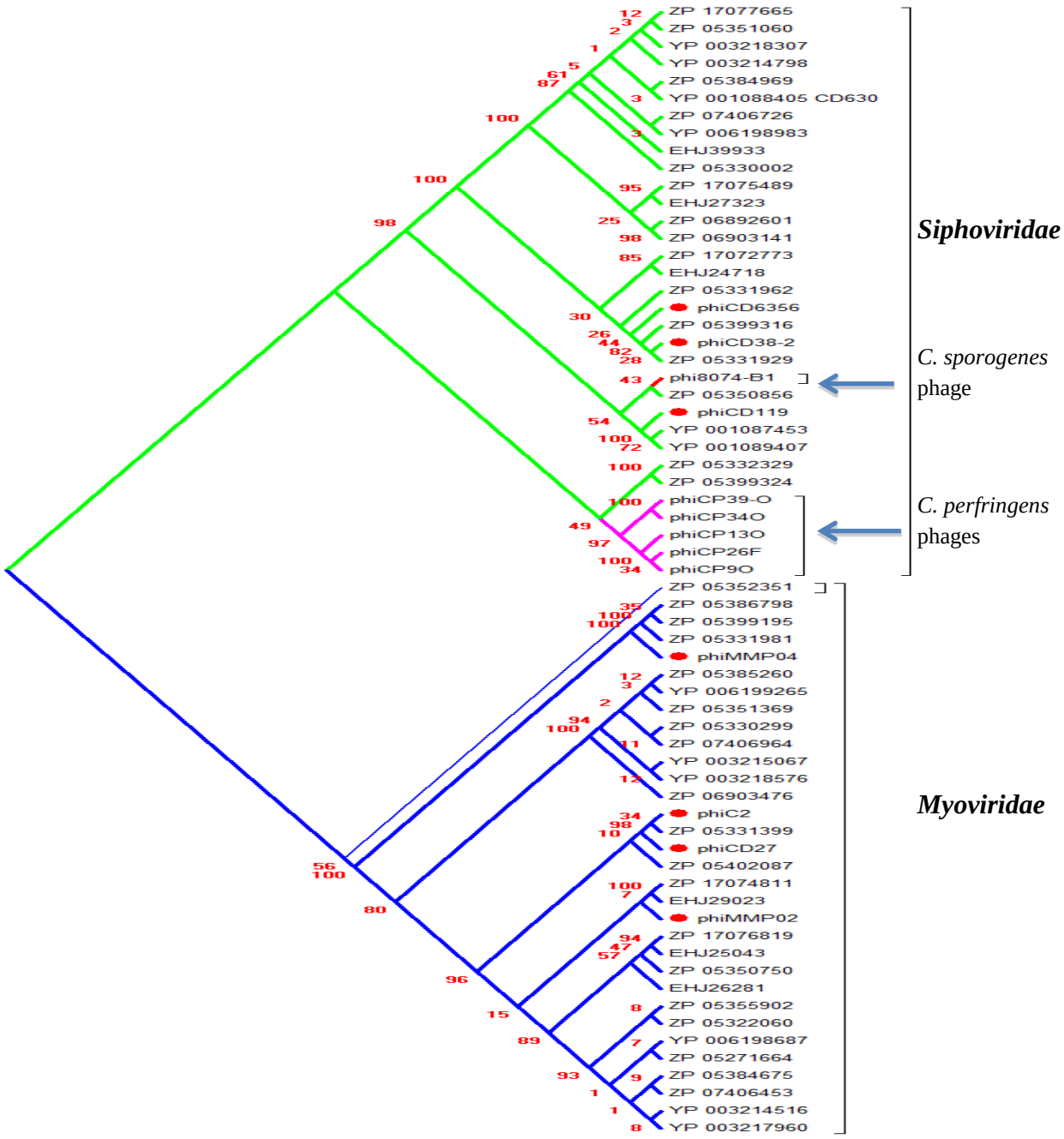


Figure 4.2. Evolutionary relationship of lysin sequences encoded peptidoglycan hydrolase, N-acetylmuramoyl-L-alanine amidase from *C. difficile* genomes and phages.

The evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. The evolutionary distances were computed using the Poisson correction method based on their amino acid sequences. The analysis involved 65 amino acid sequences including *C. difficile* phage endolysins (φCD119, φCD2, φCD27, φMMP03,φMMP04, φCD6356 and φCD38-2), *C. perfringens* phage endolysin (φ39-O, φCP13O, φCP26F, φCP34O, φCP9O) and *C. sporogenes* phage (φ8074-B1). All positions containing gaps and missing data were eliminated. There were a total of 180 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

This is not always the case; for example, ZP_05352351, the lysin carried by *C. difficile* ATCC43255 contains an amidase class 2 domain which was followed by a gene predicted to encode a holin, suggestive of the fact that this strain had previously been infected by a class of phage ,which has not been previously described in *C. difficile*. The presence of numerous lysins in the genome of multiple *C. difficile* strains located in regions lacking any other phage-related sequences suggests that these genes may represent evolutionary relic of past phage infection.

C. difficile strain 630 contains three putative endolysin sequences located in locus (YP_001087453, YP_001089407 and YP_001088405), which are homologous to previously described phage endolysins. Two of these genes (YP_001087453 and YP_001089407) are identical and share 97% amino acid identity with the endolysin from φCD119. They are located in a region of the genome of 630 which has been annotated as prophage 1 and 2 (prophage 1 CD0904-0979 and prophage 2 CD2889-2952 (Sebaihia *et al*., 2006; Goh *et al.,* 2007; Monot *et al.,* 2011).

The remaining sequence (YP_001088405) is located outside of these prophage regions and was found to have the greatest homology (62% amino acid identity) with the endolysin from φCD6356.

Interestingly , BLASTn searches revealed that the nucleotide sequence of this gene appears to be highly conserved in that a homology (96% homology or higher) has been found in all nine strains of *C. difficile* sequenced to date (Sebaihia *et al*., 2006; Stabler *et al.,* 2009; He *et al.,* 2010; Monot *et al.,* 2011).

We further confirmed the universal nature of this gene using a PCR based approach, which identified the presence of a homologous gene in 85 isolates of *C. difficile* (Chapter 5). For this reason, the lysin encoded by YP_001088405 was selected and expressed as a recombinant protein.

4.3.3 Bioinformatic analysis of the mutated endolysin (LysCD227)

During the course of this study, we generated a mutated version of the LysCD27 lysin called LysCD227, which differed from LysCD27 due to a single point deletion in the nucleotide sequence at the C-terminal domain (T at position 768). This mutation resulted in two endolysins which shared 100% homology in their enzymatically-active domain (EAD), but only 84% homology within the cell-wall binding domain (CBD) (fig 4.3).

```
LysCD27         MKICITVGHSILKSGACTSADGVVNEYQYNKSLAPVLADTFRKEGHKVDVIICPEKQFKT 60
LysCD227        MKICITVGHSILKSGACTSADGVVNEYQYNKSLAPVLADTFRKEGHKVDVIICPEKQFKT 60
                ************************************************************

LysCD27         KNEEKSYKIPRVNSGGYDLLIELHLNASNGQGKGSEVLYYSNKGLEYATRICDKLGTVFK 120
LysCD227        KNEEKSYKIPRVNSGGYDLLIELHLNASNGQGKGSEVLYYSNKGLEYATRICDKLGTVFK 120
                ************************************************************

LysCD27         NRGAKLDKRLYILNSSKPTAVLIESFFCDNKEDYDKAKKLGHEGIAKLIVEGVLNKNINN 180
LysCD227        NRGAKLDKRLYILNSSKPTAVLIESFFCDNKEDYDKAKKLGHEGIAKLIVEGVLNKNINN 180
                ************************************************************

LysCD27         EGVKQMYKHTIVYDGEVDKISATVVGWGYNDGKILICDIKDYVPGQTQNLYVVGGGACEK 240
LysCD227        EGVKQMYKHTIVYDGEVDKISATVVGWGYNDGKILICDIKDYVPGQTQNLYVVGGGACEK 240
                ************************************************************
LysCD27         ISSITKEKFIMIKGNDRFDTLYKALDFINR 270
LysCD227        ISSITKEKFIMIKGNEDLIHFIKHWILLID 270
                **************: :  : *   ::
```

Figure 4.3. Comparison of the endolysin sequences of LysCD27 (from φCD27), and mutated LysCD227. The enzymatically-active domain (EAD) and the cell-wall binding domain (CBD) are marked with blue and yellow colour, respectively. Regions of sequence homology are indicated by (*). EAD regions are identical (100% homology), CBD regions are only 84%. The sequences were aligned using ClustalW2.

In an attempt to predict the effect of this mutation on the structure of the endolysin, we constructed a 3D *in silico* model for these endolysins (Figure 4.4).

Our modelling, based on previously published crystal structure data, revealed that the enzymatic domain consisted of five helices flanked by six-twisted β-sheets connected to the cell wall binding domain, which is formed by 7 β-sheets (Mayer *et al.,* 2011) . As can be seen from the model, the single point mutation in LysCD27 was predicted to cause a structural change to make the protein more lipophilic.

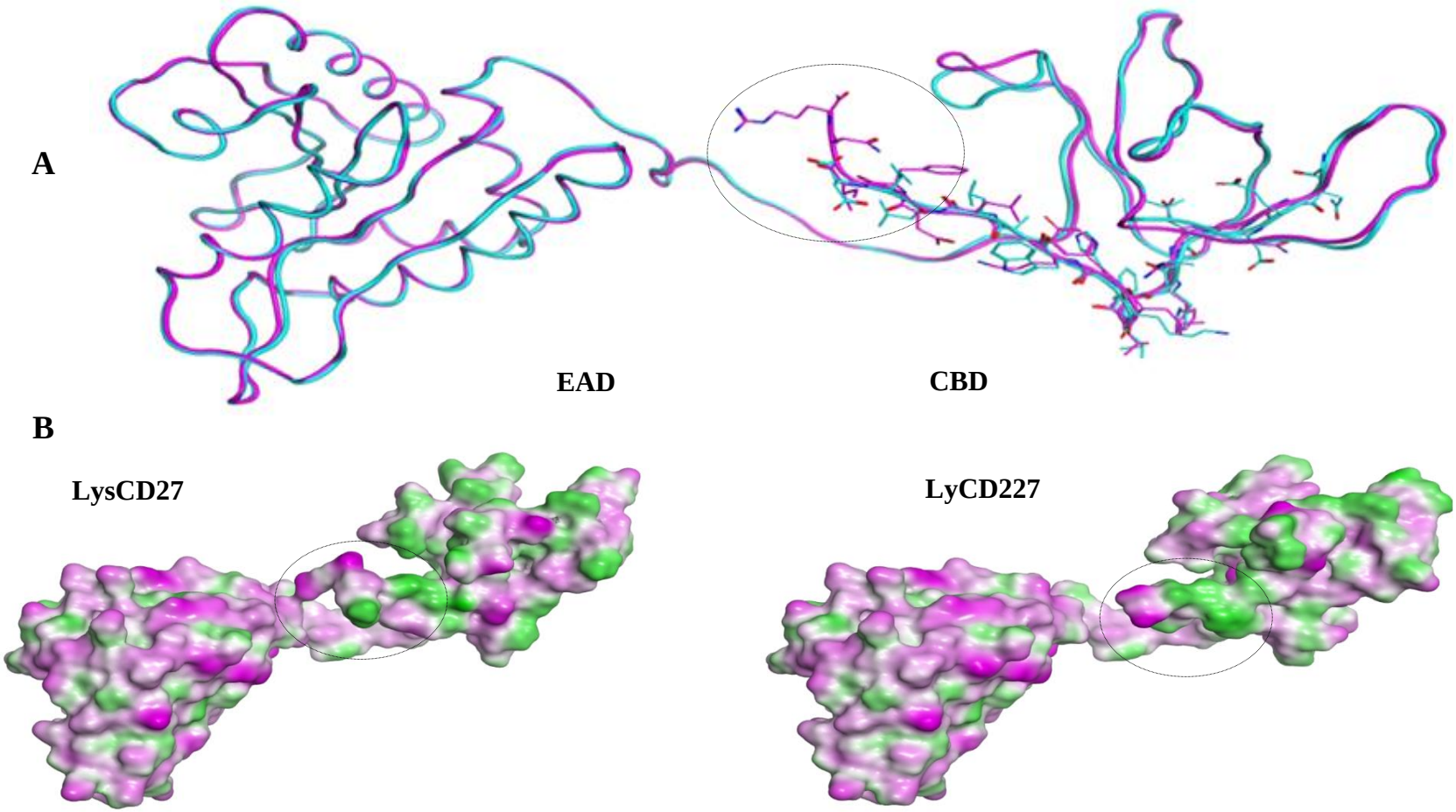


Figure 4.4. Structural comparison of endolysins (LysCD27 and LysCD227 . **A:** Alignment of endolysins showing the enzymatically-active domain (EAD) and the cell-wall binding domain (CBD). LysCD27 is purple and LysCD227 is blue. The dashed circle highlights the area containing the difference.

B: Showing the 3D structure of endolysins displaying the hydrophilic (purple), lipophilic (green) and neutrophilic (white) residues.

4.3.4 Cloning and expression of lysins of *E. coli*

4.3.4.1 DNA Manipulation

Individual protein sequences were codon optimised for expression in *E. coli* by Genscript Inc. Gene sequences were synthesised and cloned into pUC57 and then sub-cloned into the QIAgen expression vector pQE30, where they were placed downstream of a histidine affinity tag and under the control of an inducible T5 promoter. Following insertion, plasmids were sequenced to ensure the correct orientation of the gene, after which the recombinant plasmid was transformed into competent cells of *E. coli* M15[pREP4] and SG13009[pREP4] for protein expression. Positive transformants were identified by culture on LB agar plates containing ampicillin (100 µg/ml) and kanamycin (25 µg/ml).

Endolysin	Gene length (bp)	Protein length (aa)	Predicted MW (kDa)	Protein Solubility prediction
LysCD119	816	271	30.16	81.0% insolubility
LysC2	813	270	30.16	86.7% insolubility
LysCD27	813	270	30.14	85.9% insolubility
LysCD6356	807	268	29.9	79.5% insolubility
LysCD38-2	807	268	29.84	69.7% insolubility
***LysCD227**	812	270	30.12	81.8% insolubility
****LysCD630**	789	263	28.57	72.5% insolubility

Table 4.3. Showing the details of endolysin sequences used in this study, and the predicted MW which was generated using ExPASy tools (http://web.expasy.org/compute_pi/). *LysCD227 is the modified endolysin generated from phage CD27 endolysin, as described before in the methods section. **LysCD630 is the putative endolysin from *C. difficile* 630.

While the transformation efficiency of the pQE30 plasmid containing the endolysin genes was lower than that seen for the positive control plasmid (pQE-40), it was still efficient enough to yield the desired transformants (Table 4.4). M15 seems to provide a statistically significantly better transformation ratio than SG13009; therefore, *E. coli* M15 was select for further work.

Endolysin construct	Transformants /µg of plasmid DNA	
	M15	SG13009
LysCD27	$167.5\pm11X10^3$	$150.8\pm6.9X10^3$
LysC2	$192.3\pm5X10^3$	$214.4\pm2.2X10^3$
LysCD119	$321.4\pm24X10^3$	$498.8\pm22X10^3$
LysCD630	$385.5\pm20.6X10^3$	$229.8\pm8X10^3$
LysCD6356	$305.5\pm4.2X10^3$	$266.4\pm2.7X10^3$
LysCD38-2	$267.6\pm7.8X10^3$	$251.6\pm4.1X10^3$
LysCD227	$279.9\pm8.1X10^3$	$222.3\pm2.5X10^3$
pQE40	$341.5\pm459X10^3$	$286\pm232X10^3$

Table 4.4. Transformation efficacy in *E. coli* M15 and SG13009 ,using 100 µl of the overnight culture plated out on LB-agar plates for 24h (readings represent the mean of three results± sd). pQE-40 control plasmid, which expresses 6xHis-tagged DHFR, is optimised for a high-level of transformation and expression into these *E.coli* strains.

4.3.4.2 Optimisation of recombinant protein expression, analysis and purification

As described in Chapter 2, expression cultures were grown in LB broth containing ampicillin (100 μg/ml) and kanamycin (25 μg/ml) at 37°C and 300 rpm to an OD_{600nm} of 0.6. Cultures were induced by the addition of IPTG to a final concentration of 1 mM. Under these conditions, the recombinant proteins formed insoluble inclusion bodies, so the culture temperature was reduced to 27°C, which is a temperature that has been reported to reduce the formation of inclusion bodies (Zimmer *et al.*, 2002; Dhalluin *et al.*, 2005).

The expression of recombinant proteins following induction was monitored over time and the results revealed that approximately 4 h of incubation was required for sufficient protein production (Figure. 4.5 B, 4.6 B). Following 4h of incubation, expression cultures were harvested and the recombinant proteins were purified under native conditions using Nickel–nitrilotriacetic acid (Ni–NTA) affinity chromatography.

Protein bands with a molecular weight of approximately (~30 kDa), the expected size range of the recombinant lysins (as described in table 4.3), were visualised in SDS-PAGE gels stained with Coomassie blue stain following induction. Protein purification was also performed with cells harbouring the empty vector (pQE40), with the eluate serving as a negative control in the subsequent activity assays.

Typically, the protein yields obtained from 100 ml cultures of *E. coli* M15 were approximately 3, 2.5, 4, 1.5, 3.5, and 2 mg for endolysins of LysCD119, LysC2, LysCD27, LysCD6356, LysCD630, LysCD227 and LysCD38-2, respectively, as determined using the BCA protein assay (section 2.2.7).

In figures 4.5 and 4.6 we compare recombinant protein induction and purification from phage origin endolysin (LysC2) and non-phage origin endolysin LysCD630; it can be seen that LysCD630 was induced and extracted in the same way as other phage endolysins extracted with almost the same yield.

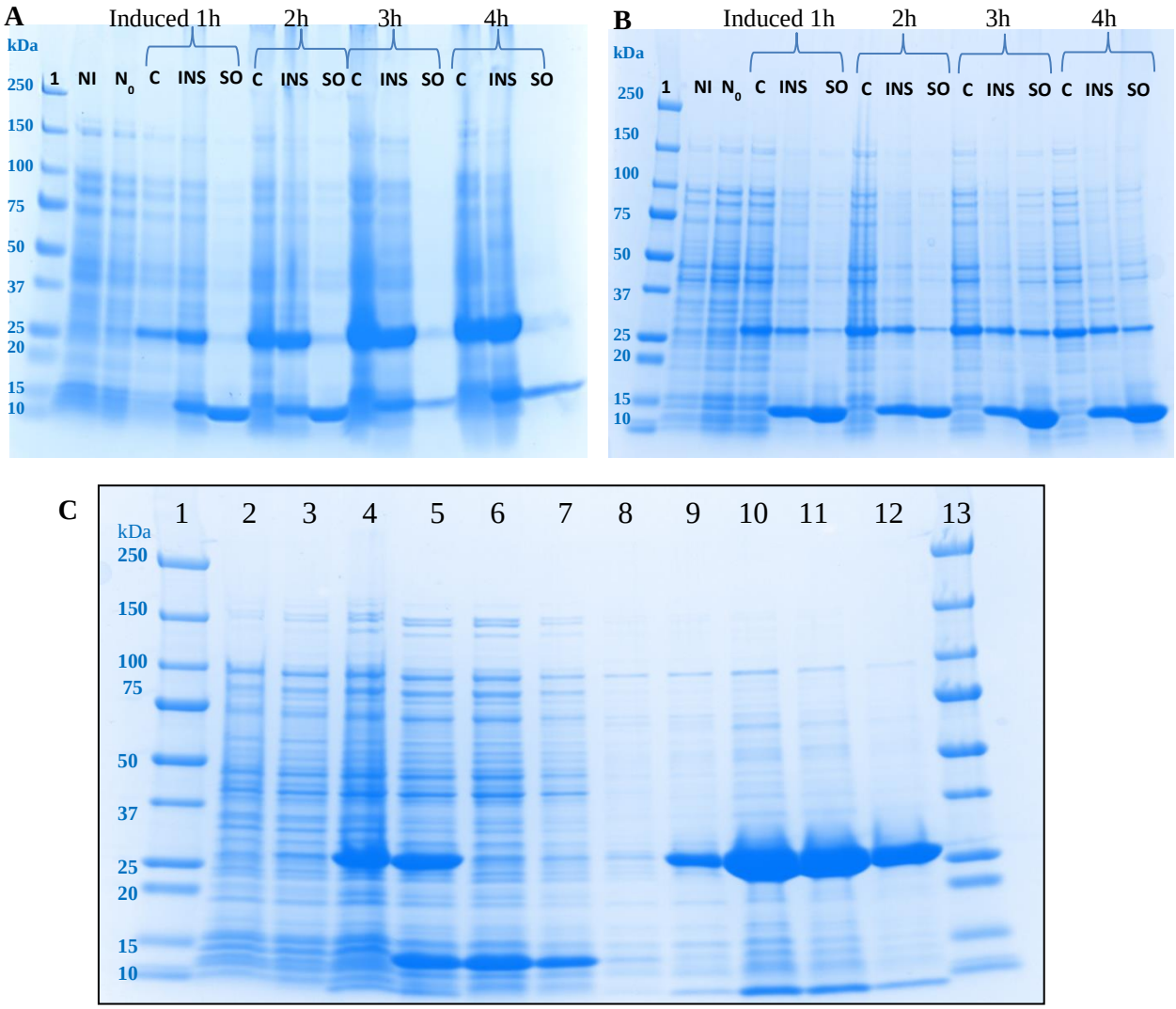


Figure 4.5. LysC2 endolysin extraction , (predicted size 30 kDa)

A) Time-course analysis of LysC2 at 37°C. Aliquots were removed post-induction at the time indicated and analysed by SDS-PAGE and Coomassie blue stain. 1: Molecular weight marker. NI: non-induced control, N_0:0 hour induced, C: crude, INS: insoluble, SO: soluble.

B) Time-course analysis of LysC2 at 27°C. Aliquots were removed post-induction at the time indicated and analysed by SDS-PAGE and stained with Coomassie blue stain. 1: Molecular weight marker. NI: non-induced control, N_0: 0 hour induced, C: crude, INS: insoluble, SO: soluble.

C) LysC2 protein extraction: Lanes: 1, marker; 2, non-induced culture; 3, 1h induced culture; 4, total protein extracts after 4 h of induction with IPTG; 5, column flow-through; 6, primary wash effluent; 7, secondary wash effluent; 8, 1st eluate (E1); 9, 2nd eluate (E2); 10, 3rd eluate (E3); 11, 4th eluate (E4); 12, buffer exchange 20mM NaPhos pH 6; 13, marker.

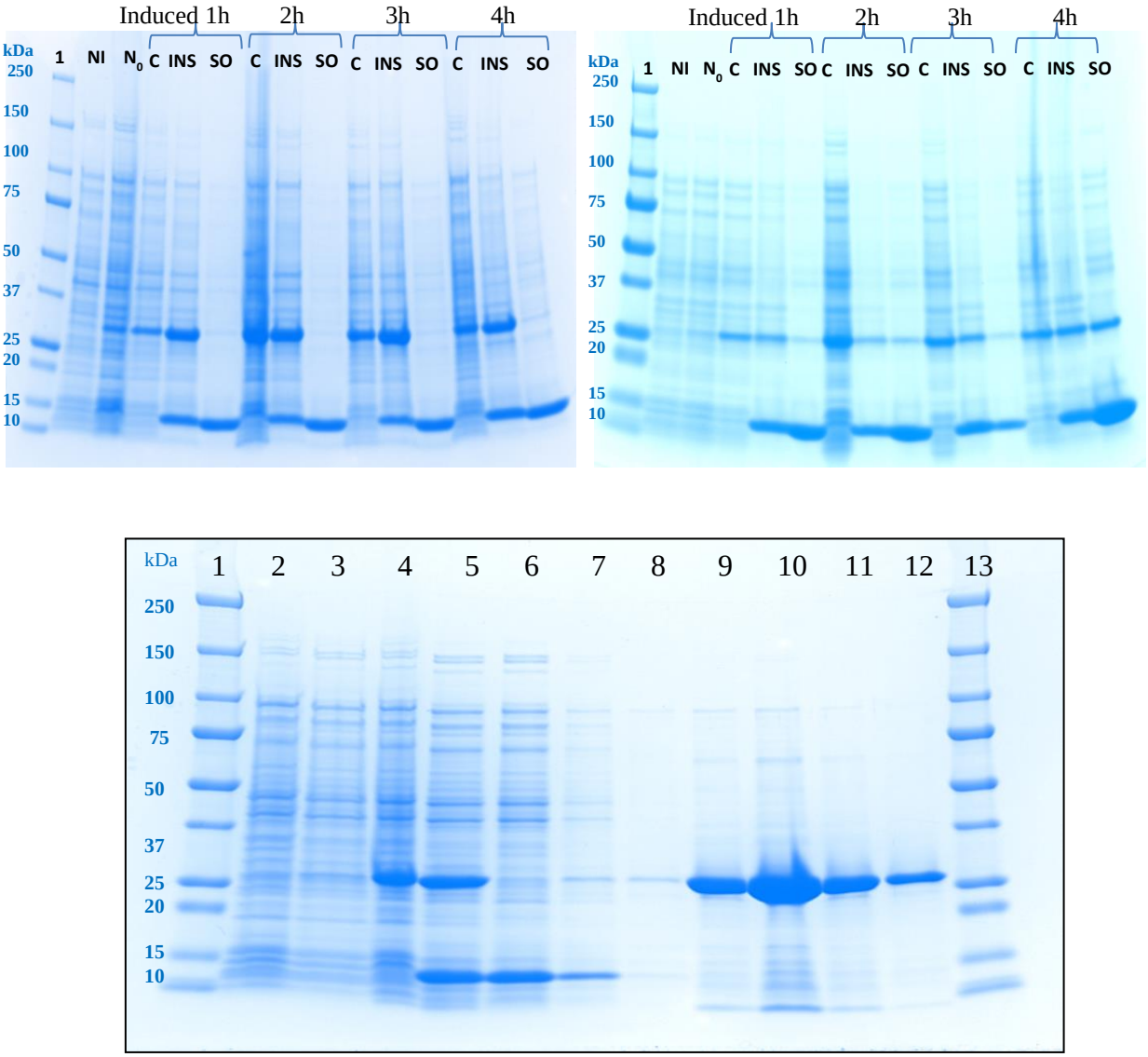


Figure 4.6. LysCD630 putative endolysin extraction , (predicted size 28 kDa)

A) Time-course analysis of LysCD630 protein at 37°C. Aliquots were removed post-induction at the time indicated and analysed by SDS-PAGE and Coomassie blue stain. 1: Molecular weight marker, NI: non-induced control, N_0: 0 hour induced, C: crude, INS: insoluble, SO: soluble.

B) Time-course analysis of LysCD630 protein at 27°C. Aliquots were removed post-induction at the time indicated and analysed by SDS-PAGE and stained with Coomassie blue stain. 1: Molecular weight marker .NI: non-induced control, N_0: 0 hour induced, C: crude, INS: insoluble, SO: soluble.

C) LysCD630 protein extraction: Lanes: 1, marker; 2, non-induced culture; 3, 1h induced culture; 4, total protein extracts after 4 h of induction with IPTG; 5, column flow-through; 6, primary wash effluent; 7, secondary wash effluent; 8, 1st eluate (E1); 9, 2nd eluate (E2); 10, 3rd eluate (E3); 11 , 4th eluate (E4); 12, buffer exchange 20mM NaPhos pH 6; 13, marker.

4.3.4.3 Protein detection: Western blot

The identity of the recombinant endolysin proteins was confirmed by probing western blots (as described in Chapter 2 section 2.2.9). Probing was accomplished with an antibody, Penta-His, which recognised the N-terminal histidine tag antibody and the secondary antibody, which detected the reciprocal proteins on the nitrocellulose membrane by producing the dark band signal, as shown in figure 4.7 .The signals were seen at the membrane region corresponding to a molecular weight of ~30 kDa. A protein marker (6x His-protein ladder) confirmed the size of the signals.

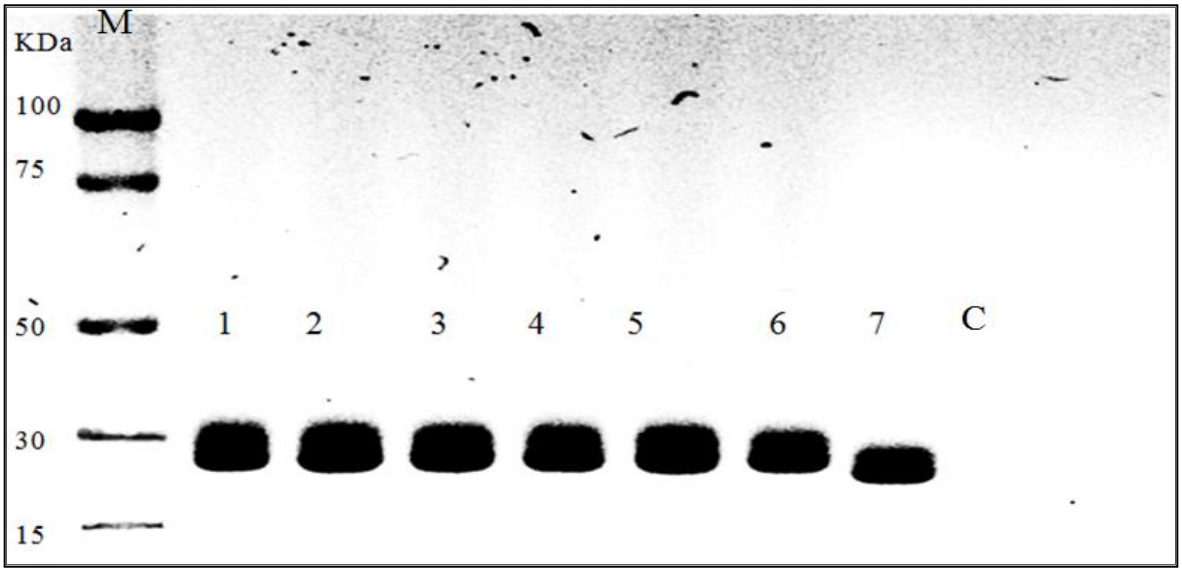


Figure 4.7. Western blot analysis of purified endolysin: (1) LysCD27 (30kDa), (2) LysCD227 (30kDa), (3) LysC2 (30kDa), (4) LysCD119 (30kDa), (5) LysCD6356 (29.9kDa), (6) LysCD38-2 (29.4kDa) and (7) LysCD630 (28.5kDa) by western blot technique. M: Pre-stained protein marker (6x His-protein ladder 100-15 kDa); C: Negative control.

4.3.5 Biochemical properties of lysins

To determine the pH optimum, the lytic activity of recombinant endolysins was assayed at pH values ranging from 4 - 9 in sodium phosphate (Na_3PO_4). All of the endolysins showed their greatest lytic activity at pH5. An example of a typical activity profile is shown in figure 4.8, which depicts the activity of the LysCD6356 endolysin against *C. difficile* NCTC11204 across a range of pH values. Lytic activity decreased at higher pHs, resulting in near inactivation at pH 9, a profile that is common amongst phage lysins (Loeffler *et al.*, 2003; Yoong *et al.*, 2004; Gu *et al*., 2011). While activity was greatest at pH 5, all subsequent assays were performed at pH 7 to determine the activity of lysins under conditions similar to physiological conditions.

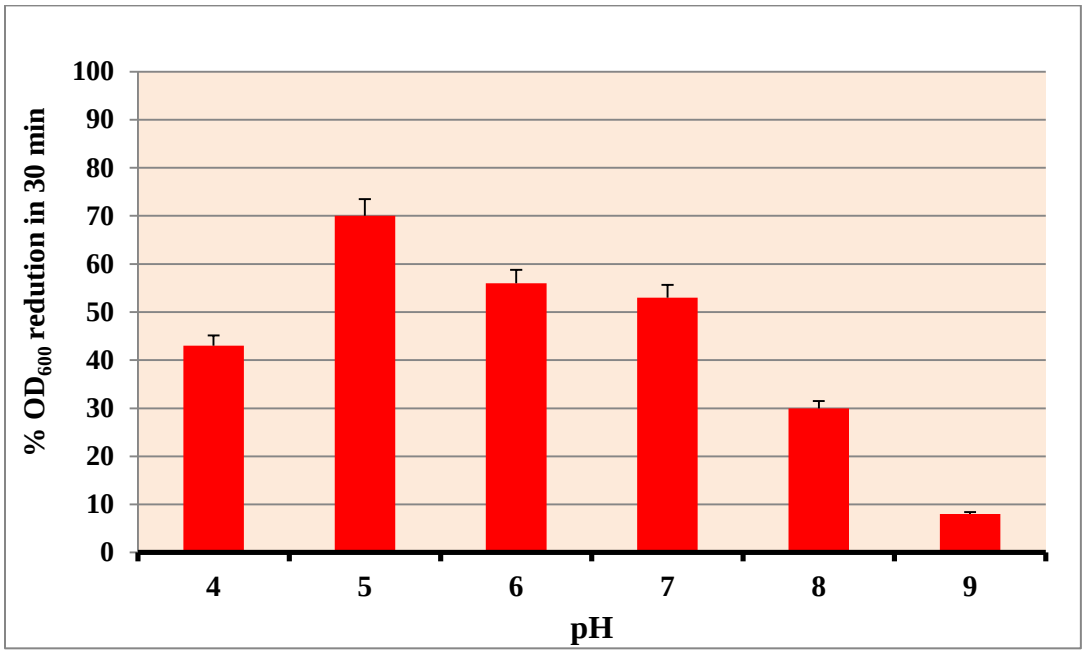


Figure 4.8. pH profile of LysCD6356 activity. LysCD6356 (7µg) activity against the *C. difficile* strain NCTC 11204 was tested at pHs 4, 5, 6, 7, 8, and 9 at 37^0C. Lysin activity was measured by an OD assay and plotted as the percentage decrease in OD_{600} in 30 min. Each column represents the mean of results from triplicate assays, and error bars indicate the standard deviation.

Temperature stability of the lysins was also determined by incubating the recombinant protein at a range of temperatures (37, 45, 50, 60, and 70°C) for 30 min prior to determining their lytic activity. An example of a typical temperature stability profile is shown in figure 4.9, which depicts the activity of the LysCD6356 endolysin against *C. difficile* NCTC11204. As can be seen from the results presented in figure 4.9, the lytic activity of LysCD6356 was lost following incubation at 60°C and 70°C.

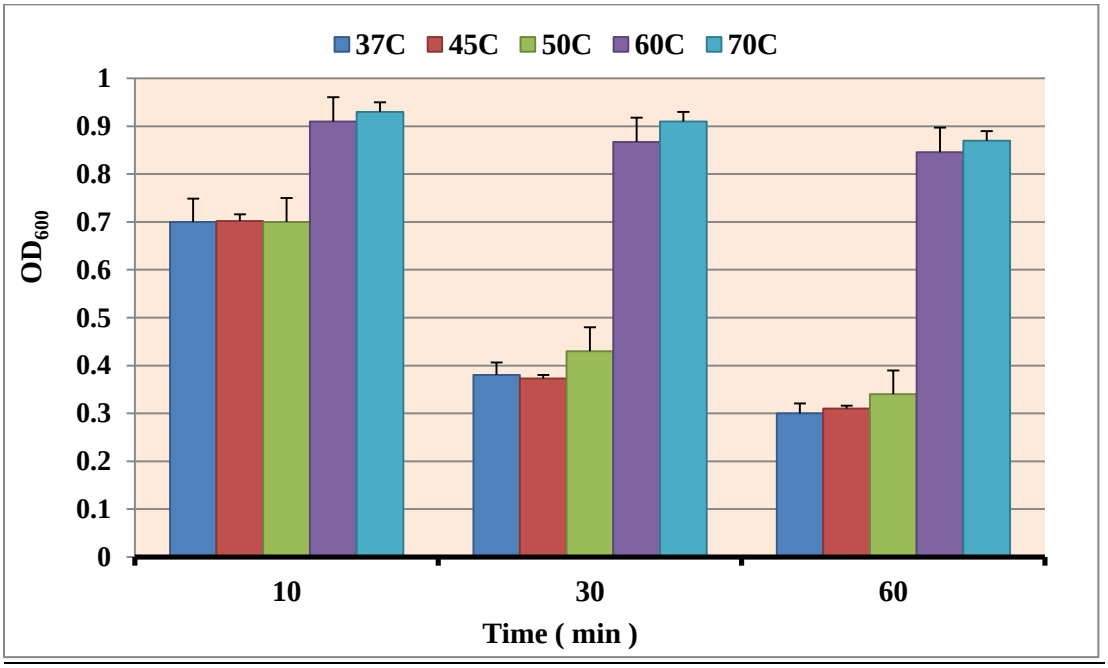


Figure 4.9. Temperature stability profile of LysCD6356, purified LysCD6356 lysin was incubated for 30 min at the indicated temperatures, followed by the addition of 7µg of lysin to a buffered suspension of *C. difficile* NCTC 11204 in 20mM sodium phosphate (pH 7) at 37^0C for 1 h. The figure reports the treated/untreated OD-ratio for each temperature at 10, 30 and 60 min intervals. Each column represents the mean of results from triplicate assays, and error bars indicate the standard deviation

To determine the effect of the suspension media on lysin activity, cells of *C. difficile* NCTC 11204 at pH7 were suspended in either 20mM Sodium Phosphate buffer, BHI broth or PBS and treated with the LysCD6356 lysin. As can be seen in figure 4.10, lytic activity was greatest when the cells were suspended in either PBS or sodium phosphate buffer. In contrast, BHI appeared to inhibit lytic activity.

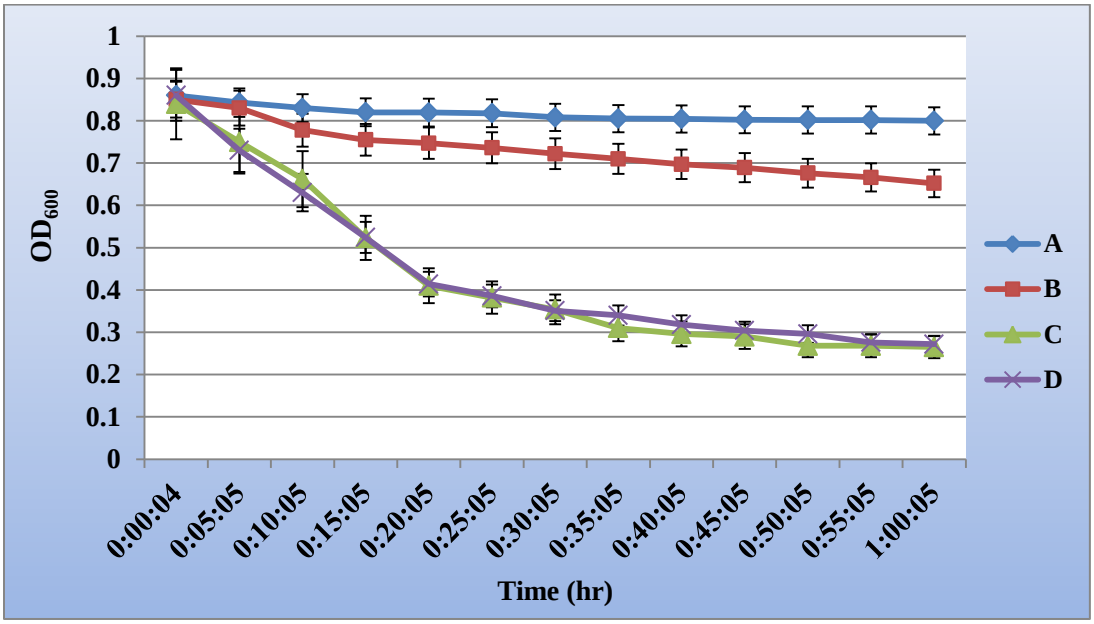


Figure 4.10. Lysis activity of LysCD6356 (7µg) with *C. difficile* NCTC 11204 at pH7 with different buffers, assays were performed at 37^0C for 1h.

A - cells with PBS only, B - cells in BHI broth with lysin, C - cells in 20mM sodium phosphate buffer with lysin, D - cells in PBS with lysin. Values are the means of results from triplicate assays ± standard deviations.

We next determined the effect of concentration on the activity of the LysCD6356 lysin when suspended in PBS at pH 7. As expected, lytic activity was proportional to the concentration of the lysin employed (Figure 4.11).

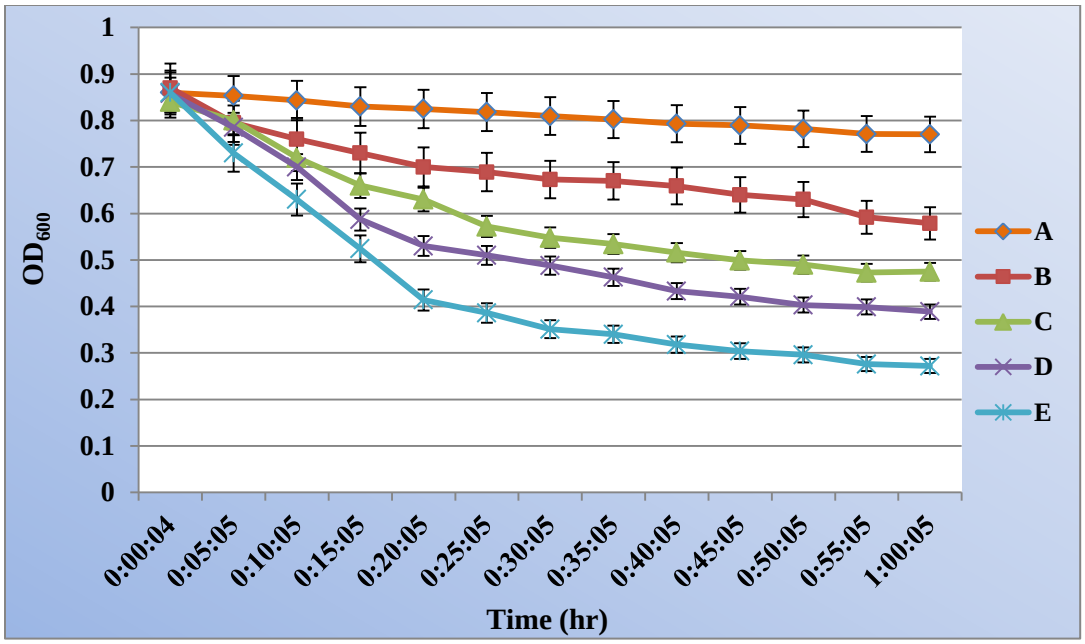


Figure 4.11. Showing the effect of concentration on the lytic activity of the LysCD6356 lysin against *C. difficile* NCTC 11204 , assays were performed at 37°C in pH7 for 1h. A: PBS, B: 0.875µg, C: 1.75 µg , D: 3.5µg , E: 7µg. Values are the means of results from triplicate assays ± standard deviations.

Long term Thermostability

In order to study the effect of storage temperature on lysine stability, LysCD6356 lysin was stored for two weeks at room temperature (ranging from 15-20°C) and at 4°C.

Aliquots were taken at periods of one day, one week and two weeks and were tested for structural integrity and biological activity (Figures 4.12 and 4.13). As can be seen from figure 4.12, the protein remained largely intact over this time period.

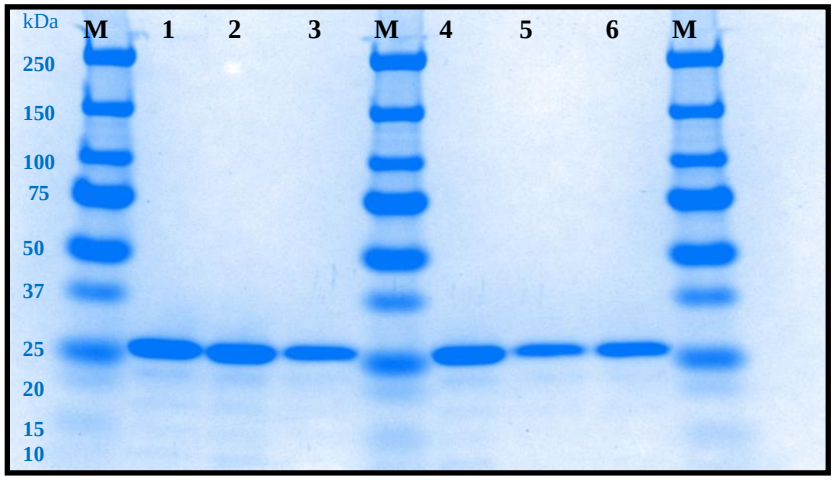


Figure 4.12. Effect of storage temperature on LysCD6356. Lysin stored at room temperature for: (1) one day; (2) one week; and (3) two weeks; lysin stored at 4°C for (4) one day; (5) one week; and (6) two weeks (M) Molecular Marker. SDS-PAGE gel of lysin stained with Coomassie blue showed that lysin bands were not affected over the time of storage.

Storage of the lysin at room temperature and at 4°C appeared to have little significant effect on the activity of lysin (figure 4.13).

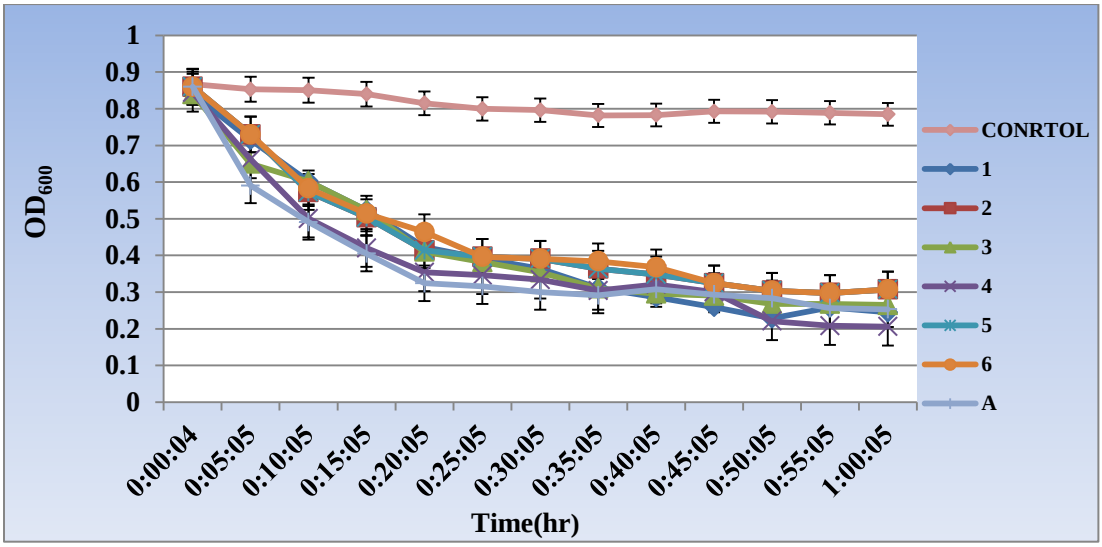


Figure 4.13. Effect of storage temperature on LysCD6356 lysin activity, a7µg of lysin was mixed with cells of *C. difficile* NCTC 11204 in 20mM sodium phosphate, which had been stored under the following conditions: (A) 0 time lysin; lysin stored at room temperatures for (1) one day; (2) one week; (3) two weeks; lysin stored at 4°C for (4) one day; (5) one week; (6)two weeks ,(control): Cells with buffer only assays were performed at 37°C, pH7 for 1h. Values are the means of results from triplicate assays ± standard deviations.

4.3.6 Biological activity

4.3.6.1 Biological activity of the phage endolysins

The ability of the endolysins to lyse a dense suspension of a clinical isolate of a hypervirulent C. *difficile* isolate, DS1813, was determined. The results in figure 4.14 show the relative activity of 7µg of lysin against a mid-log phase cell suspension when tested under the same conditions. The most active lysin was that of LysCD6356 endolysin, which decreased the optical density of the cell suspension by 66% within 30 min ($p < 0.001$). While LysCD6356 endolysin appeared to be the most active lysine, LysCD119 was the least active.

Interestingly, these activity profiles were common for 95% of the *C. difficile* strains tested in our study. We also observed no significant synergistic effects when mixtures of lysins were tested against cell suspensions.

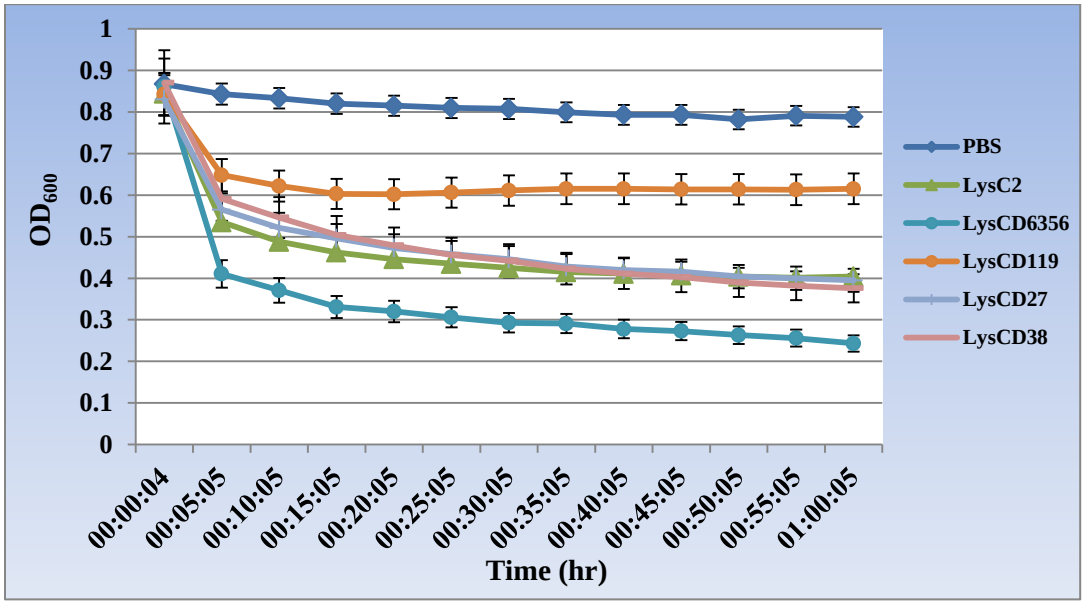


Figure 4.14 . Lytic action of all endolysins in a turbidity assay using *C. difficile* DS1813 cell suspensions prepared from mid-log phase growing cultures. Lysis assay performed using 7µg at 37°C in pH7 for 1 h. Values are the means of results from duplicate assays ± standard deviations.

The enzymatic activity of the most active lysin (LysCD6356) against a viable culture of *C. difficile* DS1813 on a CDMN agar was also determined in an agar plate. Following 24 h incubation, two clear zones of lysis were observed (Figure 4.15).

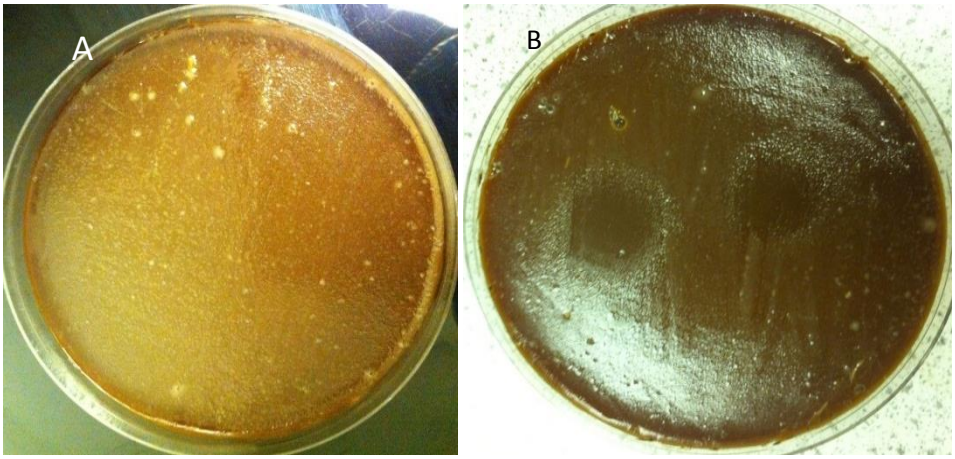


Figure 4.15. Spot assay showing the ability of LysCD6356 lysin to lyse on *C. difficile* DS1813 grown on a CDMN plate. A) Control plate, B) Test plate.

4.3.6.2 Biological activity of the endolysin from *C. difficile* genome (LysCD630)

The study presented here shows the utility of genomic sequencing of *C. difficile* in the identification of candidate lysin.

The LysCD630 protein expressed and characterised exhibits potent lytic activity *in vitro* against strains of *C. difficile*. As seen in figure 4.16, LysCD630 demonstrated lytic activity that was almost equal to those produced by *C. difficile* phage endolysin against *C. difficile* 630. Quantitatively speaking, however, the level of activity of these enzymes among *C. difficile* varied somewhat from strain to strain (see figure 4.19).

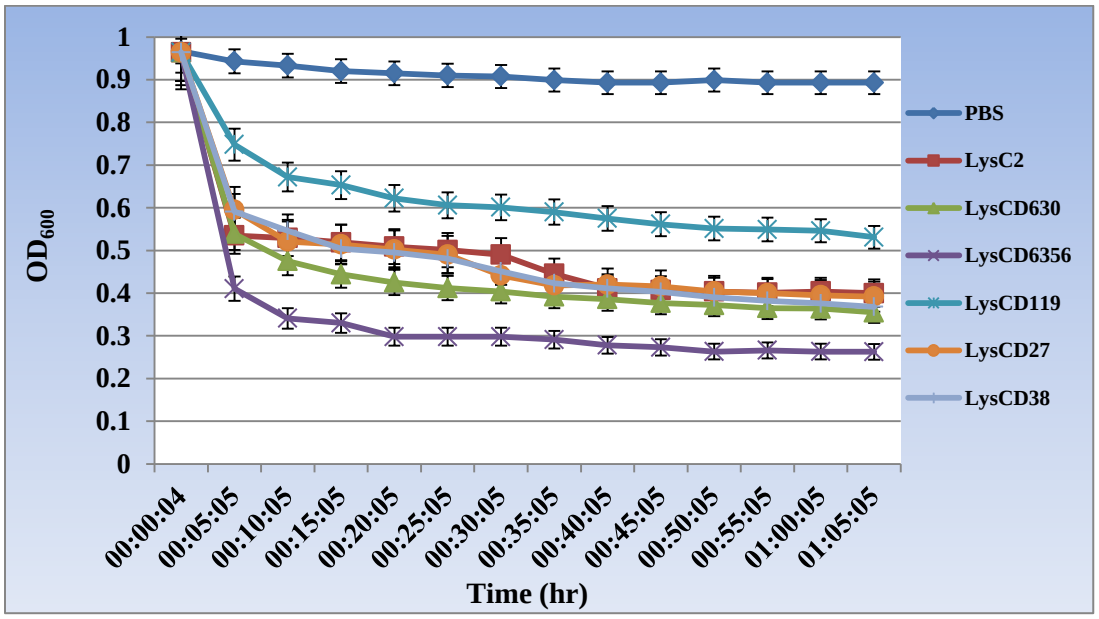


Figure 4.16. Lytic action compression between phage endolysins and putative lysin CD630 in a turbidity assay using *C. difficile* 630 cell suspensions prepared from mid-log phase growing cultures. Lysis assay was performed using 7µg at 37°C in pH7 for 1h. Values are the means of results from duplicate assays ± standard deviations.

4.3.6.3 Biological activity of the mutated endolysin (LysCD227)

Comparison of the lytic activity of the mutated version of the lysin to that of the native version against cells of *C. difficile* strain NCTC 11204, *L. monocytogenes* NCTC 5412 and *B. subtilis* ATCC 6633, found that the activity of LysCD227 was significantly enhanced compared to that of LysCD27 (Figure 4.17).

The enhanced activity of LysCD227 for these strains compared to LysCD27 suggests that the mutation in the C binding domain had altered the specificity of the lysin. Mayer *et al.* (2011) had previously observed that while *L. monocytogenes* NCTC 5412 and *B. subtilis* ATCC 6633 were unsusceptible to the full length LysCD27 lysin, the removal of the C terminal domain resulted in an enzyme which was able to lyse these bacteria (Figure 4.17).

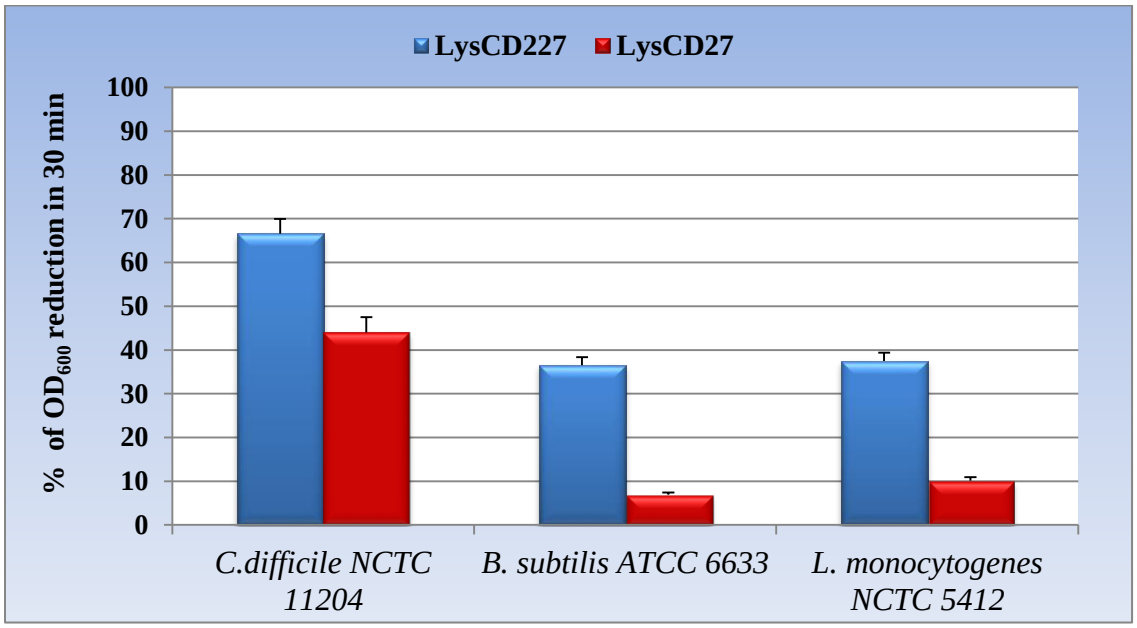


Figure 4.17. Compression of the lytic activities of the LysCD27 and LysCD227 on *C. difficile* strain NCTC 11204, *L. monocytogenes* NCTC 5412 and *B. subtilis* ATCC 6633. The activity was measured as a percent of OD reduction of bacterial cell suspensions after 30 min of incubation using 7µg of endolysin in pH7 at 37°C. Each column represents the mean of results from triplicate assays, and error bars indicate the standard deviation

Comparison of the lytic activity of LysCD227 and LysCD27 against cells of various strains of *C. difficile* supported the previous result that cell wall binding domain modification significantly (Student's *t* test; $P < 0.05$) enhanced the catalytic activity of endolysin (Figure 4.18).

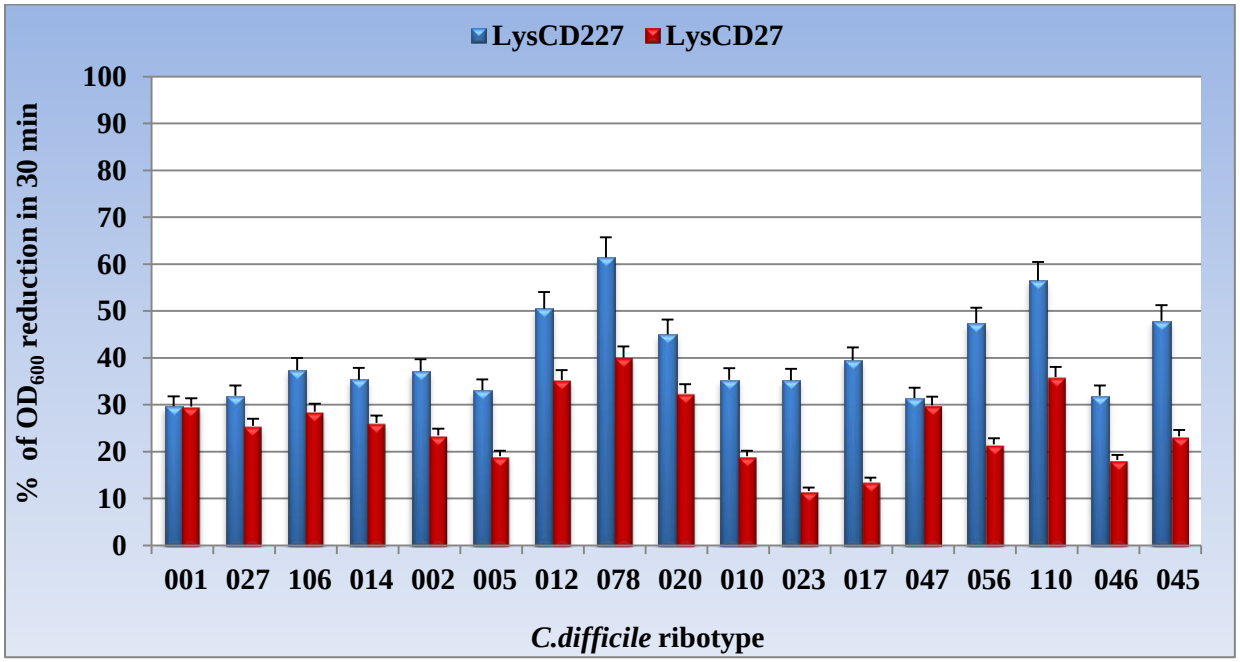


Figure 4.18. Lytic activities of LysCD227 and LysCD27 on various ribotypes of *C. difficile*. The activity was measured as percentage OD reduction of bacterial cell suspensions after 30 min of incaution using 7µg of endolysin in pH7 at 37°C. Each column represents the mean of results from triplicate assays, and error bars indicate the standard deviation

4.3.6.4 Spectrum of lytic activity

We determined the ability of all endolysins cloned in this study to target a diverse collection of clinical isolates representing the major ribotypes that are prevalent in the UK. As can be seen in figure 4.19, LysCD6356 demonstrated the greatest lytic activity (one-way ANOVA; $P < 0.05$) against all strains of *C. difficile* tested (ribotypes 001, 027, 106, 014, 002, 005, 078, 020, 010 and 023) with isolates of the 078 ribotype being the most susceptible. The second most active endolysin was LysCD38-2, which was the most active (one-way ANOVA; $P < 0.05$) against strains of ribotype 012, 110, 045 and 046, based on % of lysis activity after 30 minutes of incubation. Interestingly, both endolysins were derived from phages belonging to the *Siphoviridae* family.

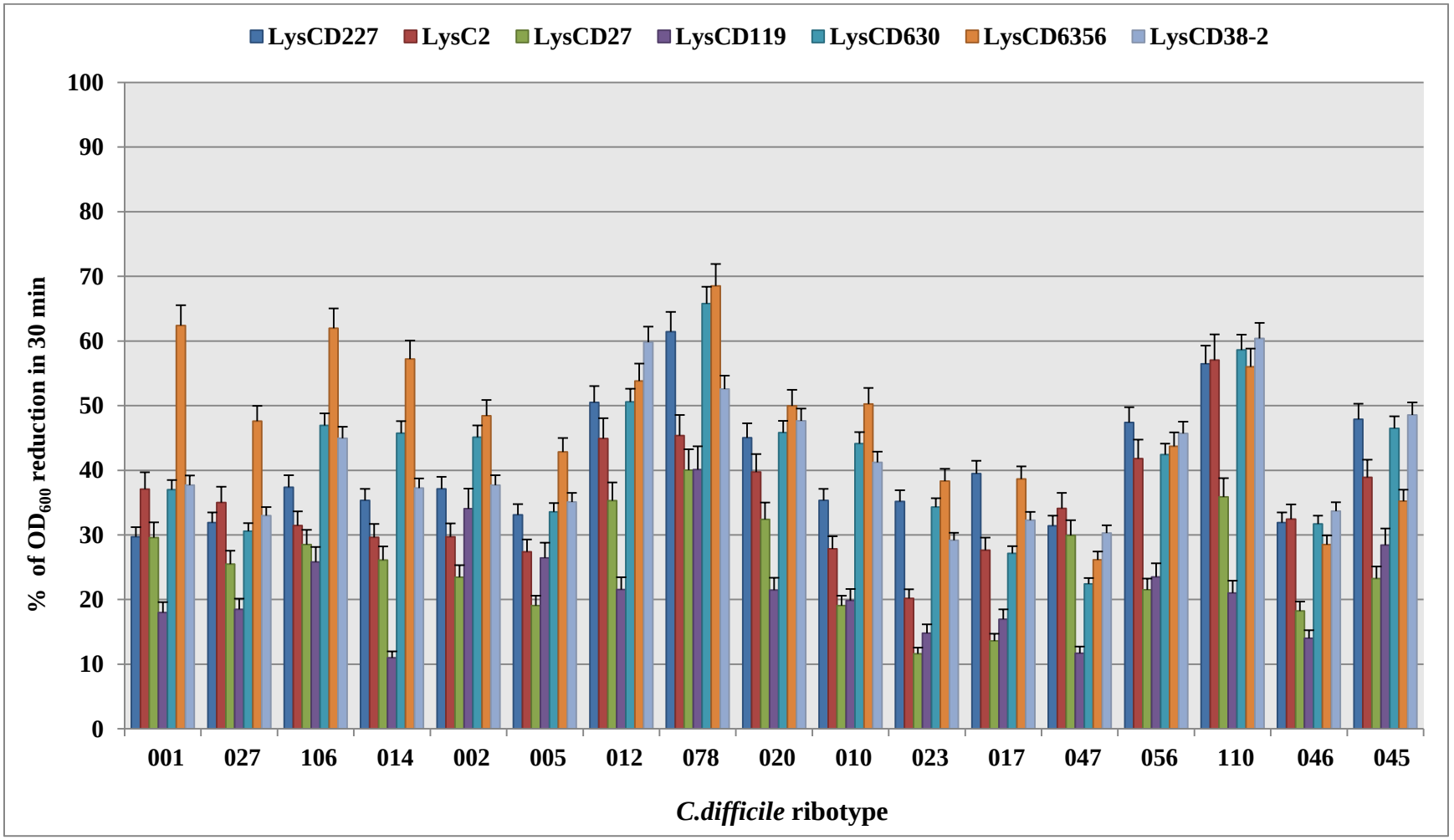


Figure 4.19. Comparison of lytic activities of lysins against a range of ribotypes of *C. difficile* (*n=5* for each ribotype) , lytic activity was measured as percent of OD reduction after 30 min of incubation using 7µg of lysin in pH7 at 37°C. These results allowed the direct comparison of lysin activities among different ribotypes of *C. difficile*. Each column represents the mean of results from triplicate assays, and error bars indicate the standard deviation.

4.3.6.5 Activity of the endolysins against other species of Gram-positive bacteria

The ability of the lysins to target other Gram-positive spore forming bacteria was determined using various strains of *C.perfringens* (Figure 4.20).

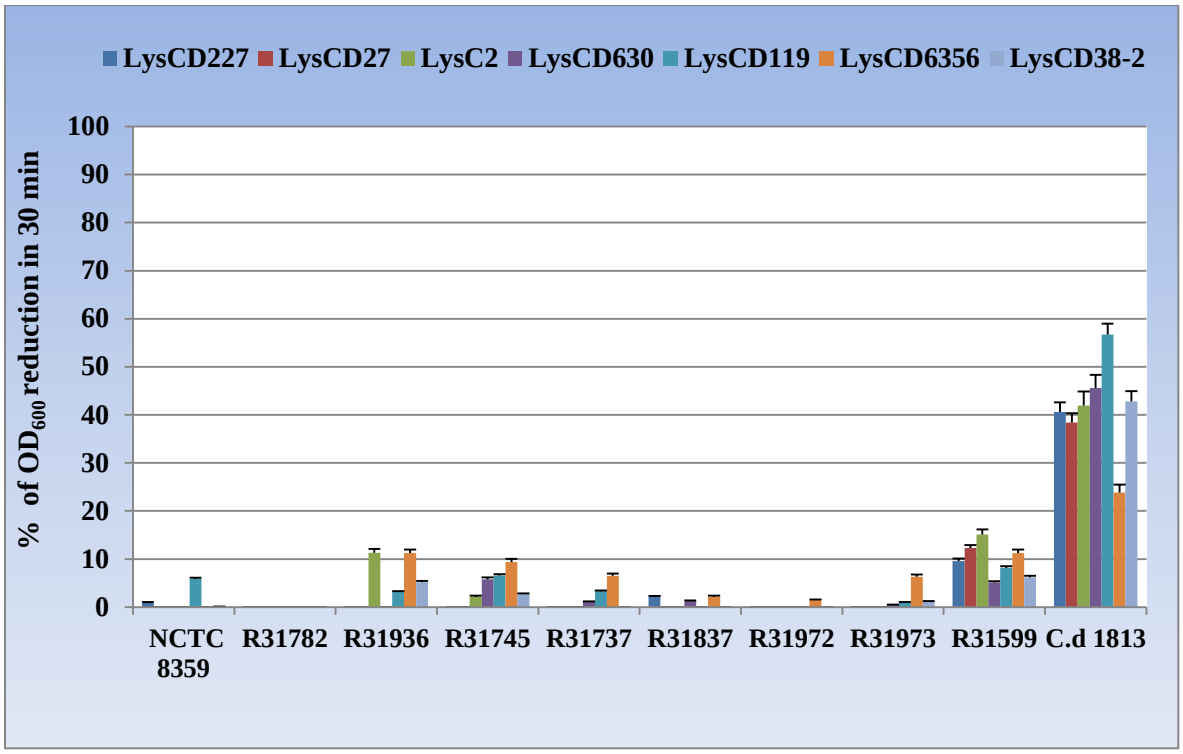


Figure 4.20. Lytic activities of our lysins on various strains of *C. perfringens*. Graph shows percent of reduction in the OD of a bacterial cell suspension after treatment with 7 µg/ml our lysin for 30 min. *C. difficile* DS1813 (ribotype 027), included as a positive sensitive strains. Each column represents the mean of results from triplicate assays, and error bars indicate the standard deviation.

The lytic activity of our lysins was also evaluated in a few clinical isolates of Gram-positive bacteria (described in Chapter2, table 2.2) by performing the lysis assay as described above. No obvious lytic activity could be determined in any of the clostridium isolates, even with the highest protein concentrations used (10 and 15 µg/ml) (Figures 4.20 and 4.21). However, *C. sordellii,* which is phylogenetically closely related to *C. difficile* (Collins *et al.*,1994), seems to be sensitive to some of our lysin; moreover, some of our lysins were surprisingly capable of lysing a number of less closely related species, such as *B. subtilis* ATCC 6633, which was significantly sensitive to lysis with LysCD227, LysCD630, LysCD119 and LysCD6356. *L. monocytogenes* NCTC 5412 was only significantly sensitive to LysCD227 and LysCD119.

Interestingly, our lysins were active against species characterised as having peptidoglycan type A1γ as described in Mayer *et al*.(2011), but not active against strains that do not have that type of peptidoglycan, such as *Micrococcus luteus* (NCTC 2665) and *Staphylococcus aureus* (NCTC 10788) (Schleifer and Kandler, 1972*;* Mayer *et al.,* 2011).

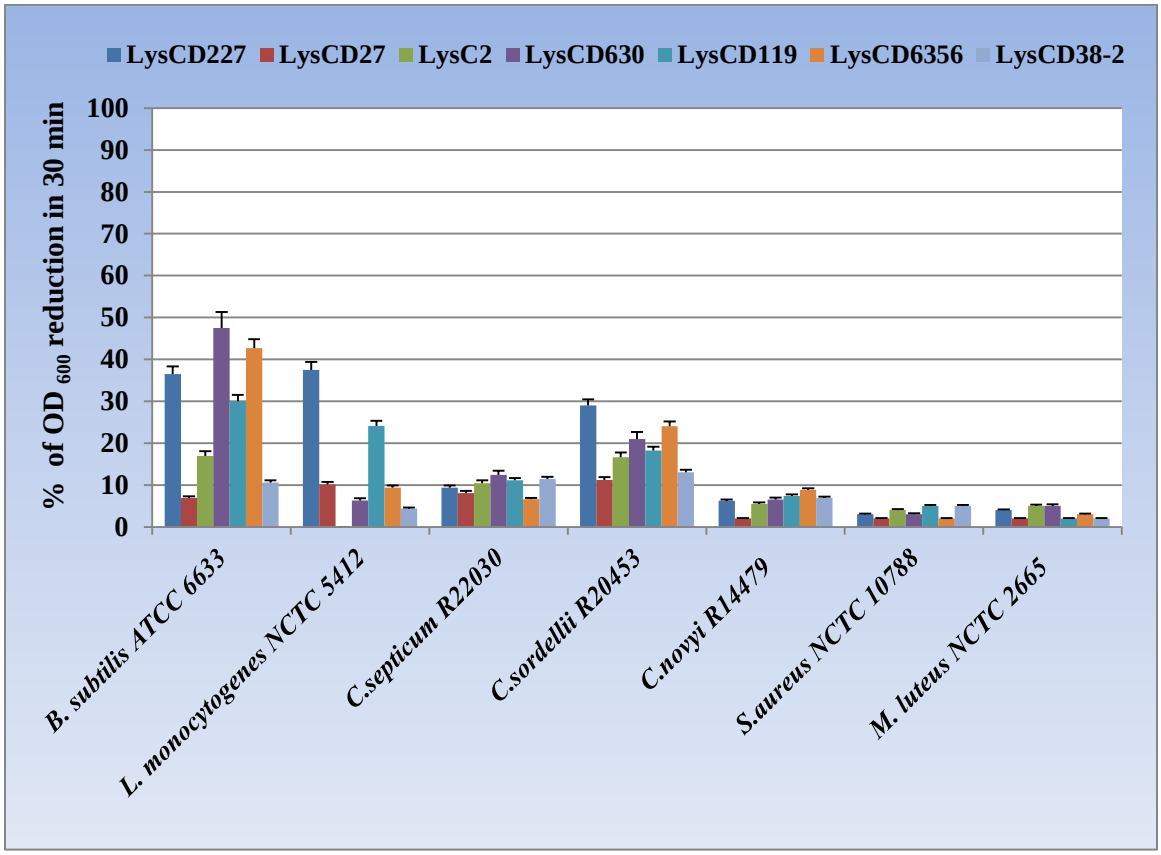


Figure 4.21. Lytic activities of our lysins on various strains of *Clostridia* and other Gram-positive strains. Graph shows percent of reduction in the OD of a bacterial cell suspension after treatment with 7 µg/ml of lysin for 30 min. Each column represents the mean of results from triplicate assays, and error bars indicate the standard deviation

4.4 Discussion

4.4.1 Endolysins isolated from *C. difficile* phages

All *C. difficile* specific phages that have been isolated to date and characterised in order to develop therapeutic agents have proven to be lysogenic phages, and no lytic phages have been isolated (Sell *et al*.,1983 ; Mahony *et al*.,1985 ; Dei, 1989; Nagy *et al.,* 1991; Goh *et al.,* 2005; Govind *et al.,* 2006; Fortier and Moineau, 2007; Mayer *et al.,* 2008; Horgan *et al*., 2010; Sekulovic, *et al*., 2011; Meessen-Pinard *et al*., 2012; Nale *et al*., 2012; Shan *et al.,* 2012) and in this study. However, these types of phages are of limited value as therapeutic agents (Gill and Hyman, 2010; Meader *et al*., 2010).

Although these phages have limited therapeutic potential in their current form, they each carry a gene encoding an endolysin which has the ability to degrade the bacterial cell wall when applied externally. Thus, these proteins could be developed and used as effective therapeutic agents against *C. difficile*. In theory, endolysins would be well suited as treatment of *C. difficile,* since they target only *C. difficile* cells without harming the normal body microbiota. In this work we have characterised five endolysins from *C. difficile:* φCD27, φC2, φCD119 (*Myoviridae* family), φCD6356 and φCD38-2 (*Siphoviridae* family).

Bioinformatic analysis of these endolysins revealed that they all contained N-acetylmuramoyl-L-alanine amidase (amidase3) N-terminal catalytic domains. This class of endolysin cut the critical amide bond between *N*-acetylmuramic acid and L-alanine and are predicted to be the most active PG degrading enzymes (Oliveira *et al*., 2012).

This type of endolysin represents one of the most widespread classes of endolysin (Fischetti, 2008). In contrast, the C-terminal domains of these endolysins lacked any homology, which was somewhat surprising given that these phages are specific for *C. difficile*.

All endolysin sequences were codon optimised for expression in *E. coli*, to maximise the production of recombinant protein. Comparison of the level of expression of codon optimised LysCD27 to that of the non-codon optimised version found a two-fold increase in expression (Mayer *et al*., 2008). During protein induction at 37°C, the recombinant proteins formed insoluble inclusion bodies located in inclusion bodies; these results correlate with those generated by a protein solubility prediction program that all of our proteins have a high insolubility percent (Table 4.3). However, the culture temperature was reduced to 27°C, which is a temperature that has been reported to reduce the formation of inclusion bodies (Zimmer *et al.,* 2002; Dhalluin *et al.,* 2005).

With regards to the level of protein expression of the other endolysins, they were expressed at a roughly similar level, ranging from 1-4 mg/100ml.

Characterisation of the biological activity of the recombinant endolysins across a range of pH values was determined. The endolysins maintained their activity over a wide range of pHs(4 to 9) including physiological pH, suggesting that they would remain biologically active in the gastrointestinal tract (GI) environment. A similar observation was made for other previously characterised *C.difficile* phage endolysin (φCD27 endolysin) (Mayer *et al*., 2008). The effect of temperature on lytic activity was also determined and was found to be stable over a range of temperatures (37-50°C). However, a sharp decrease in

activity was seen at temperatures of 60°C and above. With the exception of BHI, the buffers tested had no significant effect on the lysins activity. Thus, based on these results, we developed a standardised set of assay conditions with which to determine the level of activity of each endolysin.

While the lytic profile of each endolysin for *C. difficile* PG was similar to that described in the literature, this was not the case for LysCD27 (Loeffler *et al*., 2001; Schuch *et al.*, 2002; Son *et al*., 2012). Mayer and colleagues (2008) reported that the lytic activity of their version of LysCD27 required a 10 minute incubation step before lytic activity was initiated (Mayer *et al*., 2008). In contrast, our version of LysCD27 was able to lyse the PG almost immediately upon addition. Mayer and colleagues (2008) suggested that this delay may have been the result of incorrect folding of the recombinant lysin expressed from their *E. coli*-based system (pET15b, *E. coli* BL21). Our results suggest that our recombinant protein, which is expressed from different host *E. coli* strains to the Mayer and colleagues (2008) group, is produced in the correct form and hence, is more active. Given that one of the differences between the two systems is the codon usage of the gene, it is tempting to speculate that this is responsible for the enhanced activity (Zimmer *et al*., 2002).

While the level of activity of individual lysins varied from strain to strain, the LysCD6356 endolysin was the most active against all strains of *C. difficile* tested (ribotypes: 001, 027, 106, 014, 002, 005, 078, 020, 010 and 023). The second most active endolysin was LysCD38-2, which was the most active against strains of ribotype 012, 110, 045 and 046, based on % of lysis activity after 30 minute of incubation.

Interestingly, LysCD6356 shares a high percent amino acid identity (88%) with LysCD38-2, and they belong to same phage family and same phage group

These endolysins also showed activity against other Gram-positive strains, such as *C. sordellii,* which is closely related to *C. difficile,* and less closely related species *B. subtilis* and *L. monocytogenes*. However, they were inactive against a number of other Gram-positive species.

The host range of each endolysin was significantly broader than that of the phages themselves, as previously described for φCD27 (Mayer *et al*., 2008). Moreover, all *C. difficile* strains tested were at least sensitive to one of our endolysins.

This phenomenon support the therapeutic potential of endolysins compared to bacteriophages.

4.4.2 Endolysin isolated from the *C. difficile* genome (LysCD630)

In this work we have shown the utility of employing a bioinformatics-based approach to identify protein sequences with a putative endolysin function. This endolysin is located outside of the prophage regions. We selected an endolysin from CD630 for further study on the basis that it contained an amidase 3 N-terminal domain, has a 62% amino acid identity with known phage endolysin (φCD6356) and the homologues of this gene were found to be conserved among all *C. difficile* strains sequenced to date (Sebaihia *et al*., 2006; Stabler *et al.,* 2009; He *et al.,* 2010; Monot *et al.,* 2011). Furthermore, using PCR and primers specific to this gene, we were able to detect its presence in all 85 of the isolates of *C. difficile* examined to date (Chapter 5).

The apparent universal distribution of the lysin supports the hypothesis that it plays an important function in the physiology of *C. difficile*.

The fact that the recombinant version of these gene was indeed able to lyse the *C. difficile* cell wall in a similar manner to phage-derived endolysins, and that it appears to be present in all of the isolates of *C. difficile* we have examined to date, suggests that it could be a strong candidate for development as a future therapeutic agent.

4.4.3 Mutated endolysin (LysCD227)

Mayer *et al*. (2010) demonstrated that the isolated enzymatic catalytic domain of the endolysin of φCD27 had greater lytic activity against *C. difficile* than the full length endolysin. They also found that it had a wider spectrum of activity when the C terminal domain was removed and was able to lyse isolates of *L. monocytogenes* and *L. innocua,* which were insensitive to the full sequence. A similar pattern of lytic activity has been reported for other endolysins where the catalytic domain alone was a more effective for cell lysis than the whole endolysin (Loessner *et al*., 1999; Low *et al*., 2005; Cheng and Fischetti, 2007; Mayer *et al.,* 2010).

However, other scientific research has shown that some phage endolysin activity was decreased by removing the CBD; the only explanation for this is that these endolysins require the whole endolysin sequence in order to produce lytic activity and the binding function is present in an area between ECD and CBD with these phages (Sanz *et al*.,

1992; Loessner *et al.* ,2002; Hermoso *et al*., 2003; Sass and Bierbaum, 2007; Kikkawa *et al.*, 2007; Mayer *et al*., 2010).

The results of our study showed that a small modification in the sequence of the CBD region of LysCD227 increased the activity of the lysin when compared to LysCD27. This suggests that, in *C. difficile,* the modified CBD was not able to bind the epitope that is predominantly found on the *C. difficile* cell wall and therefore was not able to tether the lytic activity of the catalytic domain from targeting other cells.

It has been proposed that in addition to mediating the specificity of the endolysin, the CBD region prevents the lysin from damaging new bacterial hosts by ensuring that the enzyme binds irreversibly to specific sites in the cell wall peptidoglycan, which the newly released phage may seek to infect (Mayer *et al*., 2010; Oliveira *et al*., 2012). Thus, there is an evolutionary pressure to retain the CBD which is secondary to its enzymatic activity. In contrast, the peptidoglycan of Gram-negative bacteria is protected from the attention of external endolysins by an outer membrane, which would explain why most Gram-negative phage-derived endolysins are described to date lacking a CBD. Also interesting was the fact that these modification also extended in the lytic spectrum of LysCD227, which was active against other Gram-positive strains, This suggests that the LysCD27 CBD may target a bacterial cell wall epitope that is common in some Gram-positive bacterial species.

4.5 Conclusions

The recombinant endolysins examined in this study possessed broad spectrum activity against a range of different *C. difficile* ribotypes with little or no activity against other species included in this study, thus highlighting their potential as novel, specific microbial agents. Indeed, our characterization of endolysin LysCD27 and its variant, LyCD227, provides a starting point from which to investigate the ability of genetic engineering to improve the activity of these endolysin. This work demonstrated the utility of using *C.difficile* sequenced genomes as rich pools for the identification of potential endolysins

In the context of developing new therapeutics for the treatment of *C. difficile*, endolysin, in contrast to antibiotics, offers a novel approach to the destruction of the specific pathogens without damaging the normal microbiota of the GI tract.

Chapter Five

The distribution of endolysins amongst clinical isolates of *Clostridium difficile*

5.1 Introduction

5.1.1 Use of the phage gene as a marker of prophage carriage

As a consequence of the increased interest in bacteriophage research and bacteriophage therapy a number of *C.difficile*-specific temperate bacteriophages have been isolated. Of these bacteriophages only seven bacteriophages of *C. difficile*, namely φCD119, φCD2 ,φCD27 , φMMP03 , φMMP04, φCD6356 and φCD38-2, have been fully characterized at the molecular level (section 1.12). In addition to these examples, prophages have also been identified in most *C.difficile* strains sequenced to date (Sebaihia *et al.,* 2006 ; Goh *et al.,* 2007; He *et al.*,2010; Nale *et al*., 2012 ; Shan *et al.*, 2012).

Prior to the advent of the molecular age bacteriophages were isolated by inducing strains chosen at random from clinical or environmental samples with mitogens such as mitomycin C. Although these methods have yielded bacteriophages in the past they are extremely time-consuming and dependent on isolating and finding an indicator strain for propagation of the initial bacteriophage. Moreover, they cannot predict how novel the identified bacteriophage will be relative to previously characterized bacteriophages.

Several bacteriophage genes have been used to study phage diversity, as these genes encode for the major capsid protein and portal protein in the myoviruses *gp20* and *gp23* respectively (Fuller *et al*., 1998; Marston and Sallee, 2003; Short and Suttle, 2005; Shan *et al*., 2012). Phage DNA polymerase and terminase genes have also been used to assess phage diversity (Chen and Suttle, 1996; Chen *et al*., 1996 ; Marston and Sallee, 2003

;Short and Suttle, 2005 ; Shan *et al*., 2012). Furthermore, the holin gene was used to study prophage carriage in *Streptococcus pneumonia* (Romero *et al*., 2009) and the molecular phylogeny of phi29-like phages and their evolutionary relatedness (Pecenkova and Paces, 1999).

In *C.difficile* studies, only major capsid protein genes and holin genes have been used to study prophage diversity (Nale *et al.,* 2012; Shan *et al.,* 2012). None of these genes, however, are conserved enough to be targeted in all known *C.difficile* bacteriophages.

The endolysin gene may be a suitable bacteriophage gene for studying the genetic diversity of *C.difficile* bacteriophages. The endolysin gene encodes protein that is essential for breaking down the cell wall PG bond and releasing the new phage progeny. This gene is present in all the seven sequenced *C. difficile* phages and has been identified in all of the *C. difficile* genomes sequenced to date (Chapter 4) . We selected these endolysins as PCR targets due to the fact that they share a high degree of sequence homology and belong to a single class of endolysin.

For this reason we employed a polymerase chain reaction (PCR) based approach to determine the distribution of bacteriophage endolysins across a diverse collection of *C.difficile* clinical isolates. It was our hypothesis that the presence of an endolysin may be indicative of the presence of a bacteriophage. To do this we designed PCR primers which recognized the endolysins present in the *C.difficile* bacteriophage sequences which have been deposited in GeneBank.

In addition to searching for known endolysins amongst our collection of strains we also sought to identify new endolysins by treating PCR negative isolates with mitomycin C and screening for the production of the bacteriophages, the premise being that these bacteriophages may contain novel endolysins which could be of value as therapeutic agents.

5.1.2 Aim of this study

The aims of this chapter are:

- To employ a PCR-based method to screen multiple clinical isolates of *C.difficile* for the presence of homolog to known *C.difficile* bacteriophage endolysin genes.
- To characterize the distribution of bacteriophage endolysins.
- To characterize the distribution of the *C.difficile* endolysin LysCD630 described in Chapter 4.
- To identify novel bacteriophage endolysins by inducing the expression of bacteriophage from PCR negative *C.difficile* isolates.

5.2 Material and methods

In this study we examined a much bigger collection of strains than were used in the original bacteriophage induction work; strains that were used in this study are described in table 2.1 and 2.2.

5.2.1 PCR primer design and synthesis

The National Center for Biotechnology Information's (NCBI) nucleotide database at the time of searching contained the complete genome sequences of the following *C. difficile* bacteriophages; φCD119, φCD27, φC2, φCD6356 and φCD38-2. The sequences encoding the endolysin of each of these bacteriophages were used as templates to design specific PCR primers.

A further putative endolysin gene which was identified by bioinformatics tools in the genome of *C. difficile* 630 Chapter 4 (NC_009089 region: 2201141-2201929) was used as template to design PCR primers.

PCR primers with specificity for each endolysin were designed with NCBI primer-blast (www.ncbi.nlm.nih.gov) and the MWG Operon online PCR primer design tool (www.eurofinsdna.com) under the following criteria: (1) primer length was 18–25 bp in length, (2) annealing temperature was 55–65°C, (3) guanine-cytosine (GC) content of the primer was 40%–60%, and (4) amplicon size 200-700 bp in length (Table 5.1).

Endolysin gene	Primer sequence (5'->3')	Annealing temperature, ^{0}C	Product size, bp
LysCD27	F1: TGGAGCATGTACTTCTGCTGATGG R1: CGTCTGACCTGGCACGTAATCTTT	**58**	**640**
	F2:TGGAGCATGTACTTCTGCTGAT R2:CATGCGCCACCTCCTACAA	**60**	**672**
LysC2	F1: GCATGCACTTCTGCTGATGGAGTA R1: TCCGTCTGACCTGGTACGTAATCT	**56**	**638**
	F2: TGCACTTCTGCTGATGGAGTA R2: CCCAACTAAGCACATTCGCA	**59**	**574**
LysCD119	F1: GCAACACAAGCAAATAGACAAG R1: TTCACTGTCCACAAAGCATAC	**56**	**246**
	F2:TGTGGACATACAAAAACAGGAGC R2:TCGCATCCTCGTCCTTTTCC	**56**	**251**
LysCD6356	F1: CGGATGCAAGAGGGCTTTACTGGC R1: CGCATGCTCCTCCCCCTACAACA	**60**	**348**
	F2: AGGTTGTACTAACAGCAGGGC R2: AAGCCCTCTTGCATCCGATT	**58**	**374**
LysCD38-2	F1:ACAGCAGGGCATACGCTAACAGG R1: TGCGTCCGATTTTGCGCCTCT	**59**	**354**
	F2: TGAGAGTAGCACTAACAGCAG R2: CCACCCCCAACAACATAAAG	**59**	**700**
LysCD630	F1: AGCAGCCACTTTATCAGCATCACCT R1: GGAAAAGGTACAGGAGCTGTTGGG	**58**	**564**
	F2: AGCAGCCACTTTATCAGCATCACC R2: GGAAAAGGTACAGGAGCTGTTGGGT	**58**	**564**

Table 5.1. Sequences (listed 5'-3') of primers used for amplification of the *C.difficile* bacteriophages endolysin gene, derived from endolysin genes of φCD27, φC2, φCD119, φCD6356, and φCD38-2 and a putative phage endolysin gene within the genome of *C.difficile* 630 (NC_009089 region: 2201141-2201929).

To serve as positive controls previously described primers which target non-repeating sequences within the gene of toxin A were employed (Kato *et al.*, 1991). Primers that target the 16S rRNA gene of *C.difficile* were also employed as positive control primers (Gumerlock *et al.*, 1991). The sequences of these primers are shown in table 5.2.

Target	Primer sequence (5'->3')	Annealing temperature, ^{0}C	Product size, bp
Toxin A gene*	F1: GGACATGGTAAAGATGAATTC R1: CCCAATAGAAGATTCAATATTAAGCTT	**60**	**546**
	F2: GGAAGAAAAGAACTTCTGGCTCACTCAGGT R2: CCCAATAGAAGATTCAATATTAAGCTT	**58**	**252**
16S rRNA gene**	F1: CTCTTGAAACTGGGAGACTTGA R1: CCGTCAATTCMTTTRAGTTT	**55**	**270**

Table 5.2. Sets of primers were used with PCR reaction as positive control primers: (*) primers derived from (Kato *et al.*, 1991) targeting the Toxin A gene of *C.difficile* ; (**) primers derived from (Gumerlock *et al*., 1991) targeting the 16S rRNA gene of *C.difficile*.

Primers were synthesized and purified by high-performance liquid chromatography (HPLC) supplied by Eurofins MWG Operon (Ebersberg, Germany) and were supplied lyophilized, with 100 pmol/µl concentration. Prior to use the primers were resuspended in sterile distilled water to the required concentration.

5.2.2 DNA extraction using Chelex

Template DNA was extracted from selected bacterial cultures by the use of a commercially available Chelex resin-based DNA extraction kit (InstaGene Matrix; Bio-Rad). Colonies from an overnight culture of the bacteria on a blood agar plate incubated at 37°C in an anaerobic chamber were harvested and resuspended in 200µl of 5% Chelex 100 and heated at 56°C in a dry block heater (Techne DB-3D, Bibby Scientific Ltd, Staffordshire, UK), for 30 minutes. The suspension was then vortex mixed and then heated again but this time at 100°C for eight minutes. Bacterial cell debris was removed from the suspension by centrifugation at 10,000 g for three minutes

at 4^0C temperature (Eppendorf 5417R centrifuge, fixed angle rotor F45-30-11, UK) and the resulting supernatant which contained bacterial DNA was transferred to new tubes and stored at -20^0C until needed. The extraction method above was adapted from a method for purifying *C. difficile* DNA described by Arroyo *et al*. (2005).

5.2.3 Spectrophotometric quantification of DNA concentration

The concentration of DNA in a sample was estimated by measuring the absorbance of the solution at 260nm using a cuvette (Cuvette UVette 80 disposable single sealed Eppendorf ,Fisher Scientific, Leicestershire, UK) and an Eppendorf Biophotometer (Eppendorf UK Ltd, Cambridge, UK). An OD_{260} value of one is equivalent to a DNA concentration of 50µg /ml for double-stranded DNA. The purity of the sample was assessed by calculating the 260/280 nm absorbance ratio. This approximated to > 1.5 in protein-free DNA samples.

5.2.4 Optimization of the PCR

Parameters were optimized using known positive strains before we screened our collection in order to determine the optimum PCR reaction conditions. These parameters included $MgCl_2$ concentration (1.0 to 3.5 mM), primer concentration (0.2 to 2.0 µM) and annealing temperature (55 to 66°C). Each component was optimized while others were kept constant. Then concurrent optimized parameters were used in subsequent experiments; however, other parameters were used according to the manufacturer's instructions (Taq PCR Handbook, 2012; Qiagen, Crawley,UK).

5.2.5 PCR reaction conditions

PCR amplifications were undertaken using a standard PCR reaction mix, the composition of which is shown in table 5.3. The following components, CoralLoad PCR Buffer dNTP mix (10 mM each), Q-Solution, 5X Taq DNA Polymerase, and optimized $MgCl_2$ (25 mM), were obtained from the Taq PCR Core Kit (Qiagen, Crawley, UK). To confirm the success of the PCR reaction, positive control primers which recognized a conserved target present in *C.difficile* 630 were included in each PCR amplification run (Table 5.2.). DNA extracted from *C.sordellii*, a strain not known to contain any of the target endolysins, was included as a negative control.

	Components	Volume for 1 reaction (µl)	Final concentration
Master Mix	**CoralLoad PCR Buffer**	3.75 µL	1x
	dNTP mix (10 mM each)	0.6 µL	200 µM of each dNTP
	Q-Solution	6 µL	1x
	Taq DNA Polymerase	0.15 µL	2.5 units/reaction
	MgCl2 (25 mM)	2.4µL	2.0 mM
	Distilled water	14.1 µL	
	Primer mix (F&R)	1.5 µL	0.5µM each
	Template DNA	1.5 µL	
	Total volume	30 µL	

Table 5.3. PCR reaction composition

PCR amplifications were carried out in a Touchgene Gradient Thermal Cycler (Techne (Cambridge) Ltd, UK). The following programme was used; one cycle of 3 minute at 94 °C for initial denaturation followed by 35 cycles (45 second at 94 °C, 45 second at 55-63 °C and 1 minute at 72 °C) and a final extension of 10 minute at 72 °C.

PCR products were subsequently analysed by agarose gel electrophoresis.

5.2.6 DNA gel electrophoresis

Agarose gels were prepared by melting 1% (w/v) agarose **(**Sigma-Aldrich, Gillingham, UK) in 100ml of 1x Tris-borate EDTA (TBE) (1xTBE: 89mM Tris base; 89 mM boric acid; 2 mM EDTA; pH 8.0)(Fisher Scientific, UK), in a 100ml glass Duran bottle using a microwave oven .

Following heating the agarose was cooled to 45°C and 2.5µl of Safe View dye (NBS Biologicals) was added after which the agarose was poured into a gel cast mould (10x15 cm) (Bio-Rad Laboratories, UK). Following cooling the gel was placed in a tank (Bio-Rad Laboratories, UK) and overlaid with 800ml of 1% TBE buffer supplemented with 30 ul of Safe View dye. The gel was then loaded with 10 ul volume of amplified PCR product in loading buffer.

Gels were subjected to electrophoresis for 45 minutes at 100 V (constant). The size of the PCR products was determined by comparison with a GelPilot 1 kb Plus Ladder (100-10,000bp) (Qiagen, Crawley, UK).

Gel images were acquired by a CCD camera with Bio-Rad Image Lab™ software version 3.0. Each PCR assay was repeated three times for each DNA sample template.

5.2.7 Induction of bacteriophage expression from PCR negative strains of *C.difficile*

In an attempt to isolate bacteriophages which had not been previously described we employed the mitomycin C induction method described in Chapter 3 section (3.2.4) on strains which were PCR negative for bacteriophage endolysin. *C.difficile* strains from a previous study were used as a positive control strain (NCTC12727, inducible strain) and negative control strain (DS1813, non-inducible stain) . Following mitomycin C treatment the cultures were screened for the presence of bacteriophages by electron microscopy as described in Chapter 3 section (3.2.7).

5.3 Results

5.3.1 Bioinformatics analysis

C.difficile bacteriophages which have been isolated and sequenced to date have been shown to share sequence homology of at least 59%, despite the fact that they belong to different morphological families.

Genomic details for each bacteriophage are shown in table 5.4 . Table 5.5 shows the percentage of homology between these bacteriophages across the entire genome.

Bacteriophage	Accession number	Genome length (bp)	GC content (%)	Bacteriophage family
φCD119	AY 855346	53325	28.7	dsDNA, *Myoviridae*
φC2	NC 009231	56538	28.72	dsDNA, *Myoviridae*
φCD27	EU 719189	50930	29.4	dsDNA, *Myoviridae*
φCD6356	GU 94955	37664	28.4	dsDNA, *Siphoviridae*
φCD38-2	NC 015568	41090	30.83	dsDNA, *Siphoviridae*

Table 5.4. *C.difficile* bacteriophages: genomic details.

Bacteriophage	Bacteriophage	% of homology
φC2	φCD27	69.0
φC2	φCD6356	66.0
φC2	φCD119	64.0
φC2	φCD38-2	63.0
φCD27	φCD6356	64.0
φCD27	φCD119	59.0
φCD27	φCD38-2	60.0
φCD6356	φCD119	64.0
φCD6356	φCD38-2	59.0
φCD119	φCD38-2	61.0

Table 5.5. shows the percentage of homology between the *C.difficile* bacteriophages used in this study.

The gene sequences of the five endolysins were compared using the Multiple Sequence Alignment (MSA) program ClustalW2 (European Bioinformatics Institute) to determine the degree of similarity between the endolysins. As can be seen from figure 5.1. the endolysins cluster in three distinct groups, LysCD27 and LysC2, LysCD6356 and Lys CD38-2 and finally Lys CD119.

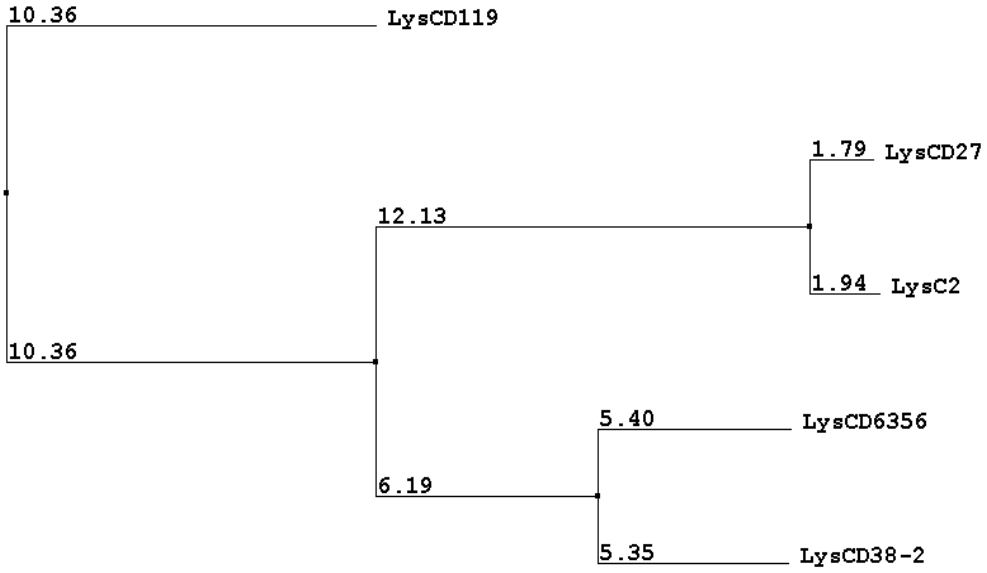


Figure 5.1. Neighbor-joining phylogenetic tree showing the relationship between the endolysins of the following *C.difficile* bacteriophages φCD119, φCD27, φC2, φCD6356 and φCD38-2. The phylogenetic tree was constructed based on multiple alignments using the DNA sequences generated from Clustalw2.

Within each group the homology of the individual lysins was determined. In the case of LysCD6356 and LysCD38-2 they shared only 89% similarity. In contrast LysCD27 and LysC2 showed the greatest similarity at 96%, suggesting that they may have evolved from a common ancestor. This high level of homology could also be a result of gene transfer between the bacteriophages and their hosts and gene transfer among the bacteriophages (Meessen-Pinard *et al*., 2012).

Comparison of the nucleotide sequences revealed that the endolysin genes were more closely related than the holin genes which have previously been used as phage genetic markers for studying the prophage diversity within *C.difficile* isolates (Shan *et al*., 2012) (Table 5.6).

Indeed the endolysin sequences were more conserved than the capsid genes which have also been used to study phylogenetic relationships between bacteriophage bacterial species (Nale *et al*., 2012).

Bacteriophage	**Bacteriophage**	**Endolysin gene homology (%)**	**Holin gene homology (%)**
φCD27	**φ**C2	96.0	87.0
φCD27	**φ**CD119	63.0	87.0
φCD27	**φ**CD6356	69.0	53.0
φCD27	**φ**CD38-2	70.0	49.0

Table 5.6. Multiple sequence alignment showing percent identity between bacteriophage endolysin and holin genes of φCD27 and other *C.difficile* bacteriophages using ClustalW2.

According to Oakley *et al.* (2011) the endolysin genes of four closely related *C.perfringens* bacteriophages were the most variable (ranging from 69 to 100% similarity) of the bacteriophage genes, possibly due to the fact they were under different selective pressures than the rest of the genome. In contrast, the holin genes were the most conserved genes (ranging from 97 to 100% similarity) among these bacteriophages (Oakley *et al.,* 2011). This does not appear to be the case for the *C.difficile* bacteriophages examined in this study.

A major aim of this study was to determine how widely each endolysin was distributed across our study population of clinical isolates of *C.difficile,* and for this reason two sets of primers were designed for each endolysin gene.

The endolysin genes from previously sequenced bacteriophages of *C.difficile* (Table 5.4) were used to design the primers employed in this study. They represented all of the *C.difficile* bacteriophage sequences deposited in GeneBank until 2012 but did not include the sequences of φMMP03 and φMMP04 which were published in 2012 (Meessen-Pinard *et al*., 2012). The specificity of individual primer sequences was confirmed by running nucleotide blast searches (NCBI).

5.3.2 PCR results

The $MgCl_2$ concentration, annealing temperature and primer pair concentration for each endolysin were individually optimized as described below. The Mg^{2+} concentration is a crucial parameter that can affect the performance of *Taq* DNA polymerase. High levels of Mg^{2+} may lead to false negative reaction whereas low levels may cause failure of the PCR reaction (Innis and Gelfand, 1990). The effect of altering the Mg^{2+} concentration on the PCR reaction using primers for the LysC2 gene is shown in figure 5.2.

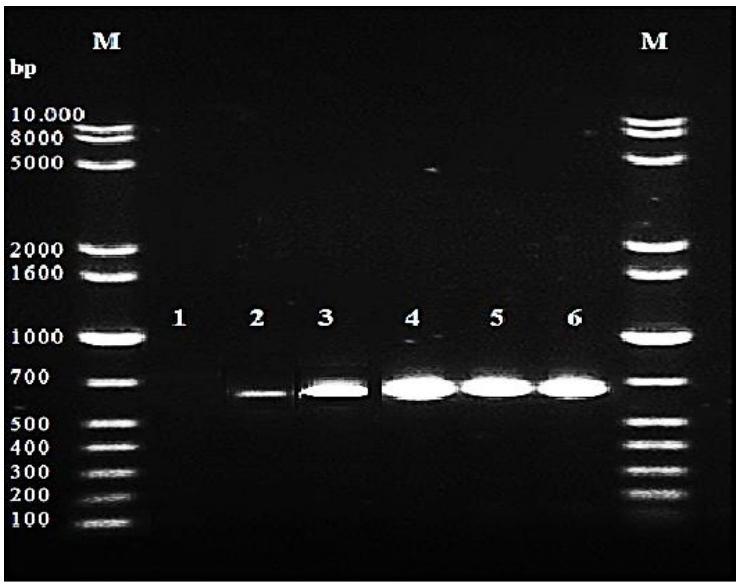


Figure 5.2. Optimization of $MgCl_2$ concentration for the detection of endolysin using LysC2 primer-1 and genomic DNA from *C.difficile* DS1813 (ribotype 027) as the template. The expected product size is 638 bp. Lane M shows a GelPilot 1 kb Plus Ladder (100-10,000bp). Lanes 1 to 6 show the PCR products obtained using (**0.5, 1.0 ,1.5 ,2.0 ,2.5 and 3.0** mM) $MgCl_2$.

From this result we concluded that the optimum $MgCl_2$ concentration was between 1.5 and 3.0 mM. A similar optimization was performed for all of the primer pairs.

The annealing temperature of each primer depends on the length and composition of the primer which may be above or below the estimated Tm (Taq PCR Handbook, 2012).

This is usually adjusted to improve the specificity and yield of PCR reaction (Wu *et al*., 1991). We determined the annealing temperature for each primer pair which gave the highest amount of specific product with the least amount of non-specific PCR product. As can be seen in figure 5.3 the annealing temperatures that yielded the greatest amount of amplicon using primer-2 for LysCD27 were between 57°C and 58°C.

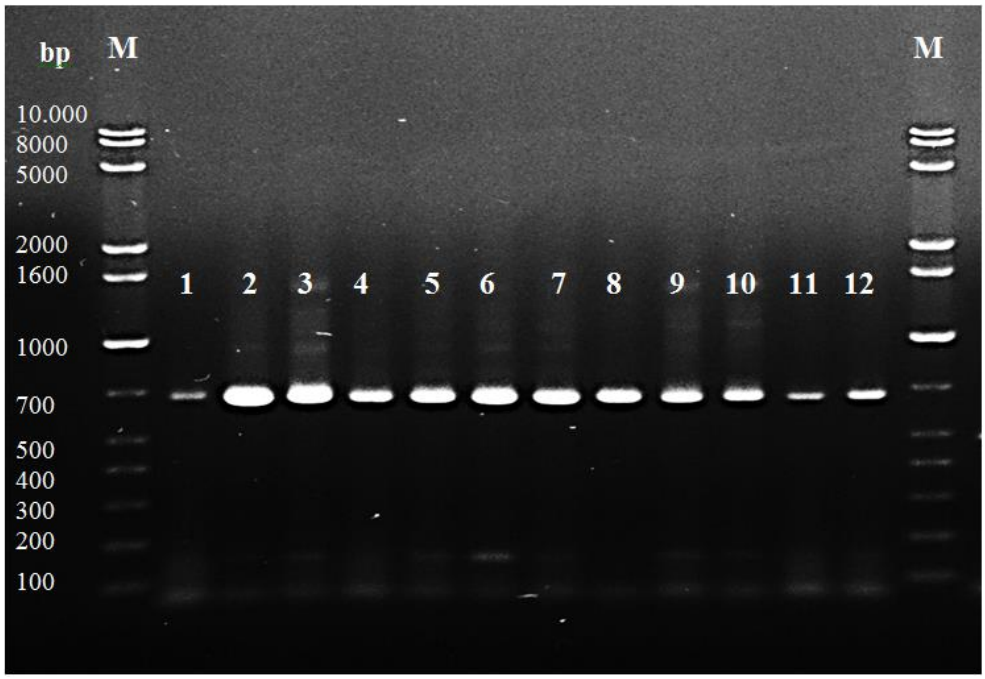


Figure 5.3. Optimization of annealing temperatures for the detection of endolysin using LysCD27 primer-2 and genomic DNA from *C.difficile* DS1759 (ribotype 001). The expected PCR product size is 672bp. Lane M: GelPilot 1 kb Plus Ladder (100-10,000bp). Lanes 1 to 12 show the PCR products obtained using (**55 to 66 ^{0}C**) annealing temperature need more detail.

The optimum annealing temperature for each primer was determined individually, but all primers seem to produce the best annealing result between 55 and 60^{0}C (Table 5.1). Finally we determined the effect of altering the primer concentration on the quality of the PCR product. As can be seen in the example shown in figure 5.4 the optimum endolysin primer pair concentrations were 1.0 and 0.5 µM.

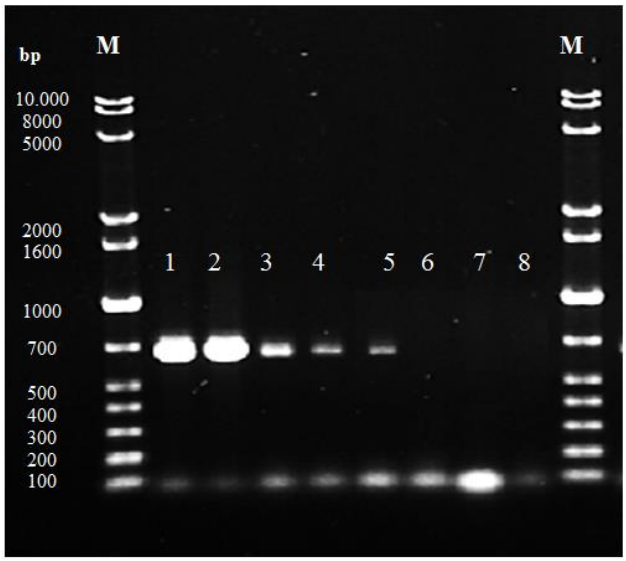


Figure 5.4 Optimization of primer concentration for the detection of endolysin using LysCD38-2 primer-2 with genomic DNA from *C.difficile* DS20291 (ribotype 027). The expected product size was 700bp. Lane M: GelPilot 1 kb Plus Ladder (100-10,000bp). Lanes 1 to 7 show the PCR products obtained using (**1.0,0.5 ,0.4 ,0.3 ,0.2 ,0.1 ,0.05 and 0.025** µM) of primer .

The optimized PCR reaction conditions for all primer pairs are those described in table 5.3 and they were employed throughout the remainder of the study.

Following PCR optimization specific oligonucleotide primers for each endolysin gene were run against genomic DNA from 85 isolates of *C.difficile*, 12 isolates of *C.perfringens* and individual isolates of *C. sordellii* R20453, *C. septicum* R22030 and *C.novyi* R14479.

The PCR products obtained following amplification using primers specific to LysCD27 are shown in figures 5.5 and 5.6. The expected size of the amplicon is 672 nucleotides. The samples were run on two separate gel due to the number of samples involved.

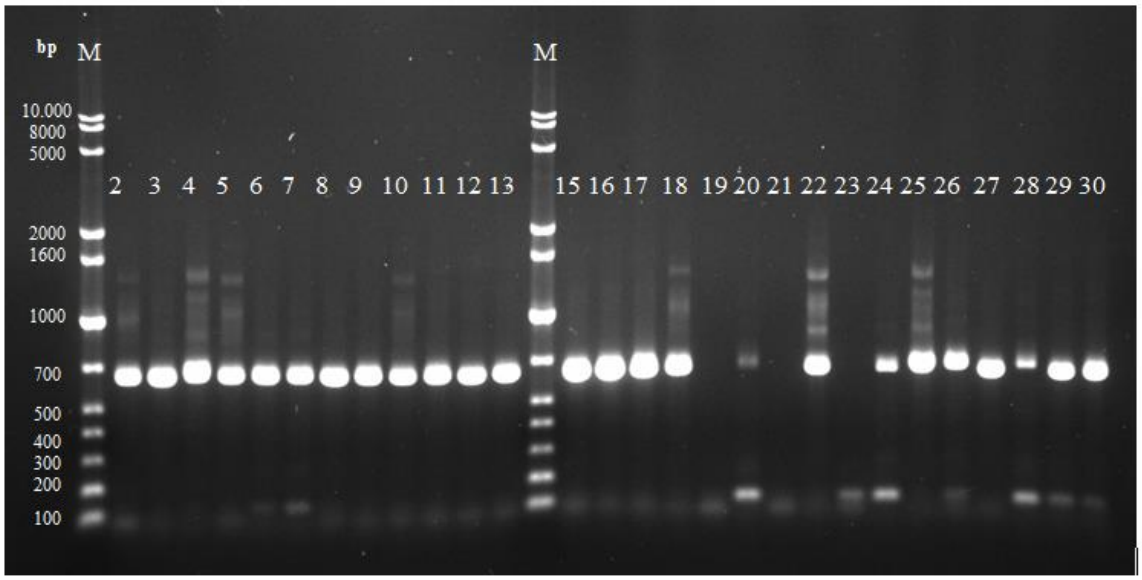


Figure 5.5. The results obtained from screening genomic DNA from isolates of *C.difficile* representing a range of ribotypes with PCR primers 2 specific to LysCD27. The expected size of the amplicon is 672 bp and PCR products were separated by electrophoresis on a 1% gel. Lane M: GelPilot 1 kb Plus Ladder (100-10,000bp).Lanes 2-6: 001 ribotype.; Lanes 7-11: 027 ribotype. Lanes12-17:106 ribotype. Lanes 18-22: 014 ribotype. Lanes 23-27: 002 ribotype. Lanes 28-30: 005 ribotype.

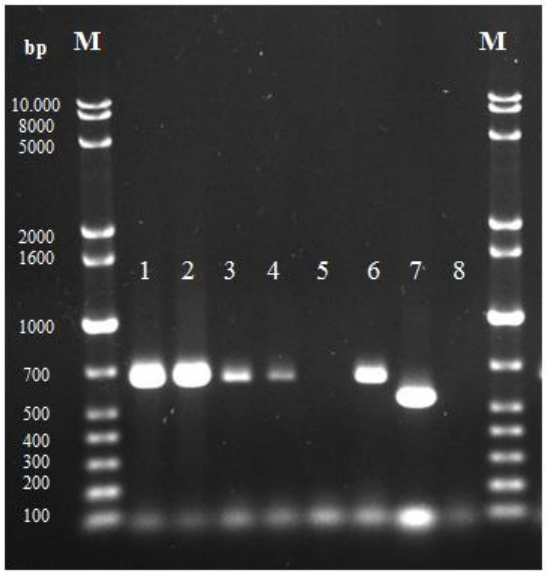


Figure 5.6. The results obtained from screening genomic DNA from isolates of *C.difficile* representing a range of ribotypes with PCR primers 2 specific to LysCD27. The expected size of the amplicon is 672bp and PCR products were separated by electrophoresis on a 1% gel. Lane M: GelPilot 1 kb Plus Ladder (100-10,000bp). Lanes 1-6: 012 ribotype. Lane 7: positive control primer-1 (toxin A gene) with *C.difficile* 630 (product size : 546 bp). Lane 8: negative control (no DNA).

Figure 5.7 summarizes the results obtained using primers which recognised the endolysins of the following members of the *Myoviridae* family: φC2, φCD27 and φCD119.

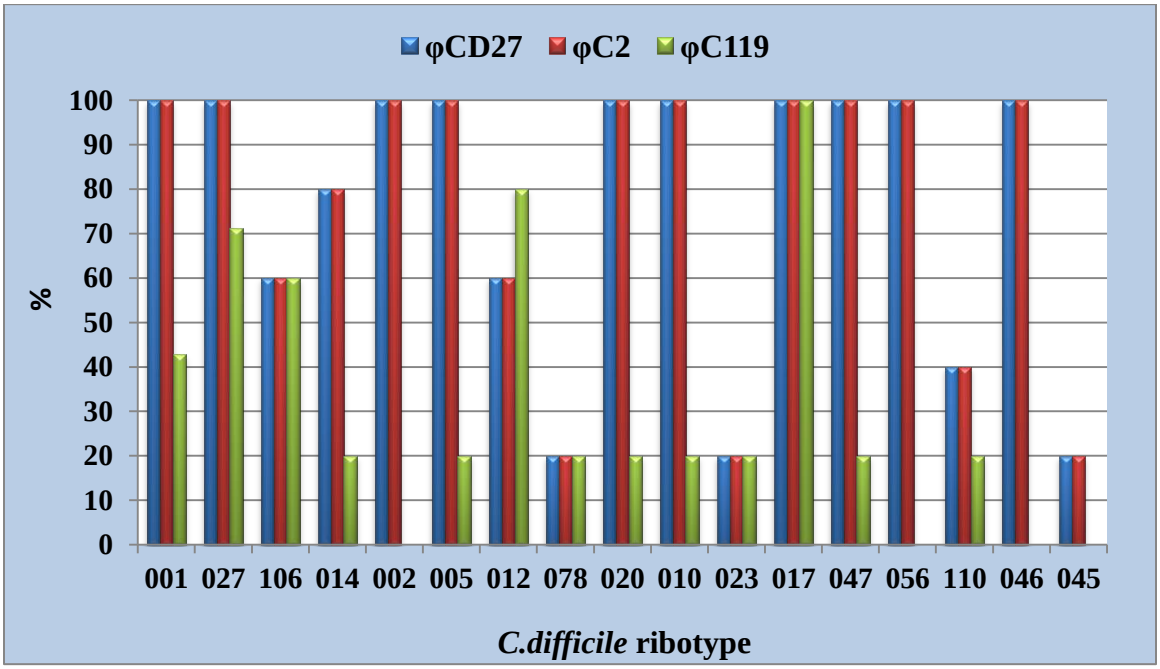


Figure 5.7. The percentage of *C.difficile* isolates of each ribotype (n=5) which produced an amplicon of the correct size suggesting the presence of a homolog to an endolysin of the *Myoviridae* bacteriophages φCD27, φC2 and φCD119.

The two bacteriophage endolysins with the greatest homology, φC2 and φCD27, occurred at an identical frequency in all of the strains tested. In contrast, φCD119 appeared less frequently and was absent from strains of the following ribotypes: 002, 056, 046 and 045.

Figure 5.8 shows the results obtained using primers which recognised the endolysins of members of the *Siphoviridae* family, φCD6356 and φCD38-2. Each lysin had the same percentage distribution for ribotypes 001, 027, 106, 002, 010, 017 and 046. The distribution across the other ribotypes was much more varied.

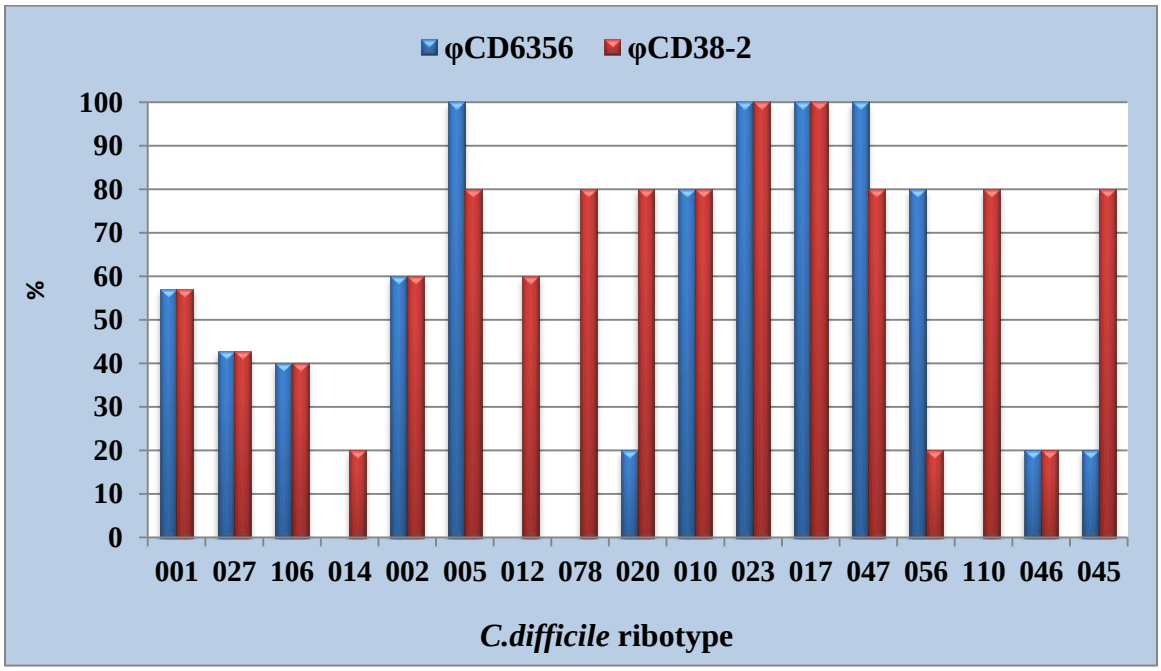


Figure 5.8. The percentage of *C.difficile* isolates of each ribotype (n=5) which produced an amplicon of the correct size suggesting the presence of a homolog to an endolysin of the *Siphoviridae* bacteriophages φ CD6356 and φCD38-2.

The frequency of endolysin detection for all of the bacteriophage derived targets across all of the *C.difficile* isolates tested in this study is summarized in figure 5.9. The most common hits were the endolysins of φCD27 and φC2 while the least commonly seen endolysin was that of φCD119.

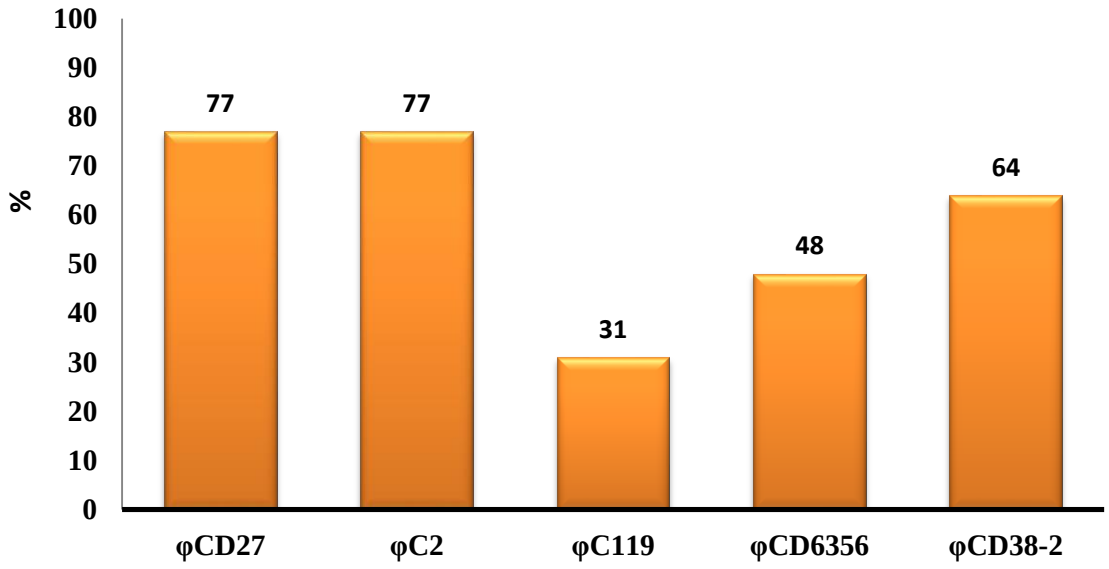


Figure 5.9. The percentage distribution of PCR hits for *C.difficile* bacteriophage endolysin genes from 85 isolates of *C.difficile*.

5.3.3 Induction of bacteriophage expression from PCR negative strains of *C.difficile*

Interestingly, we observed six isolates which failed to yield any product when probed with these primers. These strains were DS1787, R28614 (ribotype 106), R31710, R31708 (ribotype 078), R30967 (ribotype 110), and R30072 (ribotype 045) . This failure to amplify on repeated occasions (six separate reactions and on each occasion the positive controls produced the expected amplicons) suggested that either the isolates lack the endolysin gene and by extension the whole bacteriophage or that the DNA sequence recognized by the primers had altered in some manner.

To determine if these PCR negative isolates contained inducible bacteriophages we employed the mitomycin C induction method described in Chapter 3 (3.2.4). Two of the isolates, R30967 (ribotype 110) and R30072 (ribotype 045) yielded detectable bacteriophage like particles when examined by electron microscopy.

The two bacteriophages had an isometric capsid with a diameter of 97-52 nm connected to a contractile tail 180-150 nm in length and approximately 18-13 nm in width (Table 5.7). They were structurally similar to previously described *C. difficile* bacteriophages and can be classified as members of the *Myoviridae* family of the order *Caudovirales.* We have made several attempts to identify a susceptible indicator strain that supports virulent phage production of these two phages through plaque assays methods described earlier in Chapter 3. Unfortunately, no sensitive strain was found despite testing them on 85 different *C. difficile* strains

Induction of other negative strains failed to identify any new phage particles under the induction conditions described in Chapter 3, section (3.2.4).

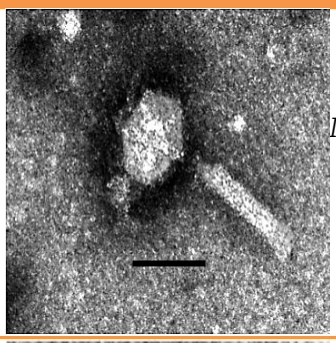
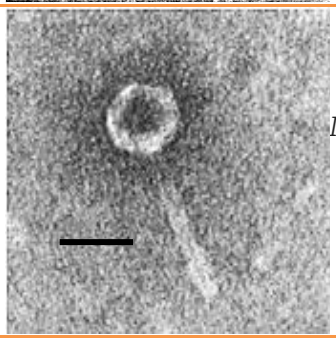
Ribotype	Strain	Bacteriophage	Morphology	Tail	Capsid
110	**R30967**		***Myoviridae***	**(180±0.5)x(18.4±0.6)**	**97.2±0.5**
045	**R30072**		***Myoviridae***	**(150±0.4)x(13.3±0.3)**	**52.7±0.6**

Table 5.5. Electron microscopy images of temperate bacteriophages isolated from PCR negative isolates of *C.difficile*, R30967 and R30072. Bacteriophages were induced with 3ug/ml mitomycin C.

5.3.4 Putative bacteriophage endolysin from *C.difficile* 630 (LysCD630)

A gene (NC_009089 region: 2201141-2201929) of 719 bp which encodes a putative phage endolysin (Sebaihia *et al*., 2006; Monot *et al*., 2010) was identified in the *C. difficile* 630 genome sequence through sequence similarity analysis using blast search (BLASTn) with *C.difficile* bacteriophage endolysins. This gene is located outside prophage regions that were identified in *C.difficile* 630 (prophage 1 and 2) (Sebaihia *et al*., 2006; Monot *et al*., 2010). Furthermore, BLASTn search revealed that a homolog (96% homology or higher) of this gene has been found in all nine strains of *C.difficile*

sequenced to date (Sebaihia *et al.*, 2006; Stabler *et al.*, 2009; He *et al*., 2010), suggesting that it has an essential function.

The endolysin protein has a structural organization similar to those seen in *C.difficile* bacteriophage endolysins with an N-terminal domain comprising an amidase 3 class enzyme and a putative C-terminal domain.

Using PCR primers designed as part of this study we found that an amplicon of the correct size (564 bp) was produced by all of the *C. difficile* isolates analysed (n=85) in the study. As can be seen in figures 5.10 and 5.11 PCR reactions using primers LysCD630-1 against genomic DNA isolated from strains of *C.difficile* representing different ribotypes yielded a product of the expected size. Testing of the remaining isolates yielded identical results.

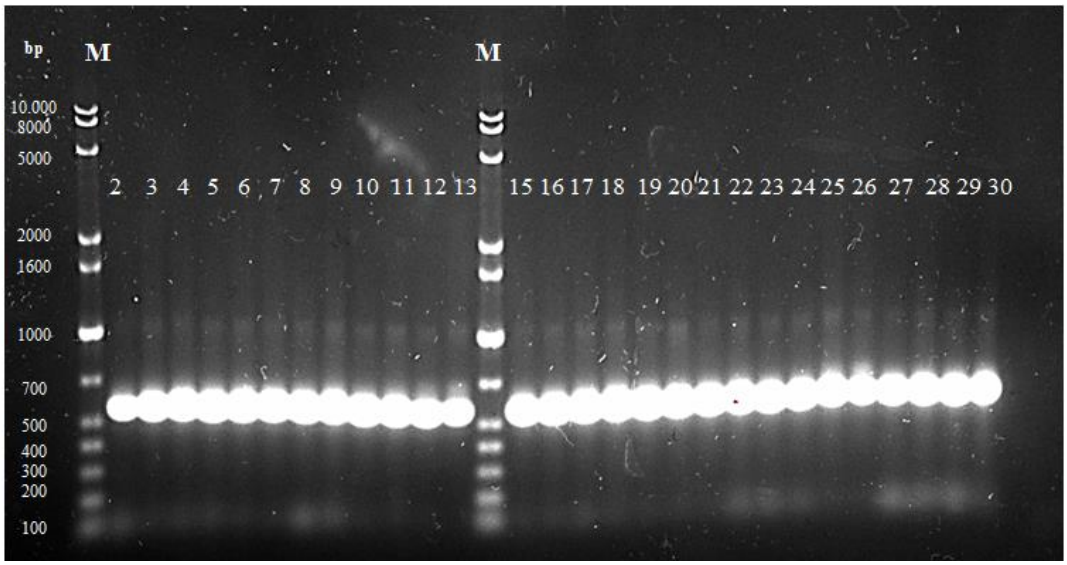


Figure 5.10. The results obtained from screening genomic DNA from isolates of *C.difficile* representing a range of ribotypes with PCR primers 1 specific to LysCD630. The expected size of the amplicon is 564 bp and PCR products were separated by electrophoresis on a 1% gel. Lane M: GelPilot 1 kb Plus Ladder (100-10,000bp). Lanes 2-6: 001 ribotype. Lanes 7-11: 027 ribotype. Lanes 12-17:106 ribotype. Lanes 18-22: 014 ribotype. Lanes 23-27: 002 ribotype. Lanes 28-30: 005 ribotype.

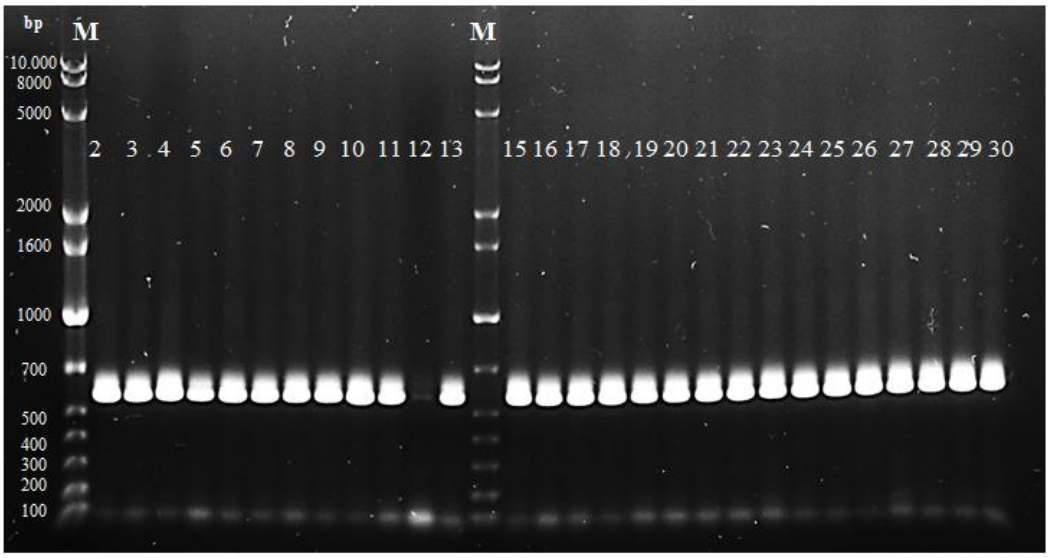


Figure 5.11. The results obtained from screening genomic DNA from isolates of *C.difficile* representing a range of ribotypes with PCR primers 1 specific to LysCD630. The expected size of the amplicon is 564 bp and PCR products were separated by electrophoresis on a 1% gel. Lane M: GelPilot 1 kb Plus Ladder (100-10,000bp). Lanes 2-3: 005 ribotype. Lanes 4-8: 012 ribotype. Lanes 9-15: 078 ribotype. Lane 12: negative control, Lanes 16-20: 020 ribotype. Lanes 21-25: 010 ribotype. Lanes 26-30: 023 ribotype.

As expected, all of the other clostridia strains failed to generate PCR products with any of the endolysin specific primers.

5.4 Discussion

5.4.1 Use of the phage endolysin gene as a marker of prophage carriage

Bacteriophages are known to have a major influence on the activity, structure, biological diversity and evolution of a microbial community (Letarov and Kulikov, 2009). Therefore by studying bacteriophages diversity across multiple *C.difficile* isolates we hope to be able to gain a better understanding of the role of these mobile genetic elements in shaping the biology of their host genome.

For example, it has been reported that some lysogenic phages are able to increase the production of toxin B in toxigenic strains of *C.difficile* while having no effect on toxin A production (Goh *et al.,* 2005b). Interestingly, a homolog of one of the genes thought to be involved in the regulation of toxin expression, *tcdE*, has been detected in the genomes of three phages belonging to the *Myoviridae* family, φC2, φC6 and φC8, suggesting that this gene could be a remnant of an ancient prophage (Goh *et al.,* 2005a).

Further evidence that phages could play a role in the regulation of host gene expression was provided by Govind and colleagues who observed a decrease in the level of expression of the genes in the pathogenicity locus (PaLoc) in strains which had been lysogenized by φCD119 (Govind *et al.,* 2009). This effect was attributed to the presence of DNA homology between *repR***,** a putative repressor gene of φCD119, and DNA sequences upstream of tcdR, which causes the down regulation of *tcdA* and *tcdb* promoters (Govind *et al.,* 2009). Furthermore, a recent study suggests that lysogenization

of a NAP1/027 isolate with φCD38-2 led to increased production of both toxin A and B toxin (Sekulovic *et al*., 2011).

In addition to influencing the expression of host genes bacteriophages are also capable of conferring resistance to infection by related lysogenic phages (or superinfection immunity), which may explain the narrow host range for *C. difficile* phages (Mayer *et al.,* 2008; Fortier and Moineau, 2007; Horgan *et al*., 2010;Nale *et al*., 2012).That said it is not uncommon for a bacterial host such as *C.difficile* 630 to harbor more than one unrelated prophage in its genome (Sebaihia *et al*., 2006; Monot *et al*., 2010), suggesting isolates of *C.difficile* are repeatedly under attack by bacteriophages.

So far, only seven bacteriophages, namely φCD119, φCD2, φCD27, φMMP03, φMMP04, φCD6356 and φCD38-2, have been isolated from *C.difficile* and studied at the molecular level. The genetic diversity of these bacteriophages and their possible role in the biology of *C. difficile* has yet to be fully determined. The studies which have been undertaken have focused on comparing the similarity of two genes which encode the major capsid protein and the holin (Nale *et al*., 2012 : Shan *et al.,* 2012). Unfortunately homologs of these genes have not been found in all of the *C.difficile* specific phages sequenced to date. For example, the genome of φCD27 (Mayer *et al*., 2008) has no recognizable major capsid gene(Nale *et al*., 2012). In a recent paper Shan and colleagues (2012) isolated phages from isolates of *C.difficile* belonging to ribotypes 078 and 220 which also lacked a recognizable major capsid gene (Nale *et al*., 2012; Shan *et al.,* 2012). Indeed the data suggest that this gene target is only conserved among bacteriophages belonging to the *Myoviridae* family (Nale *et al.*, 2012).

The holin gene is also subject to considerable diversity: for example, the holin gene of the newly sequenced *C.difficile* bacteriophage φMMP04 (Meessen-Pinard *et al*., 2012) had little similarity (41 to 50%) to previously identified holin genes (Shan *et al*., 2012).

For this reason we hypothesized that bacteriophage endolysins could serve as a more effective marker of current or past phage infection. Indeed bioinformatics analysis of all of the *C. difficile* bacteriophages sequenced to date have revealed the presence of a recognizable endolysin gene .

To determine if homologs to these genes were present in other clinical isolates of *C.difficile* we designed gene specific PCR primers and used them to screen our collection of isolates.

A homolog of at least one of the bacteriophage endolysins was found in 79 of the 85 isolates, suggesting that the majority of our study population may had been infected by a bacteriophage. The primers which gave the greatest number of positive hits (77% each) were those which recognised the endolysins present in φCD27 and φC2, two closely related bacteriophages of the *Myoviridae* family. In contrast, the endolysin of φCD119 was only detected in 28% of the isolates. Thus, despite the fact that these bacteriophages are members of the same family they differ markedly in their ability to infect the same *C.difficile* isolates.

There are at least three possible explanations as to why this occurs; the phages bind to different bacterial receptors or differ in their ability to recognize the same receptor; the host strains are currently infected with a lysogenic phage of the same family; and finally infection with a related phage in the past has generated immunity which is mediated by the CRISPR system.

As previously described (Chapter 1) the role of the CRISPR system is thought to be to prevent superinfection by other bacteriophages (Sebaihia *et al.*, 2006; Sorek *et al*., 2008; Karginov and Hannon, 2010). Thus it is tempting to speculate , based on our results that repeated infection with different bacteriophages has resulted in a bacterial population which is inherently resistant to infection with temperate phages.

In contrast, the primers with specificity to endolysins of members of the *Siphoviridae* bacteriophage family φCD6356 and φCD38-2 gave positive results with 48% and 59% of the isolates respectively. The limited results of this study suggest that infection with members of the *Siphoviridae* family is less common than that seen with bacteriophages of the *Myoviridae* family. This observation is supported by previous results which showed that of all of the *C.difficile* phages isolated to date 28 belong to the *Myoviridae* family while only 11 belong to the *Siphoviridae* (Chapter 1, table 1.2).

In addition to observing differences in the degree to which individual endolysins were detected across the entire collection of clinical isolates we also observed differences in the responses of individual ribotypes. All strains belonging to ribotype 017 appear to have been infected by all of the phages examined in this study. In contrast, strains belonging to certain ribotypes appeared to have been infected by only one phage family. For example, isolates DS1759 and 1747 (ribotype 001) gave hits with primers specific to endolysins belonging to members of the *Myoviridae* family while all of the ribotype 023 isolates (n=4) were only infected by *Siphoviridae* family phages. Therefore, it can be concluded from these results that there is no correlation between phage infection and

C.difficile ribotype. Owing, however, to the low number of samples used per ribotype, further work would be required to improve the result.

In addition to generating information concerning the distribution of bacteriophage endolysins across our collection of clinical isolates we also sought to identify new bacteriophages. Our hypothesis was that clinical isolates which gave no hits when probed with our PCR primers may carry previously unseen bacteriophages. To determine if this was the case we treated PCR negative isolates of *C. difficile* with mitomycin C and screened for the release of bacteriophages. Two of the six PCR negative isolates yielded low levels of bacteriophages upon induction. Examination by electron microscopy showed a clear resemblance to other *C. difficile* bacteriophages and their morphology indicated that each phage belonged to the *Myoviridae* family. Unfortunately the yield of each bacteriophage was such that we were unable to isolate sufficient bacteriophages to perform sequencing studies and also we failed to identify an indicator strains that supports these phage propagation. Consequently we were unable to characterize the endolysin associated with these bacteriophages.

5.4.2 Putative bacteriophage endolysin from *C.difficile* 630 (LysCD630)

This endolysin was identified owing to its sequence homology with endolysins of known *C.difficile* phages. Moreover the nucleotide sequence of this gene was found to be highly conserved (at least 96%) amongst all *C.difficile* isolates sequenced to date. As one would expect given the sequence homology, the protein sequence of this gene has a structural organization similar to those seen in *C.difficile* bacteriophage endolysin with an N-terminal domain comprising an N-acetylmuramoyl-L-Ala amidase class enzyme and a putative C-terminal domain. We sub-cloned and expressed the corresponding protein in *E.coli* (M15) and have confirmed its lytic activity against the vegetative form of a number of different strains of *C.difficile* (Chapter 4).

The presence of a homolog to this endolysin in all of the *C.difficile* isolates tested in our study suggests that this gene could represent an evolutionary relic of past phage infection of an early ancestor of *C.difficile*. Further characterization of the expression and function of this gene is required to determine what role, if any, it plays in the biology of this strain.

5.5 Conclusions

- The results of this study suggest that prophage carriage is a common event for clinical isolates of *C.difficile*.

- This approach was able to identify what appear to be novel bacteriophages.

- Propagation of these novel phages and characterisation of their endolysins could reveal new classes of endolysin with therapeutic potential.

- PCR negative strains could also be used as an indicator strain for the previously isolated phages, because such strains are assumed to be sensitive.

- We identified an endolysin from the genome of CD630 which was present in every clinical isolate of *C.difficile* examined during our study. This gene might have potential as a diagnostic marker. Owing to its universal and conserved nature further studies are warranted to determine its role in the physiology of the host bacteria.

Chapter Six

Germination-specific cortex-lytic enzymes of *Clostridium difficile*

6.1 Introduction

Dormant Spores of *Bacillus* and *Clostridium* species are extremely resistant and can survive for years; these features are largely attributed to their unique structures (Popham *et al.,* 1995). These endospores are generally formed under conditions of environmental stress such as a reduction in the availability of nutrients (Popham *et al.,* 1995). In the case of *C.difficile* this ability to form spores allows the organism to survive on contaminated surfaces for extended periods of time and increases its resistance to chemical and physical agents (Setlow, 2007). It is this feature that makes *C. difficile* an enormous challenge for infection control systems in hospitals and health care facilities; therefore, there is a need for new strategies to reduce the burden of *C. difficile* spores in contaminated environments (Nerandzic and Donskey, 2010).

The mechanisms by which this pathogen forms spores and then subsequently germinates are poorly understood, due, until recently, to a lack of appropriate genetic tools with which to study these processes (Burns *et al.,* 2010). Access to the bacterial genome sequencing projects has enabled researchers to search for homologs to genes known to be involves in the germination of other spore-forming bacteria (Burns *et al.,* 2010)

6.1.1 Structure of the *C.difficile* spore

Spores of *Bacillus* and *Clostridium* are normally surrounded by several envelope layers, as described in the introduction, and their role is to protect the organism and interact with the environment (Russell, 1990).

Structurally, the bacterial spore comprises a core containing nucleic acid and enzymes, which is surrounded by the spore's inner membrane. Next comes the cortex region, which comprises a major part of the spore and consists of a layer of spore peptidoglycan. This cortex is surrounded by an inner spore coat, while the outer surface of the spore coat is surrounded in some species by a closely fitting structure known as the exosporium (Gerhardt and Ribi, 1964; Russell, 1990; Lawley *et al.*, 2009).

6.1.2 Structure of the cortex peptidoglycan

Following the triggering of germination, spore lytic enzymes are activated, which break down the structure of the peptidoglycan within the spore cortex (Moir, 2006). While little is known about the composition of the peptidoglycan of the cortex of *C.difficile* spores, it has been reported that the peptidoglycan of the spore and vegetative forms of *B. subtilis* differ markedly (Atrih and Foster, 1999). In the case of *B.subtilis* spore peptidoglycan, only 3% of *N*-acetylmuramic acid (NAM) residues carry peptide chains which are cross-linked via tetrapeptide side chains (Figure 6.1). A further 26% of NAM residues have tetrapeptide side chains, 20% have L-alanine side chains, while a small percentage carry tripeptides (Leighton and Doi, 1971; Popham *et al.*, 1996; Atrih and Foster, 1999). In addition, Muramic acid δ-lactam, which has no side chains, accounts for almost half of the muramic acid residues in spore peptidoglycan (PG), and as a consequence, spore PG has considerably less structural integrity than vegetative cell PG (Warth and Strominger, 1969; Warth and Strominger, 1972; Smith, Blackman and Foster, 2000; Meador-Parton and Popham, 2000; Dowd, *et al.*, 2008).

A similar peptidoglycan structure is also found in the spore cortex of *B. cereus*, *B. megaterium* and *C. botulinum* spores (Popham *et al.*, 1996; Atrih and Foster, 2001). In *C.perfringens* spore coat PG, no NAM residues have single alanine side chains and some NAM are connected with dipeptide terminating in D-glutamate and with tripeptide terminating in diaminopimelic. Like *B. subtilis* s, only 3% of NAM residues are cross linked (Marquis and Bender, 1990; Makino and Moriyama, 2002).

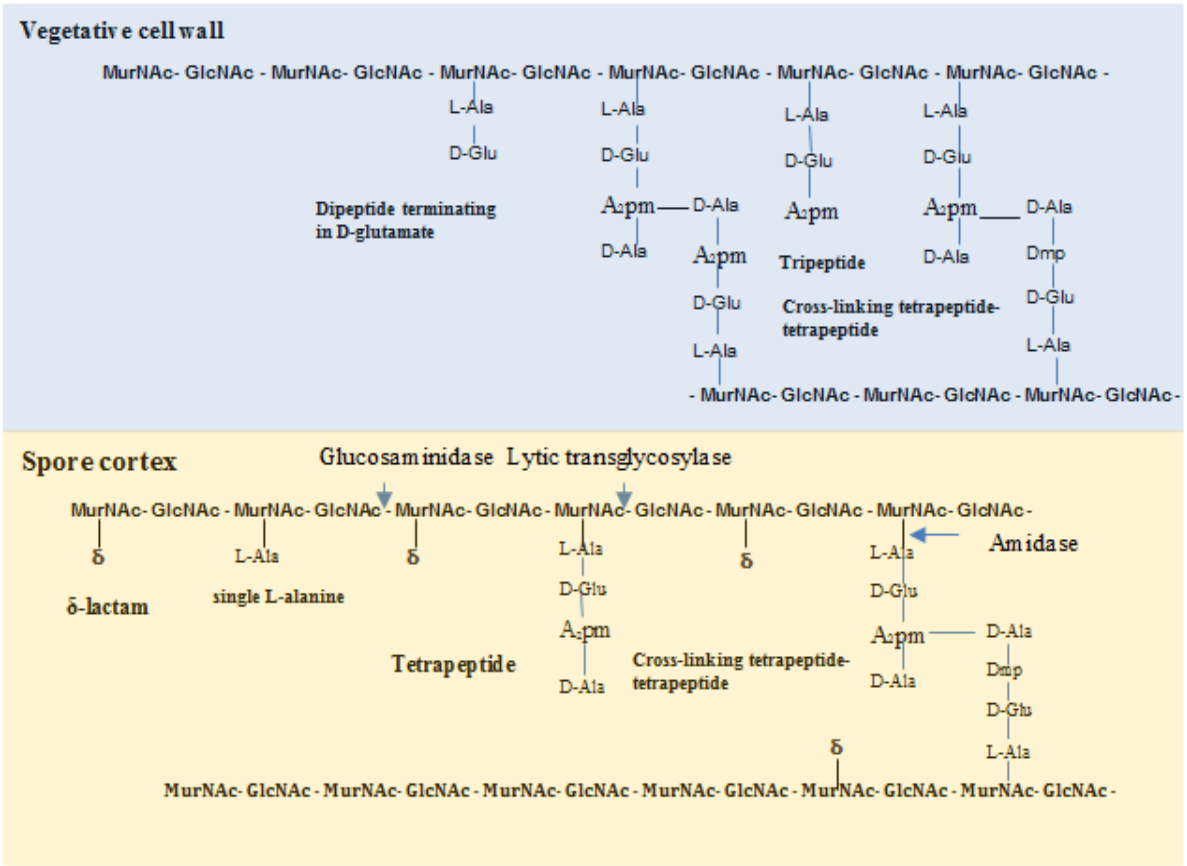


Figure 6.1. Simplified schematic structure of the peptidoglycan composition of the vegetative cell wall and spore cortex of *B.subtilis* , showing the representative action of germination-specific cortex-lytic enzymes (GSLEs) (glucosaminidase and lytic transglycosylase) involved in cortex lysis during germination. GlcNAc, *N*-acetylglucosamine; MurNac, *N*-acetylmuramic acid; L-Ala, L-alanine; D-Glu, D-glutamic acid; meso-A_2pm, meso diaminopimelic acid; D-ala, D-alanine; *δ*, *δ*-lactam. Modified from Smith, *et al*, (2000) ,Atrih and Foster (2002) and Kumazawa *et al*., (2007).

6.1.3 Germination-specific cortex lytic enzymes

During germination, hydrolysis of the cortex is essential to allow the production of new vegetative cells and to initiate outgrowth (Paredes-Sabja *et al.*, 2009a). Hydrolysis is believed to be mediated by a group of recently identified enzymes called germination-specific cortex-lytic enzymes (GSLEs) (Makino and Moriyama, 2002; Okamura *et al.*, 2000; Urakami *et al.*, 1999). According to Makino and Moriyama (2002), these enzymes can be classified into two types based on their activity: spore cortex-lytic enzymes (SCLEs) and cortical fragment-lytic enzymes (CFLEs). Both SCLEs and CFLEs recognise the muramic-δ-lactam of spore peptidoglycan as a substrate but have no activity against vegetative cell peptidoglycans (Makino and Moriyama, 2002) and they are believed to be located in inner spore coat (Bagyan and Setlow, 2002).

In the case of *B. subtilis*, two SCLEs capable of degrading spore peptidoglycan muramic-δ-lactam have been identified (Popham *et al.*, 1996). The first is called SleB and is located in the inner and outer layers of the spore. It possesses lytic transglycosylase Muramidase activity and requires a second enzyme called YpeB for its activation (Boland *et al.*, 2000; Masayama *et al*., 2006).

The other SCLE is called CwlJ and is found in the spore coat in association with GerQ during sporulation; it requires CaDPA for activation and is essential for CaDPA-mediated germination (Paidhungat *et al.*, 2001; Bagyan and Setlow, 2002). Inactivation of both *cwlJ* and *sleB* results in a spore that is unable to complete cortex hydrolysis during nutrient germination (Ishikawa *et al.*, 1998).

Unlike the SCLEs of *Bacillus*, *Clostridium* SCLEs are less well-defined. In the case of *C.perfringens*, SleC (Miyata *et al*., 1995) and SleM (Chen *et al*., 1997) have been identified in all strains of the organism that have been sequenced to date and have been shown to possess the ability to degrade cortex peptidoglycan (PG) *in vitro* (Shimizu *et al*., 2002; Myers *et al*., 2006).

A recent study has demonstrated that only SleC is essential for cortex hydrolysis and viability of *C.perfringens* spores, whereas SleM is not essential (Kumazawa *et al*., 2007; Paredes-Sabja *et al.*, 2009a). Of considerable interest, in the context of this study, is the fact that the SleC of *C.perfringens* retains its lytic activity following extraction from the spore coat (Gould and Hitchins, 1963; Gombas and Labbe, 1981; Makino and Moriyama, 2002).

6.1.4 Characterisation of SleC of *C. perfringens*

During the process of *C.perfringens* spore development, SleC is produced at stages II to III of forespore formation (Masayama *et al*., 2006). It is thought to be synthesized as a 50 kDa precursor of 438 amino acids comprising four domains (Figure 6.2). These are an N-terminal presequence (113 amino acid residues); an N-terminal prosequence (35 amino acid residues); a mature enzyme with amidase activity (264 amino acid residues); and a C-terminal prosequence of 25 residues (Miyata *et al*., 1995; Urakami *et al*., 1999; Makino and Moriyama, 2002).

During sporulation (stages IV to VI) the N-terminal presequence and the C-terminal prosequence are cleaved to generate an inactive proenzyme with a mass of 35 kDa , which is localized outside the cortex layer (Miyata *et al*., 1995; Urakami *et al*., 1999). During the germination, The active form of this enzyme is a 31 kDa fragment (C_{31}), which is achieved by the proteolytic cleavage of an N-terminal prosequence from the 35 kDa fragment by a germination specific protease (GSP) (Urakami *et al*., 1999). This enzyme is a cysteine-dependent serine protease, which cleaves C_{35} at Val-149 and Val-150 to generate the mature form of the enzyme (Figure 6.2). GSP is thought to be located close to the spore cortex in proximity to the major cortex-lytic enzymes (Shimamoto *et al.*, 2001). The production of this protease is encoded by three genes, *cspA*, *cspB* and *cspC*, which are located upstream of the *sleC* gene (Shimamoto *et al*., 2001; Xiao *et al*., 2011). In addition, it has recently been shown that the germination-specific serine protease CspB is essential for converting inactive pro-SleC to an active enzyme (Paredes-Sabja, *et al.*, 2009). However, the mechanism of regulation of these genes is currently unclear (Paredes-Sabja *et al.*, 2009; Paredes-Sabja *et al.*, 2011).

6.1.5 Spore cortex-lytic enzymes of *C.difficile*

In the last decade, the genome sequences of a number of *C. difficile* strains have been completed, enabling researchers to search for homologs to genes known to be involved in *Bacillus* and *Clostridium* spore germination

A study undertaken by David *et al*. (2009) found that *C. difficile* 630 contained a gene (CD3563) with homology to the cwlJ and *sleB* genes of *B. subtilis* in that it shared 30%

and 45% amino acid identity respectively. A second *C.difficile* gene, CD0552, of 238 residue proteins has also been annotated as sleB but shares no similarity with any known SCLE (Sebaihia *et al*., 2006; Monot *et al*., 2011). A third *C.difficile* gene, CD0551, shares 53% amino acid identity with the SleC gene of *C. perfringens* (David *et al.*, 2009). Furthermore, equivalent homologs of these genes are also present in other *C.difficile strains*, such as the BI/NAP1/027 strain, R20291 (David *et al.*, 2009). Interestingly *C.difficile* contains no gene with homology to *C. perfringens* SleM.

The essential role of SleC in the germination of *C.difficile* spores was recently confirmed by David and colleagues, who found that while the inactivation of CD3563 (cwlJ/sleB-homolog) and CD0552 (sleB homolog) had no effect on the germination of both *C.difficile* 630 and R20291, the inactivation of CD0551 (sleC homolog) resulted in a mutant that was unable to form colonies following the induction of germination (David *et al.*, 2009).

Furthermore, genome analysis has revealed the presence of homologues of the Csp proteases of *C.perfringens* in the genome sequence of several *C. difficile* strains, such as 630 and R20291 (David *et al.*, 2009). For example, the genome of *C. difficile 630* strains contains CD2246 (annotated as *cspC*) that has 32% amino acid identity to both CspA and CspC of *C. perfringens* and CD2247 (annotated as *cspBA*) that shows 36% and 35% identity to *C. perfringens* CspA and CspB, respectively (David *et al.*, 2009). Thus it would appear, on the basis of gene homology, that *C.perfringens* and *C.difficile* share a common means of SleC activation and spore lysis.

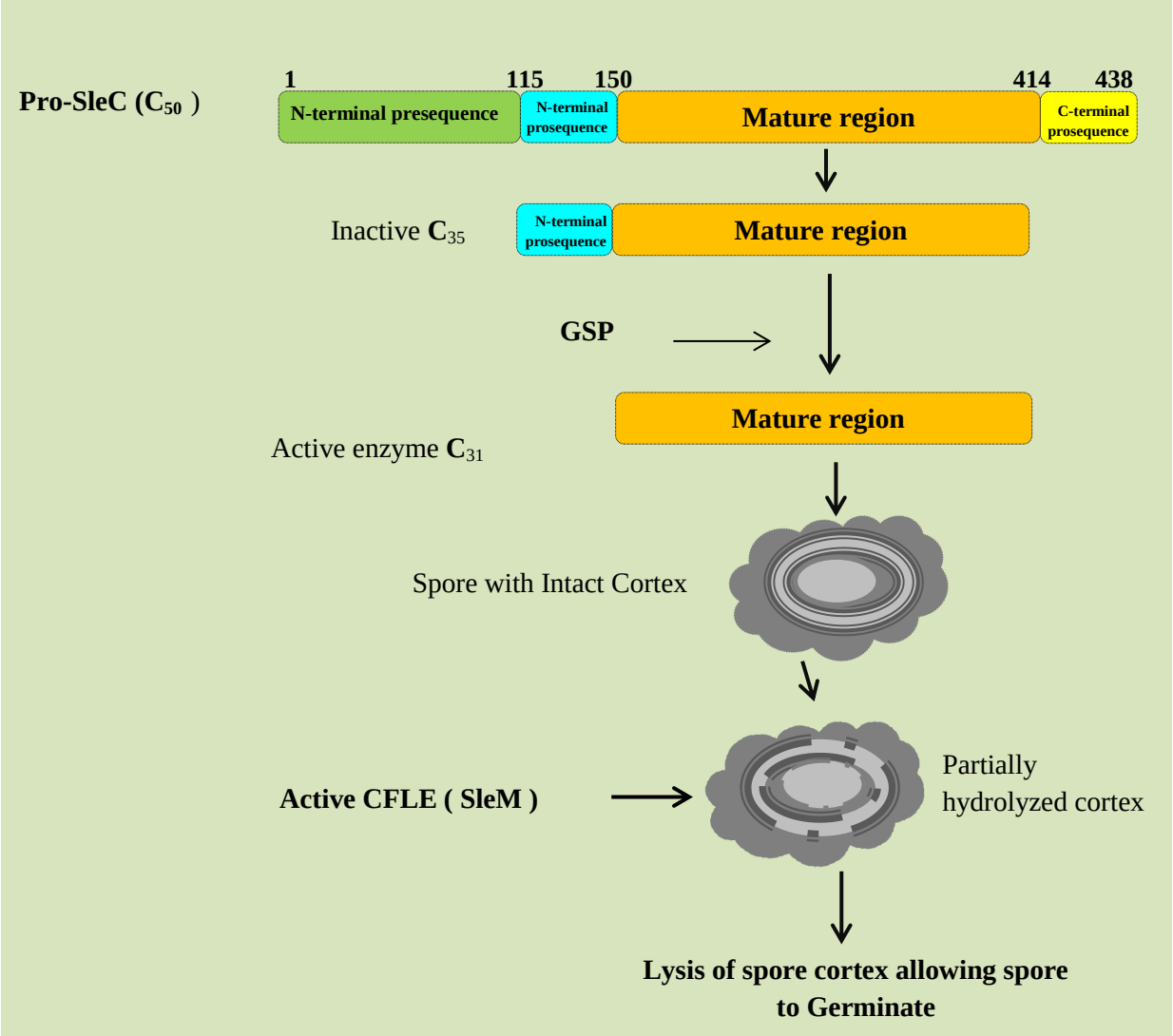


Figure 6.2. Schematic representation of a possible mechanism of cortex hydrolysis during germination of *C. perfringens* spores. Four domains in the diagram of pro SleC (C_{50}) represent the N-terminal presequence (115 aa), the N-terminal prosequence (35aa), the mature region (264aa) and the C-terminal prosequence (25aa) respectively. For full spore germination, lysis of Cortex PG is required, which is mediated by Spore Cortex-Lytic Enzymes (SCLE) and Cortex Fragment-Lytic Enzymes (CFLE). In the spore cortex, a pro SleC enzyme is activated by germination-specific proteases (GSP) by cutting its pro-site, causing partial hydrolysis of spore cortex, then (CFLE) SleM can function. Modified from Okamura *et al.* (2000), Makino and Moriyama (2002) and Xiao *et al.* (2011).

6.1.6 Aim of this study

The aims of this chapter are:

- To employ the SleC lysin of *C. perfringens* as a model to identify, clone and characterise the SleC of *C.difficile*
- To determine the ability of recombinant *C.difficile* SleC to lyse spore peptidoglycan

6.2. Materials and methods

This cloning project was approved by Cardiff University's Genetic Modification Safety Committee (GMSC) and the Health and Safety Executive (HSE) as a containment level 1 procedure.

6.2.1 Microbiological materials and methods

6.2.1.1 Bacterial strain, growth conditions and media

The *C. difficile* strains employed in this study were DS1813, CD630 and R20291. These strains were grown on CDMN agar, BHI agar and in BHI broth under anaerobic conditions at 37^0C as described in Chapter 2 (section 2.1.3).

C. perfringens strains R31973, R31782 and R31742 were cultivated in blood agar, BHI agar and BHI broth under an anaerobic environment at 37^0C as described in Chapter 2 (section 2.1.3). All strains were kindly donated by the Anaerobe Reference Laboratory, Cardiff.

E. coli M15 and SG3009 were used throughout as a cloning host, and were supplied as part of the QIA *express* System (Qiagen Ltd, Crawley, UK). All *E. coli* strains were cultured using Luria-Bertani (LB) broth or agar, as appropriate, at 37 °C, with 300 rpm shaking for liquid cultures. Media was supplemented as appropriate with kanamycin (25 μg/ml) and ampicillin (100 μg/ml).

6.2.1.2 Spore preparation

Spores of *C. difficile* strains were prepared as described previously in Chapter 2 section 2.1.6. *C. perfringens* spores were prepared by growth for 48 h on a BHI agar plate (supplemented with 5% w/v defibrinated horse blood; Oxoid, Basingstoke, UK). Plates were incubated anaerobically at 37^0C, after which two to three colonies were subcultured into sterile BHI broth (45ml) for five days. After incubation, cultures were repeatedly washed with ice-cold distilled H_2O (dH_2O) until no cell debris or vegetative cells could be observed following spore and Gram staining. The spores were resuspended in distilled water to a final optical density of 600 nm (OD_{600}) of 0.8 –1.0 and stored at -20°C.

6.2.1.3 Preparation of decoated spores

Spore produced as previously described were decoated chemically by agents which rupture disulfide bonds in the spore's coat (Gould and Hitchins, 1963) using the method described elsewhere (Popham *et al.* 1995 ; Paredes-Sabja *et al.*, 2009a ; David *et al*.,2009).Briefly, spores were suspended in 1 ml of 50 mM Tris-HCl (pH 8.0), 8 M urea 1% (wt/vol), Sodium dodecyl-sulfate and 50 mM dithiothreitol to an OD_{600} of 20 and incubated at 37°C for 90 min. Following this incubation, decoated spores were washed three times in PBS (Popham *et al.*, 1995). To confirm that the decoating process had no effect on spore viability, treated and untreated spores were serially diluted and plated onto BHI agar plates supplemented with lysozyme (1 µg/ml) and incubated overnight at 37 °C under anaerobic conditions.

6.2.1.4 Preparation of exosporium deficient *C.difficile* spores

C. difficile spores were sonicated in order to remove the exosporium using the method described by Redmond *et al.*, (2004), with some modifications. Briefly, prior to sonication, the spores were re-suspended to an optical density at 600 nm (OD600) of 1 in 1 ml of 50 mM Tris-HCl (Fisher Scientific, Leicestershire, UK) and 0.5 mM EDTA buffer (Acros Organics, Geel, Belgium) adjusted to pH 7.5.

To determine the optimal condition, the spores were sonicated using a Sanyo Soniprep 150 ultrasonic disintegrator (Sanyo Gallenkamp, Uxbridge, UK). During sonication, the tip of the probe was kept just below the surface of the samples, which were kept on ice at all times. The settings for the sonicator were as follows:

Amplitude (microns): 12 and 15

Number of bursts : 4, 7, 9 and 16

Length of bursts: 50 seconds and 1 minute

The samples were allowed to cool for 98 seconds in ice between each pulse.

Sonicated spores were then centrifuged at 9,000xg for 5 minutes. On completion, the spore pellets were washed once in 50 mM Tris-HCl, 0.5 mM EDTA (pH 7.5) and stored at 4^0C for further analysis.

6.2.1.5 Transmission Electron Microscopy (TEM)

In order to check the efficacy of the decoating and exosporium removal procedures, normal and treated spores were subject to TEM analysis.

The samples were diluted in sterile distilled water, placed on Formvar-coated grids and examined after negative staining in 2% methylamine tungstate. After 15-30 seconds the excess stain was removed using filter paper. The grids were then washed twice with water and dried by filter paper. Samples were examined immediately using a Phillips EM 208 transmission electron microscope with an accelerating voltage of 80 kilovolts (kV), and magnifications between 20,000x and 55,000x were used.

All TEM materials and chemicals were supplied by Agar Scientific (Essex, UK).

6.2.2 Bioinformatics analysis of spore cortex-lytic enzyme sequences

Using previously published data on the structure and function of the SleC of *C.perfringens S40* as a starting point, we identified homologous regions with the amino acid sequence of CD0551(annotated as *sleC*), a gene within the nucleotide sequence of *C.difficile* 630 which shares 53% amino acid identity (David *et al.*, 2009). The nucleotide sequences of the bases encoding the amino acid regions corresponding to the full-length (C_{50}) and catalytic domains (C_{31}) of *C.perfringens* and *C.difficile* were downloaded from the National Center for Biotechnology Information (NCBI) database and subjected to bioinformatics analysis. Table 6.1 shows the details of these fragments.

	Fragment	NA sequence Region	Gene length (bp)	Reference
***C. perfringens* S40 SleC (# D45024)**	**C_{50}**	852-2168	1317	(Miyata *et al.*, 1995)
	C_{31}	1297-2091	795	(Okamura *et al.* ,2000).
***C.difficile 630* SleC (#AM180355)**	**C_{50}**	656969-658240	1272	(Sebaihia *et al.*, 2006)
	C_{31}	656969-657778	810	This study

Table 6.1. *C. perfringens S40* and *C.difficile 630* C_{50} and C_{31} nucleotide sequences details used in this study (C_{50} = the whole sequence of SleC).

6.2.3 Molecular biological materials and methods

6.2.3.1 Optimisation of recombinant protein expression, analysis and purification

Specific SleC sequences designed during this project were codon optimized to enhance expression from *E.coli* using a commercial algorithm (Genscript Inc). Sequences were then chemically synthesized and cloned into pUC57 by Genscript, after which they were sub-cloned into pQE30 (QIAgen Inc.) to generate N-terminal histidine tagged fusion proteins.

The constructed plasmids were subsequently transformed into chemically competent *E. coli* M15 [pREP4] and SG13009 [pREP4] as described in Chapter 2 (2.2.4). Recombinant proteins were expressed, analyzed, and purified under native conditions, as described in Chapter 2 (2.2.5 and 2.2.6)

6.2.3.2 Germination-specific protease (GSP) extraction and recombinant SleC protein digestion

Germination-specific protease (GSP) was extracted from spores of the following strains: *C. difficile* CD630, R20291, DS1813 and *C. perfringens* R31973, R31782, R31742 using the method described by Urakami *et al*., 1999 and Okamura *et al*. (2000). A spore packed weight of 100-400 mg was resuspended in 4 ml of 0.25 M KCl/50 mM Potassium phosphate pH 7.0 (SIGMA, UK) containing 0.2% CHAPS 3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate (Thermo, UK) at 30°C for 2 h (Urakami *et al*., 1999). The cell debris was then removed by centrifugation (14,000 g for 10 minutes at 4°C, Beckman Coulter Optima L-100K centrifuge, rotor 14) and the supernatant was heated at 52°C in a water bath (Fisher Scientific, Leicestershire, UK) for 20 minutes to inactivate any SCLE that might be present. Following heating, GSP was recovered from the supernatant by centrifugation (14,000 g for 10 min at 4°C). The extract was used as GSP solution without further purification.

To digest the full length lysin (C_{50}), 50 µg/ml of recombinant protein was mixed with 30µl of GSP extract in a 1.5ml microcentrifuge tube and incubated at 32°C for 1 h. To determine whether the recombinant protein had been cleaved, 40 ul of digested protein was mixed with 40 ul of 2% SDS-PAGE sample buffer and denatured using a heating block at 95°C for 5 minutes, after which 10 ul volumes were loaded onto pre-cast gels of 4–20% Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad). Gels were then subjected to electrophoresis for 45 minutes at 150 V (400 mA), as described in Chapter 2 (section 2.2.8).

Following separation, gels were washed twice with SDS and strained with Coomassie stain for 24 h at room temperature followed by de-staining with 25% methanol until all background stain had been removed as described in Chapter 2 (section 2.2.8).

6.2.3.3 Proteins identified by MALDI TOF/TOF mass spectrometry

In order to determine the identity of the protein fragments generated following exposure to GSP, individual bands were extracted from 4–20% Mini-PROTEAN® TGX™ SDS-PAGE gels and subjected to trypsin digestion and matrix-assisted laser desorption-ionization–time-of-flight (MALDI–TOF) mass spectrometry (MS) by the CBS Proteomics Facility, Cardiff University.

Sample preparation

Individual bands were manually excised from the SDS-PAGE gels and the resulting 1.5 mm gel plugs were placed in a 96-well plate prior to trypsin digestion. Sequencing grade modified trypsin (Promega UK Ltd) at a concentration of 6.25 ng/ul in 25mM (NH_4HCO_3) was added and the samples were incubated for 3 h at 37ºC. The resulting dried peptides were resuspended in 50% (v/v) acetonitrile in 5ul 0.1% (v/v) trifluoroacetic acid (TFA) for analysis using MS. A 0.5ul aliquot corresponding to 10% of the material was spotted onto a 384-well MS plate, allowed to dry and then overlaid with ά-cyano-4-hydroxycinnamic acid (CHCA, Sigma, UK; 0.5 ul prepared by mixing a 5 mg matrix with 1 ml of 50% (v/v) acetonitrile in 0.1% (v/v).

MALDI TOF/TOF mass spectrometer analysis

Mass spectrometry was performed using a MALDI TOF/TOF mass spectrometer (Applied Biosystems 4800 MALDI TOF/TOF Analyzer; Foster City, CA, USA) with a 200 Hz solid state laser operating at a wavelength of 355nm (Medzihradszky *et al.*, 2000; Bienvenut *et al.*, 2002; Glückmann *et al.*, 2007; Brennan *et al.*, 2009). MALDI mass spectra and subsequent MS/MS spectra of the eight most abundant MALDI peaks were obtained following routine calibration. Common trypsin autolysis peaks and matrix ion signals and precursors within 300 resolution of each other were excluded from the selection, and the peaks were analysed starting with the strongest peak. A total of 800 laser shots were averaged (mass range 700–4000 Da; focus mass 2000) to produce positive-ion reflector mode spectra. In MS/MS positive ion mode, 4000 spectra were averaged based on a 1 kV collision energy (collision gas was air at a pressure of 1.6×10^{-6} Torr) and default calibration.

Combined PMF and MS/MS queries were performed using the MASCOT Database search engine v2.1 (Matrix Science Ltd, London, UK) (Perkins *et al.*, 1999), embedded into Global Proteome Server (GPS) Explorer software v3.6 (Applied Biosystems) on the Swiss-Prot database (download date 13/07/2012), and TrEMBL (download date 22/07/2011). Searches were restricted to bacteria domain taxonomy with trypsin specificity (one missed cleavage allowed); the tolerances for peptide identification searches were set at 50 ppm for MS and 0.3 Da for MS/MS. Modification of cysteine by iodoacetamide was employed as a fixed modification, with methionine oxidation as a variable modification. Search results were evaluated by manual inspection and conclusive

identification confirmed whether high quality tandem MS (good y-ion) data had been produced for ≥2 peptides (E value $p < 0.05$ for each peptide; overall $p < 0.0025$) or one peptide (only if E value was $p < 0.0001$).

6.2.4 Assay of SCLE activity

The lytic activity of recombinant cortex-lytic enzyme fragments was determined by measuring the decrease in optical density of coated and decoated spore suspensions following the addition of lysin using the methods described by Miyata *et al*. (1995a) and Okamura *et al*. (2000). with the following modification : spores, resuspended in 200ul of 40 mM potassium phosphate (pH 7) to an OD_{600} nm of 0.8–1.5, were mixed with 100 µg/ml of recombinant protein in the wells of a Bioscreen C System multiwall plate, and were incubated at 32°C with shaking (every 2 min). Changes in optical density (OD_{600}) were measured every 2 min over a 3 h period. Chicken egg white lysozyme (SIGMA, UK) at 100µg/ml was used as a positive control and spores with no enzyme were used as a negative control. Lytic assays were repeated three times, every time a new spore preparation and recombinant protein extraction was used

6.3 Results

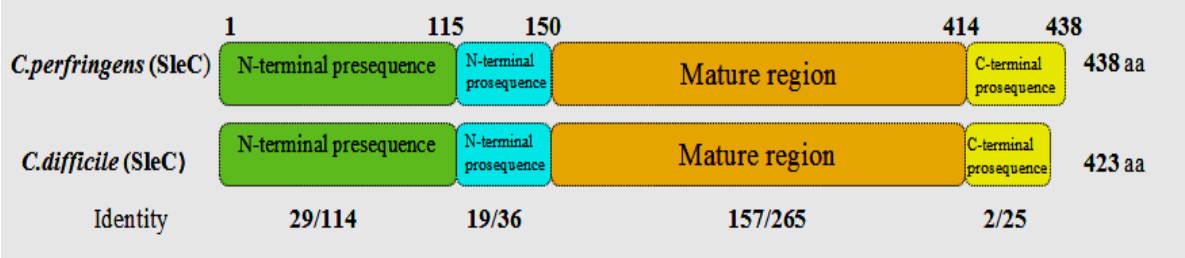
6.3.1 Bioinformatics results

A gene (1272 bp) encoding a Spore cortex-lytic enzyme pre-pro-form was identified in the *C. difficile* 630 genome sequence through sequence similarity analysis with *C. perfringens* S40. The protein sequence of this gene (#CAJ67384) was found to share 53% amino acid identity with the SleC gene (#BAA08081) of *C.perfringens S40* (David *et al.*, 2009). Multiple sequence alignment of these sequences using the ClustalW2 analysis programme revealed the presence of amino acids that were highly conserved, particularly in the region that encoded the enzymic center (mature region), as can be seen in figure 6.3.

Interestingly, our BLASTn searches revealed that the nucleotide sequence of this gene appears to be highly conserved in that a homology (98% homology or higher) has been found in all nine strains of *C.difficile* sequenced to date (Sebaihia *et al*., 2006; Stabler *et al.*, 2009; He *et al*., 2010).

	Fragment	Gene length (bp)	Protein length (aa)	MW(kDa)	Protein Solubility prediction
***C. perfringens* S40 SleC**	**C_{50}**	**1317**	**438**	**49.65**	**77.6%**
	C_{31}	**795**	**265**	**30.34**	**87.3%**
***C.difficile* 630 SleC**	**C_{50}**	**1272**	**423**	**47.34**	**74%**
	C_{31}	**810**	**270**	**30.45**	**66.9%**

Table 6.2. Showing the details of spore cortex-lytic enzyme sequences used in this study, and the predicted MW, which was generated using ExPASy tools (http://web.expasy.org/compute_pi/) and protein solubility, estimated according to Wilkinson-Harrison (Wilkinson and Harrison, 1991; Davis *et al.*, 1999; Harrison, 2000) using their web-server (http://www.biotech.ou.edu/).




```
C.perfringens (SleC)      MAVGGLKVQCFSEQRRYIPIDKCKVKITPTGEDGVAVGDTIE-LYTDDTGSTDTIELDAP 59
C.difficile (SleC)        MQDGFLTVSIIDATN-NRPIQNAVVNIYSMSNGSQSSSTLYQNLRSNESGQVTGLVLPAP 59
                          *  * *.*. :.  .   **::. *:* . .:.. : .   : * ::::*..  : * **

C.perfringens (SleC)      PIENSNQPG-TIPYSFAEVIVEREGFLPVAVNGVQIYPSRIALQNVNLP----ETRGYYR 114
C.difficile (SleC)        DVDYSLQPSDVRPYSQYIVEAIADGYETVVIEGTQLLATIEARQGVPMSPRTRSKRSFSR 119
                           :: * **. . ***   * .  :*: .*.::*.*:. .:  * *.* :.    ..*.: *

C.perfringens (SleC)      Q-EEVIDIQPNRLVGNFPPKIPEAEEKELPPPKGTVVLPEPVVPEYIVVHNGRPNDNSVA 173
C.difficile (SleC)        QSELIFDIGEHTLYGTYPPKIPESNLKPLPPPTGFVVLDNPVVPEFIVVHDGLPEDSSAP 179
                          * * ::**  : * *.:******:: * ****.* *** :*****:****:* *:*.*..

C.perfringens (SleC)      NYKVNYKDYIKNVACCEIFSTWSENTIRANVYAIISFTLNRIYTEWYRGKGKNFDITNST 233
C.difficile (SleC)        NYWIPFKEYIKNIASSEIYSTWPEQTIYANVIAIISFTLNRVFTEWYRNKGYNFTITSTT 239
                          ** : :*:****:*..**:***.*:** *** *********::*****.** ** **.:*

C.perfringens (SleC)      AFDHAFSYGRNFYDNISRIVDEIFSTYMKR-FNSKQPLLAQYCDGINVQCPGWMTQWGSK 292
C.difficile (SleC)        AYDHKFINNRNLFEPINVVVDAIFNTFIKRPPTSRQPLLAQYCDGQKSQCPDQMTQWGSK 299
                          *:** *  .**::: *. :** **.*::**  .*:********** : ***. *******

C.perfringens (SleC)      YLGDEGKVPYDILTSFYGDDLELKSAKKVKGSPRSYPGYTLKTGYSGEPVRVIQEQLNAI 352
C.difficile (SleC)        DLGDQGYDYESILRYFYGDEIVFERAPIVSGVPVSFPGTTLQVGSSGQYVRTIQNQLNAI 359
                           ***:*    .**  ****:: :: *  *.* * *:** **:.* **: **.**:*****

C.perfringens (SleC)      SRAYPLIPKIAVDGKYGPKTREAVKTFQKIFNLPQTGEVDYATWYKISDVYVAVTKIAEL 412
C.difficile (SleC)        SNSYPAVPKVIEDGIYGTDTENAVKIFQGIFGLPQSGVVDFKTWYEISRVYVATTRIASL 419
                          *.:** :**:  ** **..*.:*** ** **.***:* **: ***:** ****.*:**.*

C.perfringens (SleC)      RSSVEKKIFYPPTIMDRRENVPKIIY 438
C.difficile (SleC)        N----------PLI------------ 423
                          .          * *
```

Figure 6.3. A) Schematic representation of the different regions of SleC in the *C.difficile 630* gene homolog using the SleC of *C.perfringens S40* as a guide (Okamura *et al.*, (2000). Homologous regions are highlighted by different colors are shown with number of identical amino acid. Overall, the amino acid sequences of *C.perfringens* SleC and *C.difficile* SleC are 53 % identical. **B)** Alignment of the amino acid sequence of *C. perfringens* SleC aligned *C.difficile* SleC identified by BLAST analysis. Alignment was produced with ClustalW2; identical amino acids are marked with a star. Positions of the GPS cleavage sites are indicated by arrows.

The structural similarity of the SleC produced by *C.perfringens* and *C.difficile* strain 630 suggests that the SleC produced by *C.difficile* may also function as a spore cortex-lytic enzyme.

6.3.2 Cloning and expression of spore cortex-lytic enzymes

6.3.2.1 DNA Manipulation

Individual protein sequences were codon optimized, synthesized and cloned into pUC57 and then sub-cloned into the QIAgen expression vector pQE30. Plasmid was transformed into competent cells of *E.coli* M15[pREP4] and SG13009[pREP4] for the protein expression. Positive transformants were identified by culture on LB agar plates containing ampicillin (100 µg/ml) and kanamycin (25 µg/ml).

The transformation efficiency of pQE30 plasmid containing the cortex-lytic enzyme gene sequences (*C.perfringens* C_{50} ,C_{31} and *C.difficile* C_{50}, C_{31}) for competent cells of *E. coli* M15 and SG13009 were determined and were found to be lower than that seen for the positive control (pQE-40). M15 seems to provide a statistically significantly better transformation ratio than SG13009; therefore, *E. coli* M15 was select for further work.

Construct	Transformants /µg of plasmid DNA	
	M15	SG13009
***C. perfringens* C_{50}**	$152.7 \pm 33X10^3$	$180 \pm 10.2X10^3$
***C. perfringens* C_{31}**	$202 \pm 2.6X10^3$	$143.2 \pm 1.7X10^3$
***C. difficile* C_{50}**	$355.5 \pm 7.2X10^3$	$353.8 \pm 6.2X10^3$
***C. difficile* C_{31}**	$264.5 \pm 5X10^3$	$146.4 \pm 4X10^3$
pQE40	$435.8 \pm 440X10^3$	$294.6 \pm 400X10^3$

Table 6.3. Transformation efficacy in *E. coli* M15 and SG13009 , using 100 µl of the overnight culture plated out on LB-agar plates for 24h (readings represent the mean of three results± sd). pQE-40 control plasmid which expresses 6xHis-tagged DHFR is optimized for high-level transformation into these *E.coli* strains.

6.3.2.2 Optimization of recombinant protein expression, analysis and purification

Fragments from *C. difficile* (C_{50} and C_{31}) and *C. perfringens* (C_{50} and C_{31}) were subcloned into the BamHI and Sall restriction sites of the expression vector pQE-30 in *E. coli* M15 (pREP4). The expression of individual histidine tagged recombinant proteins was induced when the culture reached an OD_{600} 0.6 by the addition of 1mM IPTG as described in Chapter 2 (section 2.2.5). The induced culture was incubated for a further 4 h at 27^0C, which represents the optimum induction conditions for each recombinant protein as determined by experimentation (Figures 6.4, 6.5, 6.6 and 6.7).

Recombinant proteins were purified using affinity chromatography under native conditions. Protein bands were analyzed by SDS-PAGE using pre-cast gels 4–20% Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad). Bands with molecular weights of approximately 47 and 30 kDa (the predicted sizes of *C.difficile* C_{50} and C_{31}) and 49 and 30 kDa (the predicted sizes of *C.perfringens* C_{50} and C_{31}) were visualized in SDS-PAGE gel stained with Coomassie stain following IPTG induction (Figures 6.4, 6.5, 6.6 and 6.7). The expression and purification of each recombinant protein fragment is summarized in the following series of SDS-PAGE Coomassie stained gels (Figures 6.4, 6.5, 6.6 and 6.7).

Figure 6.4 showed the expression and purification of the *C.difficile* C_{50} recombinant protein. Time-course analysis of protein expression revealed that prior to, and 1 h post, IPTG induction there was no obvious expression of C_{50}. In contrast, 4 h post induction, a

thick band of approximately 47 kDa in size can be observed, indicating that C_{50} protein was successfully expressed.

Three rounds of protein binding and elution were sufficient to recover ~ 1-2mg of recombinant protein from 100 ml of starting culture, as determined using the BCA protein assay (section 2.2.7). It was noted that the recombinant proteins precipitated in the recommended elution buffer and therefore the purified His-tagged protein were dialyzed against 20 mM sodium phosphate buffer (NaPhos) (pH6) and stored at 4^0C before further use.

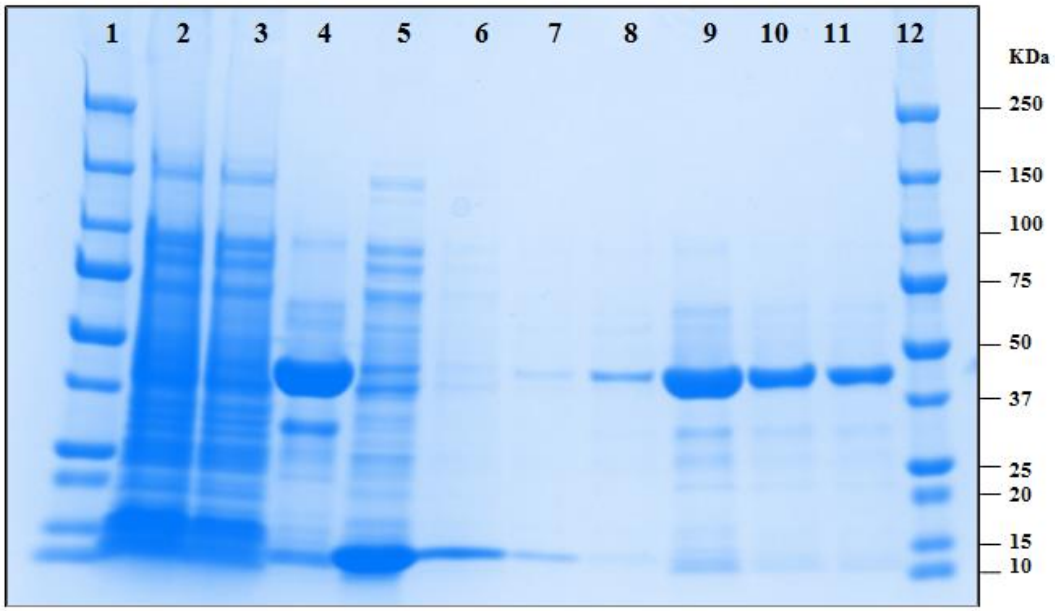


Figure 6.4. Expression and purification of *C.difficile* C_{50} protein (expected size 47.34 kDa): Lanes: 1, Molecular marker ; 2, non-induced culture ; 3, 1h induced culture; 4, total protein extracts after 4 h of induction with IPTG; 5, column flowthrough; 6, primary wash effluent; 7, secondary wash effluent; 8, 1st eluate (E1); and 9, 2nd eluate (E2); 10, 3rd eluate (E3) ;11 , buffer exchange 20mM NaPhos pH 6; 12, Molecular marker

In figure 6.5, the expression and purification of the *C.difficile* C_{31} recombinant protein is summerised. The culture and induction conditions were identical to those used to express C_{50}.

After 4 h, a thick band of approximately 30 kDa was observed (lane 4), suggesting that C_{31} protein was successfully expressed. Three rounds of protein binding and elution were sufficient to recover ~ 0.8-1.5mg (BCA assay) of recombinant protein from 100ml of starting culture. The purified protein was finally resuspended in 20mM sodium phosphate buffer pH 6 with minimal loss of protein concentration.

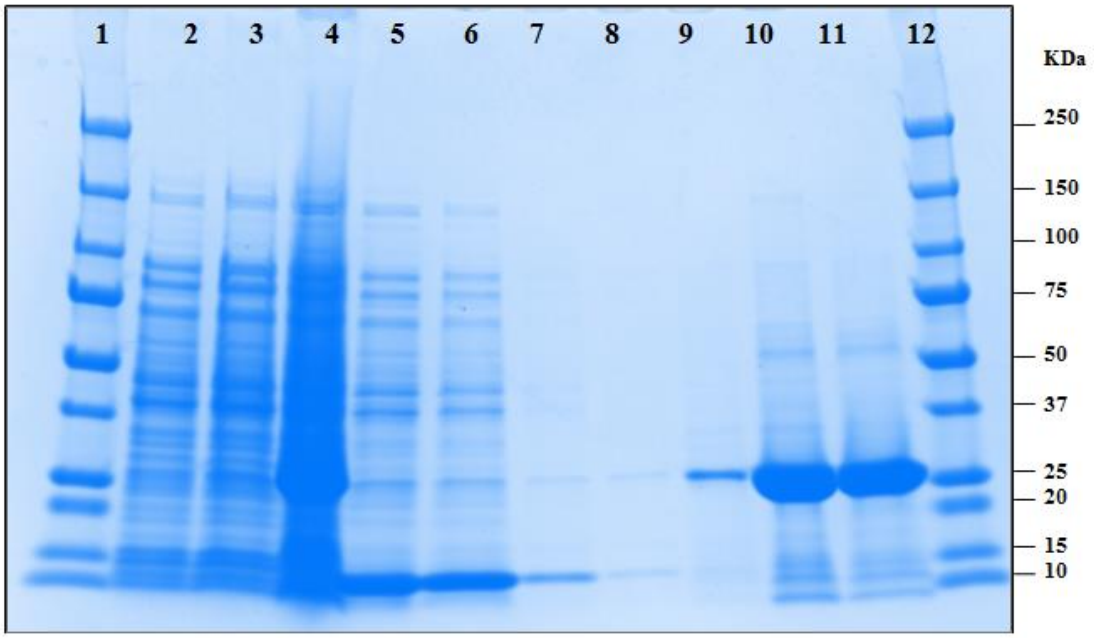


Figure 6.5. Expression and purification of *C.difficile* C_{31}(expected size 30.45 kDa), Lanes: 1, marker; 2, non-induced culture; 3, 1h induced culture; 4, total protein extracts after 4 h of induction with IPTG; 5, column flowthrough; 6, primary wash effluent; 7, secondary wash effluent; 8, 1st eluate (E1); and 9, 2nd eluate (E2), 10, 3rd eluate (E3) , 11 , buffer exchange 20mM NaPhos pH 6; 12, marker

Figures 6.6 and 6.7 show the expression and purification of recombinant *C.perfringens* C_{50} and C_{31}. The optimal induction and expression conditions were the same as those for the *C.difficile* proteins, with maximal expression occurring after 4 h of induction, at which time thick bands of approximately 49 and 30 kDa in size were observed, suggesting that both proteins were successfully expressed.

Three rounds of protein binding and elution were sufficient to recover ~ 0.5 -1mg of C_{50}, and 0.5–1mg of C_{31} from 100 ml of culture (BCA assay, section 2.2.7) . As was the case with the other proteins, *C.perfringens* C_{50} and C_{31} precipitated in the elution buffer and therefore the purified His-tagged protein were dialyzed against 20 mM Sodium phosphate buffer (pH6) and stored at 4^0C before further use.

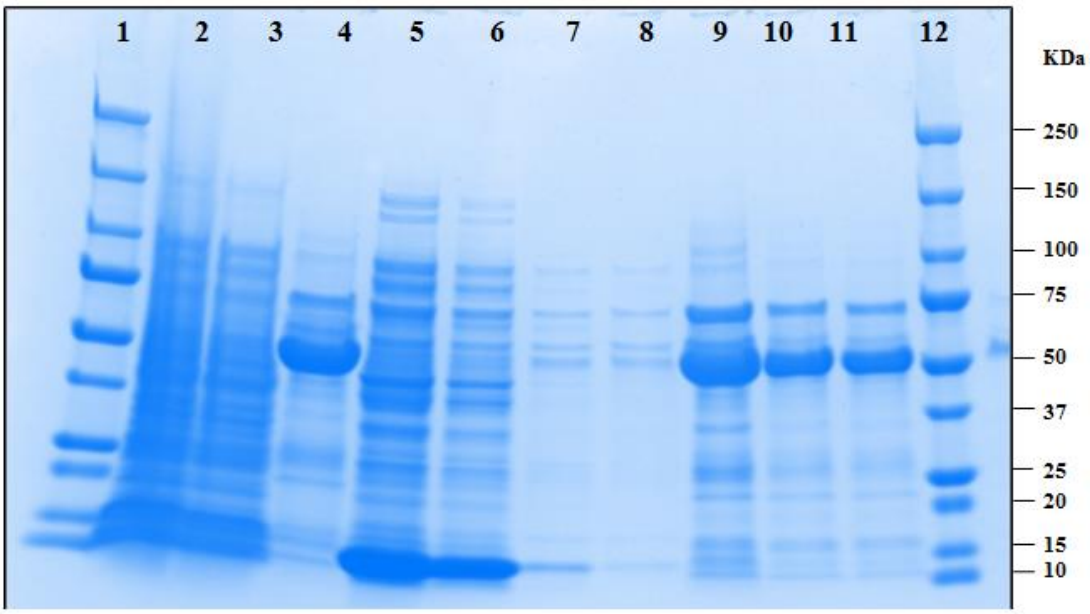


Figure 6.6 Expression and purification of *C.perfringens* C_{50} (expected size 49.65 kDa) : Lanes: 1, marker; 2, non-induced culture; 3, 1h induced culture; 4, total protein extracts after 4 h of induction with IPTG; 5, column flowthrough; 6, primary wash effluent; 7, secondary wash effluent; 8, 1st eluate (E1); 9, 2nd eluate (E2); 10, 3rd eluate (E3); 11, buffer exchange 20mM NaPhos pH 6; 12, marker

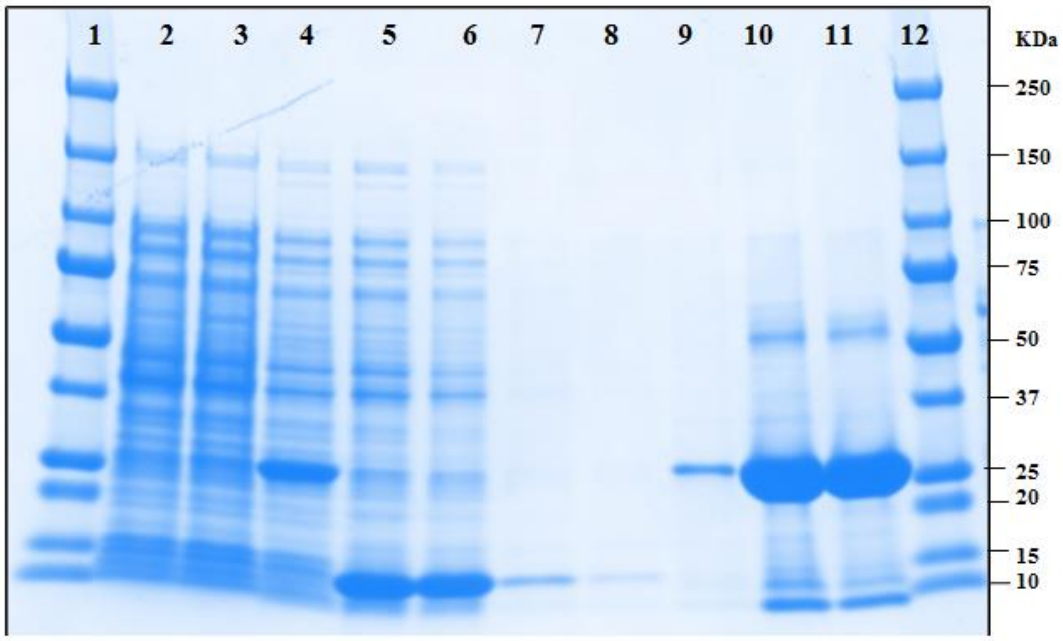


Figure 6.7. Expression and purification of *C.perfringens* C_{31} (expected size 30.34 kDa): Lanes: 1, marker; 2, non-induced culture; 3, 1h induced culture; 4, total protein extracts after 4 h of induction with IPTG; 5, column flowthrough; 6, primary wash effluent; 7, secondary wash effluent; 8, 1st eluate (E1); 9, 2nd eluate (E2); 10, 3rd eluate (E3); 11 , buffer exchange 20mM NaPhos pH 6; 12, marker

6.3.2.3 Protein detection: Western blot

The identity of the recombinant spore lytic proteins were confirmed by probing western blots with an antibody, Penta-His, which recognized the N terminal histidine tag (as described in Chapter 2 section 2.2.9). As can be seen in figure 6.8, signals were observed which corresponded in size to the expected molecular weights of 47, 30, 49 and 30 kDa. A protein marker (6x His-protein ladder) confirmed the size of the signals.

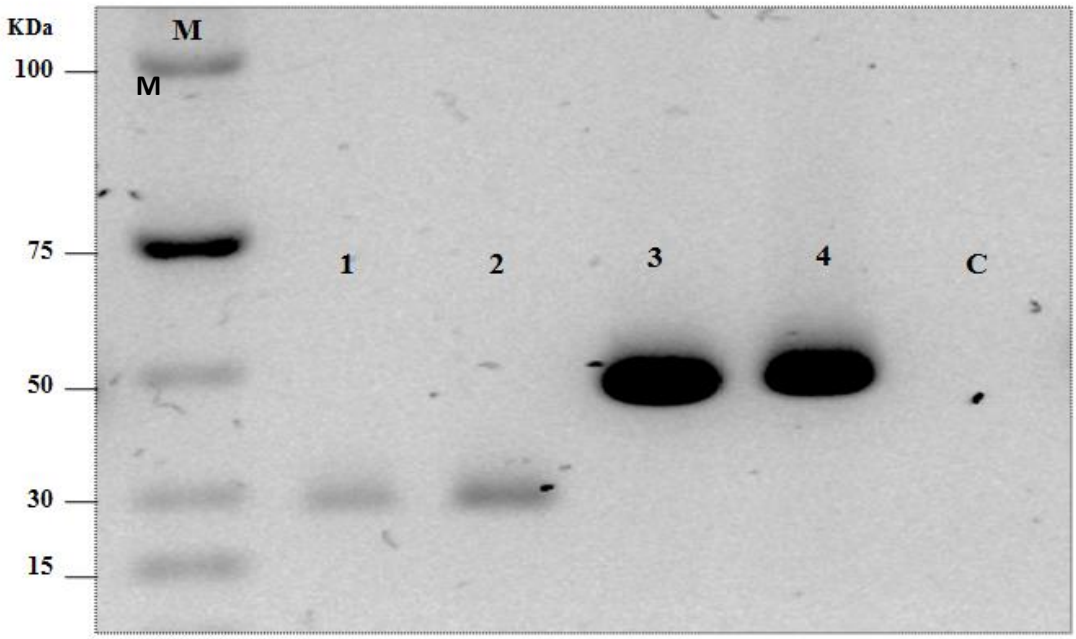


Figure 6.8. Western blot analysis of purified spore cortex-lytic enzymes, (1) *C.difficile* C_{31} (30 kDa) (2) *C.perfringens* C_{31} (30 kDa) (3) *C.difficile* C_{50} (47 kDa) and (4) *C.perfringens* C_{50} (49 kDa), Western blot technique. M: protein marker (6x His-protein ladder 100-15 kDa); C: negative control (buffer only).

6.3.4 GSP activation of recombinant lysin proteins

The GSP extracted from *C.perfringens* S40 spores cleaves full-length SleC (C_{50}) to generate a protein band 34kDa (Okamura *et al.* 2000), which was determined to have a lytic activity. We determined whether GSPs isolated from spores of a range of *C.perfringens* (R31973, R31782, R31742) and *C.difficile* (CD630, R20291, DS1813) isolates behaved in a similar manner and were able to cleave the recombinant C_{50} protein derived from *C. perfringens* and *C. difficile*.

As can be seen from the results presented in figure 6.9, only one extract obtained from spores produced by *C.difficile* DS1813 was able to digest the C_{50} of *C. perfringens* to

produce bands of ~ 34 kDa and ~ 16 kDa. The failure of the other spore extracts to cleave the recombinant protein is intriguing, although it could reflect problems with the spore extraction method.

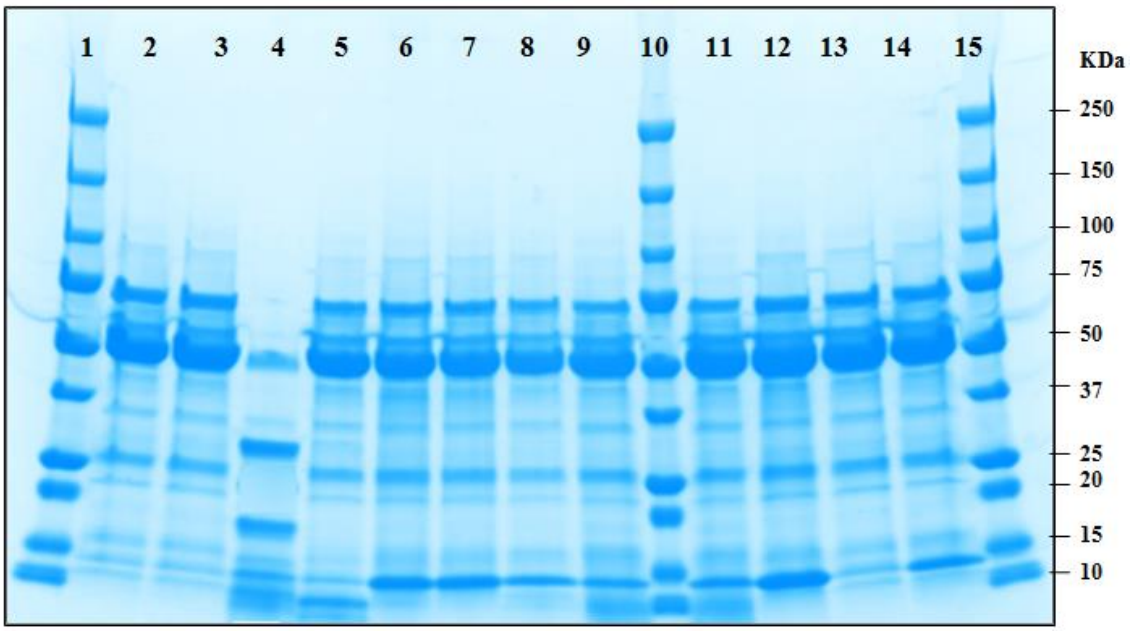


Figure 6.9. Treatment of recombinant full length SleC (C_{50}) from *C.perfringens* with Germination specific proteases (GSP) isolated from *C.perfringens* and *C.difficile* spores. 1: Molecular weight marker, 2: Nondigested protein, 3: protein with buffer; 4: protein with GSP from DS1813; 5: protein with GSP from 630, 6: protein with GSP from DS20291; 7: protein with GSP from R31742; 8: protein with GSP from R31972; 9: protein with GSP from R3174510: Molecular weight marker; 11: protein with GSP from DS20291; 11: protein with GSP from 630; 12: protein with GSP from R31782; 13: protein with GSP from R31742; 14: protein with GSP from R31972; 15: Molecular weight marker.

Interestingly, we saw a similar pattern of activity when the various spore extracts were mixed with the *C.difficile* recombinant C_{50} in that only the extract from *C.difficile* DS1813 appeared to cleave the recombinant protein (Figure 6.10).

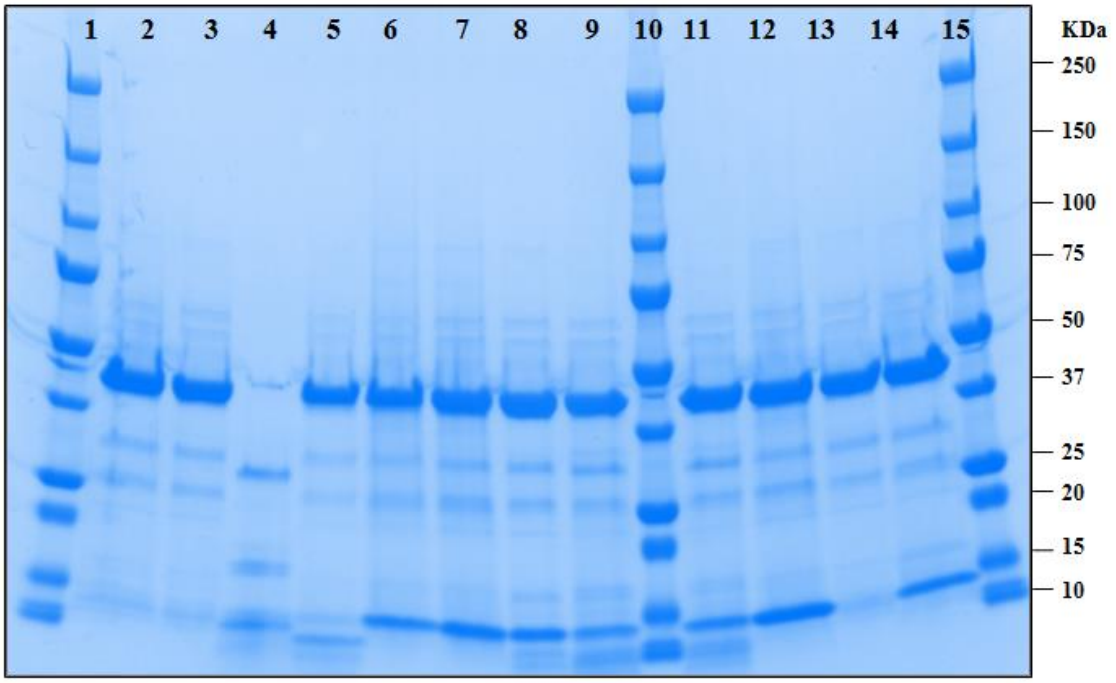


Figure 6.10. Treatment of recombinant full length SleC (C_{50}) from *C.difficile* with Germination specific proteases (GSP) isolated from *C.perfringens* and *C.difficile* spores. 1: Molecular weight marker; 2: Nondigested protein; 3: protein with buffer;, 4: protein with GSP from DS1813;, 5: protein with GSP from 630; 6: protein with GSP from DS20291; 7: protein with GSP from R31742; 8: protein with GSP from R31972; 9: protein with GSP from R31745; 10: Molecular weight marker; 11: protein with GSP from DS20291;, 11: protein with GSP from 630; 12: protein with GSP from R31782; 13: protein with GSP from R31742; 14: protein with GSP from R31972; 15: Molecular weight marker (250-10 kDa)

6.3.5 Identification of SleC digestion products using MALDI TOF/TOF mass spectrometry

Individual SleC digestion products resolved on SDS-PAGE gels (Figures 6.11 A and B) were excised, trypsin-digested and analyzed by MALDI TOF/TOF MS, with the resulting MS data being queried using Mascot. The sequences and expected values (Exp) for each peptide are displayed in table 6.4 and appendix 3.

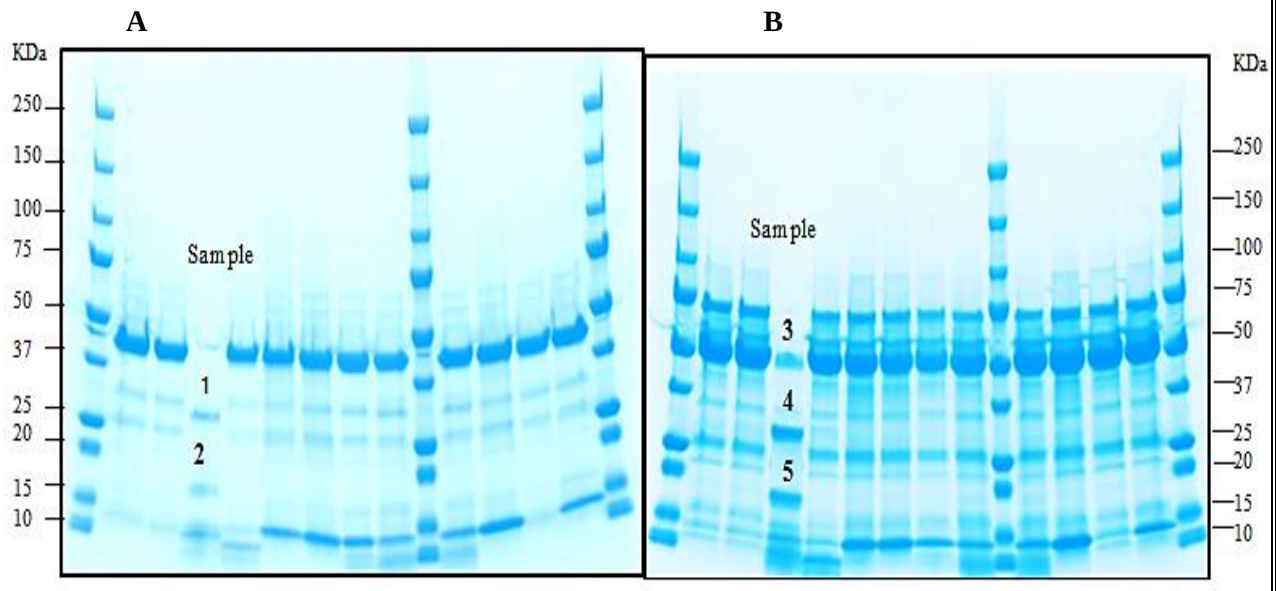


Figure 6.11. Showing the site of samples generated from treatment of recombinant full length SleC (C_{50}) from *C.difficile* (A) and *C.perfringens* (B) with GPS isolated from *C. difficile* and *C.perfringens* strains, as described above in figures 6.9 and 6.10.

Sample	ID category	Name and Accession Number	Mascot score	Sequence coverage (%)	Total number of peptides (MS and MS/MS)	Total number of peptides with MS/MS data	Peptide sequences	e value
1	1	A	291	28	9	3	YFYGDEIVFER	5.50E-07
							APIVSGVPVSFPGTTLQVGSSGQYVR	1.70E-05
							VFTEWYR	0.079
2	1	A	471	31	10	3	DLGDQGYDYESILR	5.60E-09
							YFYGDEIVFER	8.60E-07
							APIVSGVPVSFPGTTLQVGSSGQYVR	1.70E-06
3	1	B	322	31	17	2	VIQEQLNAISR	0.0012
							IALQNVNLPETR	0.0051
4	2	B	194	25	6	2	VIQEQLNAISR	1.10E-05
							IYTEWYR	0.0053
5	1	B	256	23	12	2	IALQNVNLPETR	1.20E-05
							EGFLPVAVNGVQIYPSR	0.00083
							VQCFSEQR	0.078

Table 6.4. MALDI-TOF/TOF mass spectrum of an in-gel tryptic digest of the spore cortex-lytic enzyme band cut from the 1-D SDS–PAGE gel in Figure 6. 11

Spore cortex-lytic enzyme pre-pro-form OS=*Clostridium difficile* (strain 630), accession No ACQ188Z5_CLOD6
Spore cortex-lytic enzyme SleC OS=*Clostridium perfringens* CPE str, Accession No B1RFT4_CLOPE

Mass spectrophotometer analysis of recombinant protein digestion fragment bands revealed the presence of amino acid sequences which corresponded to the published amino acid sequences of the SleC of *C.difficile* and *C.perfringens* (Table 6.4 and appendix 3).

6.3.6 Production of exosporium deficient and decoated spores

To determine the activity of recombinant lytic enzymes against spore peptidoglycan, we removed the outer layers of the spore to enhance the ability of the lysins to access the cortex. To achieve this aim, *C.perfringens* spores were decoated while *C.difficile* spores were subjected to decoating and sonication. The efficiency of each process was determined using TEM. Electron microscopy of negatively stained *C.difficile* DS1813 spores revealed the presence of a core, cortex, coat and exosporium (Figure 6.12).

To determine whether the recombinant lysins were able to lyse spores lacking the exosporium, spores of *C.difficile* DS1813 were subjected to sonication. As can be seen from the images of negatively stained spores presented in figure 6.13 (A and B), a sonication regimen of 50 seconds bursts over 16 cycles (amplitude15 μm) with each burst separated by 98 seconds of cooling on ice successfully remove the exosporium without visually damaging the underlying spore structure. In contrast, images of decoated spores of *C. difficile* DS1813 confirmed the absence of the coat and exosporium (Figure 6.13 C and D).

Coated and decoated spores of *C. perfringens* R31782 were also negatively stained and examined using TEM. As can be seen from figure 6.14, untreated spores appeared as expected with a core, cortex and coat. In contrast, the decoated spores lacked the coat.

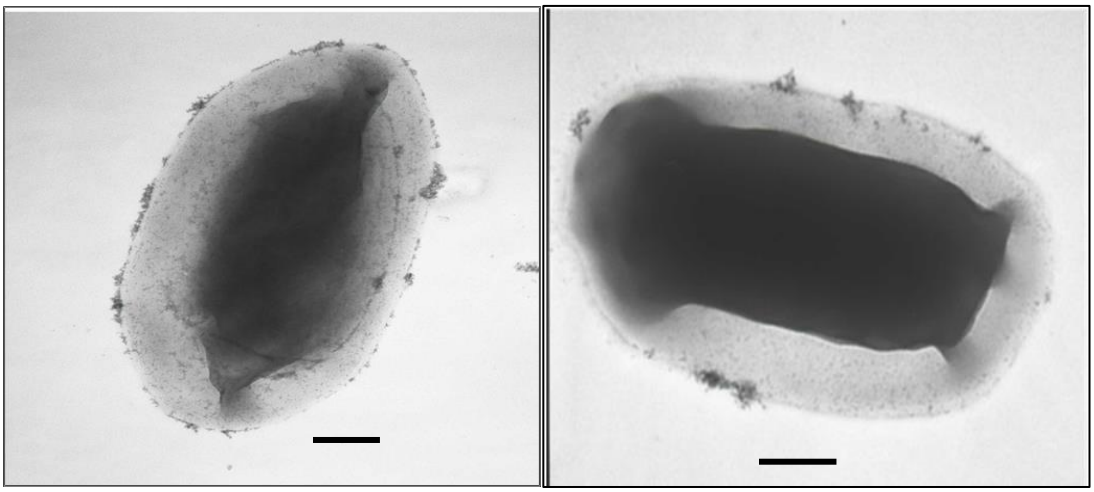
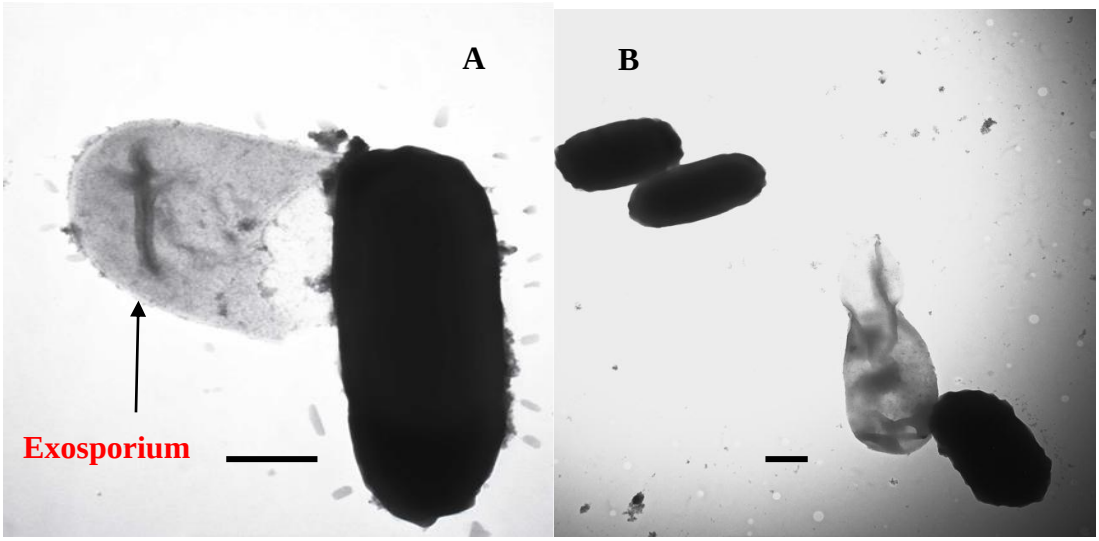


Figure 6.12. TEM of untreated *C. difficile* DS1813, stained negatively with 2% methylamine tungstate. (Scale bar: 100nm), Representation from 20 field of view



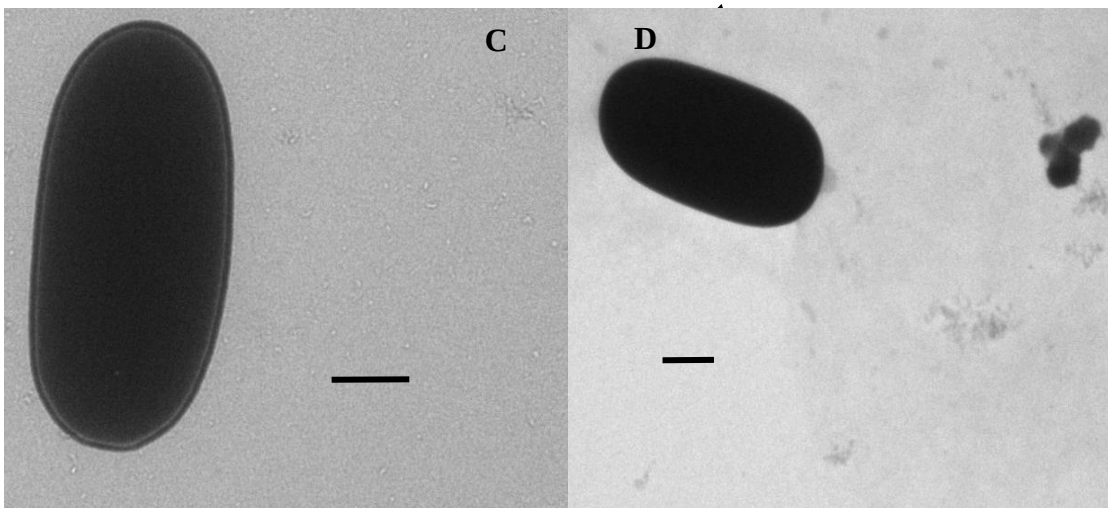


Figure 6.13. A , B)TEM of sonicated *C.difficile* DS1813 spores , stained negatively with 2% methylamine tungstate

C, D)TEM of decoated *C. difficile* DS1813 spore, stained negatively with 2% methylamine tungstate (Scale bar :300nm) ; Representation from 20 field of view

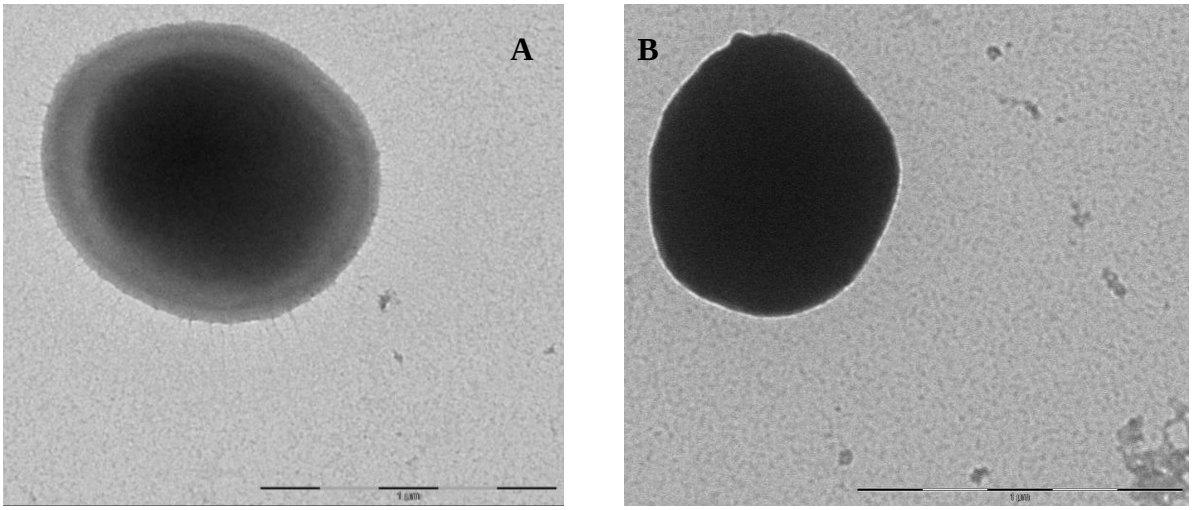


Figure 6.14. TEM of a normal coated (A) and decoated (B) *C. perfringens* R31782 spore, stained negatively with 2% methylamine tungstate
(bar :1µm) * Representation from 20 field of view.

6.3.7 Lytic activity of spore cortex-lytic enzymes

In this assay, coated and decoated spores of *C.perfringens* R31782, and for *C. difficile,* DS1813, coated, decoated and sonicated spores were used. Samples were subjected to digestion by the following lysins: full length SleC (C_{50}), the active section of the enzyme (C_{31}) and the product of the C_{50} digested by GPS with 34 kDa, which will be referred to as (C_{34}).

Lysozyme at 100µg/ml was employed as a positive control. Untreated coated and decoated spores of the corresponding strain were employed as negative controls.

Decoated spores of *C.perfringens* were considerably more sensitive to lysozyme and C_{31} than were their intact counterparts, suggesting that the outer layers of the spore act as a permeability barrier (Figures 6.15 and 6.16).

The full length recombinant *C.perfringens* SleC (C_{50}) showed no lytic activity against decoated spores of *C.perfringens*, whereas C_{31} and C_{34} produced significant lytic activity ($P < 0.05$ for both) after 3 h of incubation, the recombinant C_{31} being the most active (Figure 6.16).

It is important to mention that we looked at the activity of *C.difficile* DS1813 spore extract containing GPS alone against coated and decoated spores of *C.perfringens* and observed no evidence of lysis in this assay.

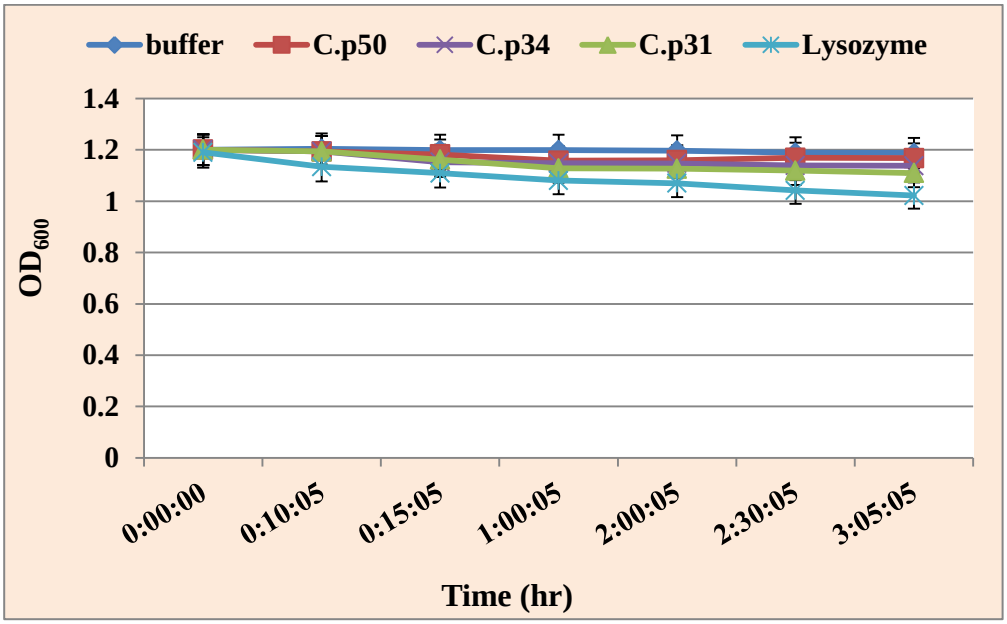


Figure 6.15. Lytic activity of the C_{50},C_{34},C_{31} of *C.perfringens* proteins and lysozyme against normal coated spores of *C.perfingens* R31782. Values are the means of results from triplicate assays ± standard deviations.

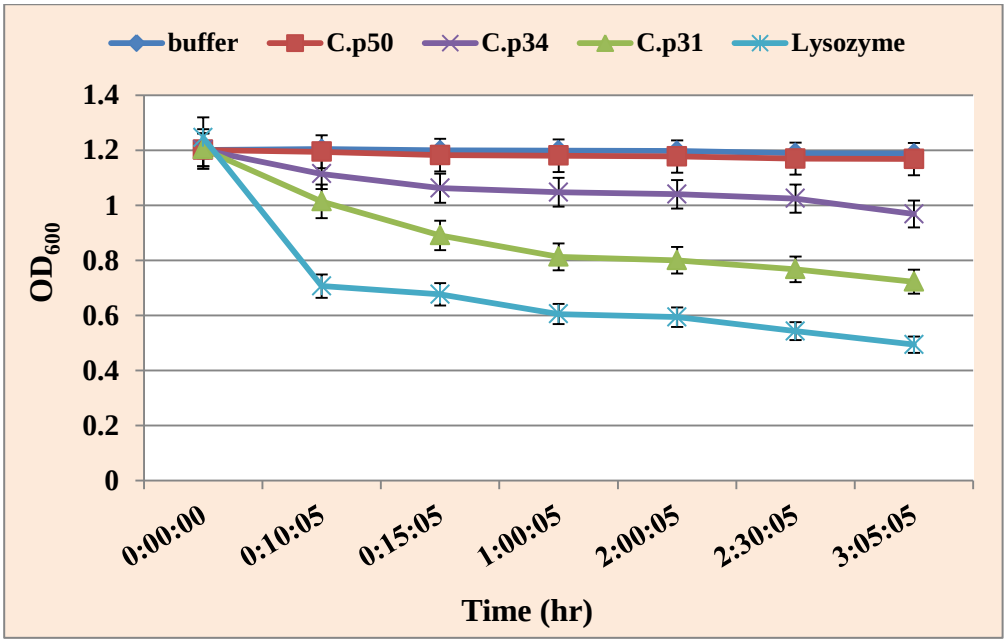


Figure 6.16 Lytic activity of the C_{50},C_{34},C_{31} of *C.perfringens* proteins and lysozyme against decoated spores of *C.perfingens* R31782 . Values are the means of results from triplicate assays ± standard deviations.

We subsequently determined the activity of C_{50},C_{34} and C_{31} from *C.perfringens* against *C.difficile* spores (coated, decoated and exosporium deficient) and observed no evidence of lysis. This result is important, as it suggests that while the SleC of *C.perfringens* and *C.difficile* are similar, either they may have different spore peptidoglycans or the two lysins may recognise different sites. This was also the case for the *C.difficile* lysins against *C.perfringens* spores (coated and decoated).

We next determined the ability of full length *C.difficile* SleC (C_{50}), its GPS generated digestion product, C_{34} and recombinantly expressed C_{31} to lyse intact spores of *C.difficile* (Figure 6.17). As can be seen, little activity was detected, suggesting that as was the case for *C.perfringens*, the spore coat of *C.dfficile* acts as a permeability barrier.

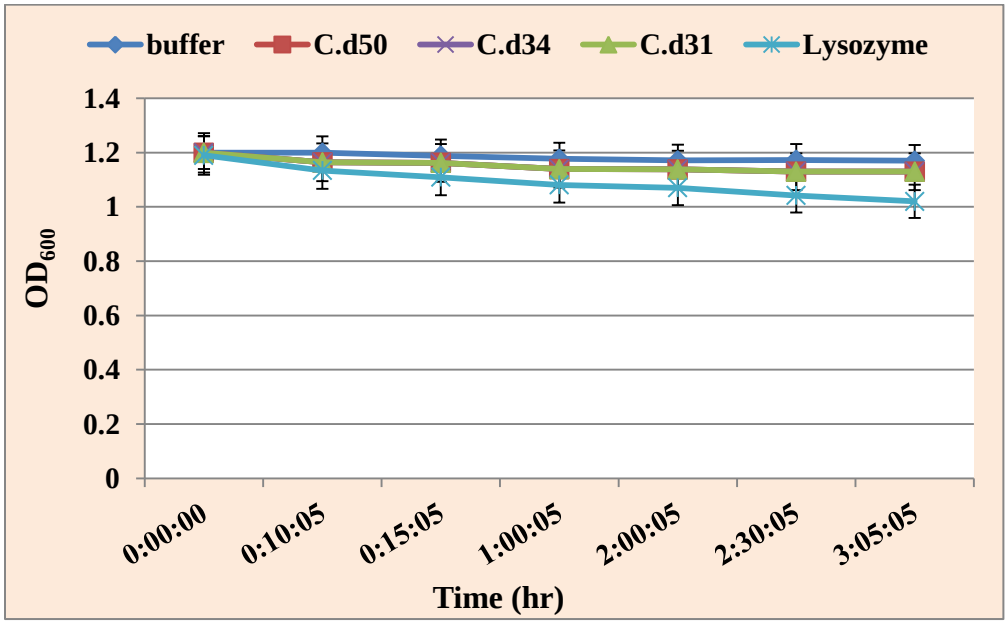


Figure 6.17 . Lytic activity of the C_{50},C_{34},C_{31} of *C.difficile* proteins and lysozyme against normal coated spores of *C.difficile* DS1813 . Values are the means of results from triplicate assays ± standard deviations

A similar lack of activity was observed for spores of *C.difficile* in which the exosporium had been removed, suggesting that the outer layer of the spore does not represent a significant permeability barrier for spore lysins (Figure 6.18).

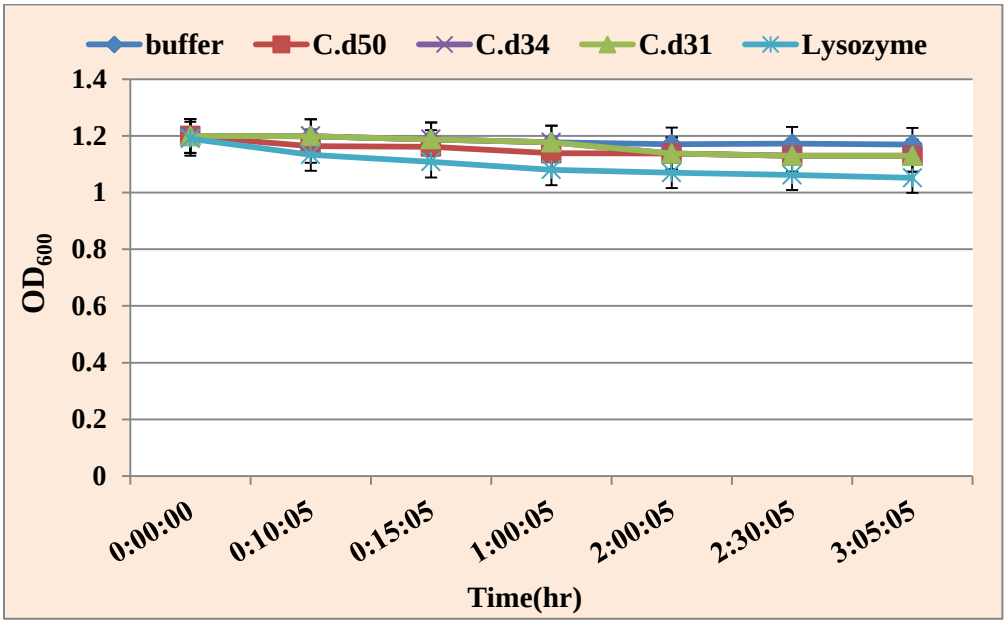


Figure 6.18. Lytic activity of the C_{50},C_{34},C_{31} of *C.difficile* proteins and lysozyme against exosporium deficient spores of *C.difficile* DS1813. Values are the means of results from triplicate assays ± standard deviations

As was the case for *C.perfringens* spores, it required the removal of the spore coat to achieve lysine-mediated spore degradation. As can be seen from figure 6.19, in *C.difficile* C_{31} was considerably more active than C_{50} and C_{34}.

Despite the fact that *C.difficile* C_{31} was inactive against coated spores, with decoated spores it showed significant activity ($P < 0.05$) in comparison to pro form C_{50}

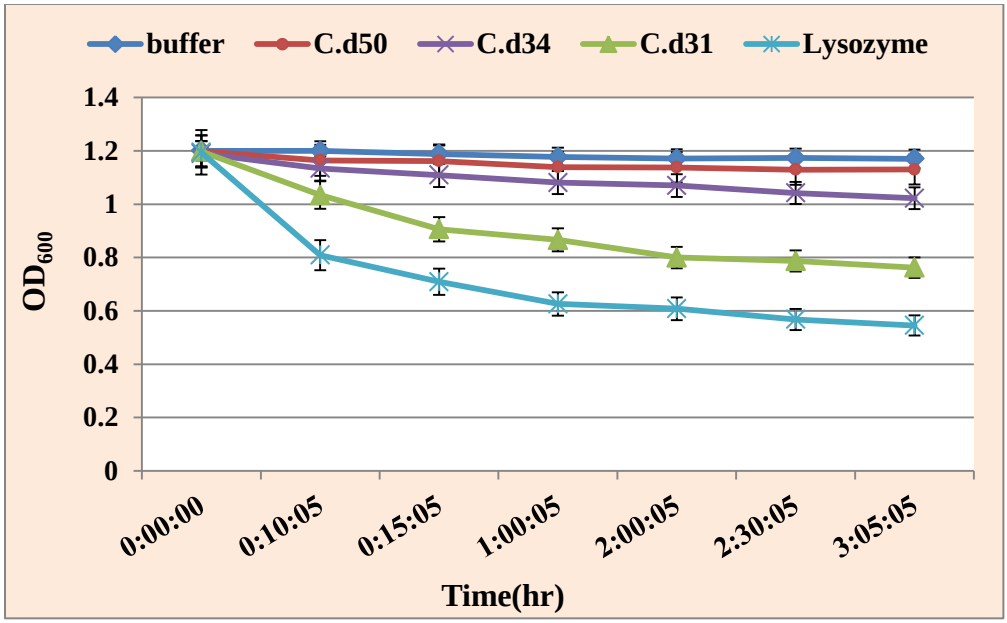


Figure 6.19 . Lytic activity of the C_{50},C_{34},C_{31} of *C.difficile* proteins and lysozyme against decoated spores of *C.difficile* DS1813. Values are the means of results from triplicate assays ± standard deviations

6.4 Discussion

Little is known about the process that triggers *C.difficile* spore germination and the subsequent lysis of the spore cortex. While a bioinformatic comparison between the genome sequence of *C.difficile* and other Gram positive spore formers has found no homology to known germination receptors, it has revealed homology to spore peptidoglycan lytic enzymes such as SleC (David *et al.,* 2009; Ramirez *et al*., 2010). Indeed, a recent study showed that the inactivation of the gene in *C.difficile* prevented germination (David *et al.,* 2009).

In this study, we identified and characterized a putative spore cortex lytic enzyme encoded by the SleC gene of *C.difficile* 630 with homology to the SleC gene of *C. perfringens*. This was achieved by codon optimizing the gene sequences of each gene for expression in *E.coli* and subsequently expressing the full-length (C_{50}) and catalytic domain (C_{31}) of each gene. The lytic activity of each recombinant protein against coated and decoated spores of *C.perfringens* and against coated, decoated and exosporium deficient (sonicated) spores of *C.difficile* was determined.

To be biologically active, the full length SleC fragment (C_{50}) is cleaved during sporulation to generate the inactive proenzyme (C_{35}), which is then proteolytically cleaved by a spore associated protease (GSP) to generate the active form C_{31} (Urakami *et al*., 1999).

Miyata and colleagues were able to isolate the SleC in an active from (C_{31}) from germinated spores, which showed lytic activity against coatless spores (Miyata *et al*., 1995a).

Furthermore, a study by Urakami and colleagues reported that the inactive form of SleC (C_{35}), which was isolated from dormant spores, exhibited lytic activity against coatless spores only after treatment with a GSP (Urakami *et al*., 1999).

However, the previous study reported that the full length SleC sequence (C_{50}) *C.perfringens* (as a recombinant protein) produced only one fragment with ~ 34 kDa after treatment with GSP: this fragment also showed lytic activity against decoated spores (Okamura *et al*., 2000).

Our results revealed that the full length recombinant SleC (C_{50}) derived from *C.perfringens* and *C.difficile* were digested by GSP but only by the spore extract obtained from *C.difficile* DS1813. The failure of spore extracts from other strains examined in this study to lyse SleC could be due to the inefficiency of the GSP extraction method. That the spore extract of *C.difficile* DS1813 produced similarly sized digestion products supports the hypothesis that the SleC homolog in *C.difficile* plays a similar role to that of SleC in the germination of *C.perfringens*. Indeed, these data support the idea that *C.difficile* and *C.perfringens* may share a common ancestor.

In our study, spores were decoated chemically by agents that rupture disulfide bonds in the spore coat, such as urea and DTT (Gould and Hitchins, 1963). The TEM results showed that this method was able to remove the spore coat and no broken spores were

found. With exosporium deficient *C.difficile* spores, spores were sonicated at 15 μm, in 50 seconds bursts over 16 cycles. We also demonstrated by TEM results that this method successfully removed spore exosporium without damaging the spore coat, maintaining the spore ultrastructure: this result is supported by previous results seen by Escobar-Cortés *et al.* (2013).

Lysozyme was used in this study to gauge the lytic activity of our spore lysin enzymes, since it is well known as a PG hydrolase and is widely used (Heffron *et al.,* 2011). Under our reaction conditions, lysozyme exhibited a high level of activity compared to those seen by the spore lysin enzymes and was only active against decoated spores of *C.difficile* (Escobar-Cortés *et al*., 2013) and *C.perfringens* (Ando, 1979)

The biological activity of our recombinant proteins was determined against coated and decoated spores of *C.perfringens* spores and against coated, decoated and exosporium deficient spores of *C.difficile*. We found that only the C31 form of both lysins was able to break down spore peptidoglycan, but to do so required the removal of the exosporium and spore coat. As with other studies, our work confirmed that the spore coat was the major barrier to access of externally applied lysins (Escobar-Cortés *et al.,* 2013).

Further biochemical characterisation of these enzymes would be required in order to achieve the full lytic activity.

We also found that the activated products of SleC failed to lyse the peptidoglycan of vegetative bacteria, while phage derived endolysins also lacked activity against decoated spores of *C.difficile*. This finding confirms that fundamental differences exist in the structure of vegetative and spore derived peptidoglycan.

6.5 Conclusions

- The SleC of *C.difficile* is similar to that of *C.perfringens* and plays a key role in mediating germination.
- The spore coat acts a permeability barrier to externally applied SleC, which could hinder its future development as a therapeutic agent.
- The peptidoglycans of the vegetative and spore form of *C.difficile* differ in structure.

Chapter Seven

General discussion

7.1 Introduction

Clostridium difficile produces large oval sub-terminal spores which facilitate environmental survival due to their resistance to heat, drying, and exposure to commonly used disinfectants. Healthy adults are protected from *C.difficile* colonization and disease by normal human colonic microbiota; however, disruption of this protective community by broad spectrum antibiotics allows spores, acquired via the faecal-oral route, to colonize and germinate in the small bowel upon exposure to bile acid, and thereafter to produce vegetative cells which proliferate in the colon, where they bring about the pathological effects of the disease.

While the current treatments of choice, metronidazole and vancomycin, are generally effective in promoting recovery, the increased incidence of CDI and treatment failure associated with antibiotic resistance combined with the emergence of hypervirulent strains highlights the need to develop approaches which specifically target the pathogen without causing collateral damage to the protective microbiota. Several non-antibiotic approaches have been investigated, such as toxin-binding resins, probiotics, immunotherapy and faecal bacteriotherapy. While promising, these approaches have yet to demonstrate their efficacy in large-scale controlled studies.

In this study, we attempted to isolate lytic bacteriophages capable of specifically targeting *C.difficile* in a manner that enables the re-establishment of a protective bacterial ecosystem.

We also investigated the potential of employing recombinant bacteriophage derived endolysins as therapeutics capable of targeting vegetative *C.difficile*. To complement this effort, we investigated the feasibility of using a spore cortex lytic enzyme (SleC) as a means of directly targeting the spore form of the bacterial cell.

7.2 Bacteriophages and their products as potential therapeutics

Our attempts to isolate a virulent, lytic bacteriophage specific for *C.difficile* that could form the basis of a future therapeutic were not successful. This result was consistent with earlier attempts, which also failed to isolate such a phage (Goh *et al*., 2005b; Horgan *et al*., 2010). Failure to isolate such a phage could in part be due to the fact that the host cell spends much of its time as a spore and as such lacks the receptors that enable phages to initiate an infection (Goh *et al*., 2005b). It is also believed that there is a selective pressure for *C.difficile*-specific bacteriophages to be lysogenic: the spore provides protection for the bacteriophage against harsh environmental conditions while the bacteriophage carries genes with functions that can be adapted and used by the host (Stewart and Levin, 1984; Letarov and Kulikov, 2009; Horgan *et al*., 2010). Based on our results, when attempting to recover *C.difficile*-specific bacteriophages from the environment, it may make more sense to first isolate the bacteria and then attempt to recover any bacteriophages that they may contain. Such an approach would differ markedly from approaches previously reported in the literature.

While we failed to isolate lytic phages, we were able to induce the expression of temperate phages, with twelve *Myoviridae* family phages being recovered from twenty-three clinical isolates following treatment of mid-log phase cultures of *C.difficile* with mitomycin C . This mitogen was found to be more efficient than norfloxacin, possibly due to its different mechanism of action (Ingrey *et al*., 2003; Shan *et al.,* 2012; Meessen-Pinard *et al.,* 2012).

The expression of temperate phages of *C.difficile* has been shown to occur at every stage of cell growth, with phages being recovered during the early (Fortier and Moineau, 2007), mid-log (Shan *et al*., 2012) and stationary phases (Sell *et al.,* 1983; Goh *et al*., 2005; Mayer *et al.,* 2008; Horgan *et al.,* 2010). Some strains of *C.difficile* have even been reported to release phages following sporulation (Nale *et al*., 2012). In common with these studies, we found that induction during mid-log phase growth yielded the greatest phage recovery, for reasons that have yet to be determined.

Our ability to recover phages from the clinical isolates examined in this study was greatly hindered by our inability to identify a susceptible indicator strain which could be used to propagate newly induced prophages. Other researchers have reported a similar problem (Fortier and Moineau, 2007; Nale *et al*., 2012). One possible explanation is that *C.difficile* has developed immunity to phage infection through the use of mechanisms such as the CRISPR system, where the bacterial cell retains a small region of DNA from the infecting phage, which it uses to inhibit subsequent infection with related phages (Barrangou *et al.,* 2007; Deveau *et al.,* 2008; Sebaihia *et al*., 2006).

Evidence to support the hypothesis that *C.difficile* isolates have been exposed to constant rounds of bacteriophage infection is provided in Chapter 5. Our results suggest that bacteriophage infection is a common event amongst clinical isolates of *C.difficile*, an observation supported by the fact that all strains of *C.difficile* sequenced to date were found to harbour at least one phage at their genome (Sebaihia *et al.*, 2006 Goh *et al.*, 2007, He *et al.*, 2010; Nale *et al*., 2012; Shan *et al.*, 2012).

In this study, we found that seventy-nine of the eighty-five clinical *C. difficile* isolates examined by PCR gave at least one positive hit, with some isolates responding to all five primer pairs, suggesting a history of multiple infections. Interestingly we found no correlation between phage infection and *C.difficile* ribotype.

Thus it is tempting to speculate, based on our results, that repeated infection with different bacteriophages has resulted in a bacterial population that is inherently resistant to infection with temperate phages, which may explain the difficulty we have faced in finding an isolate of *C.difficile* that supports virulent phage production (Goh *et al.*, 2005b; Fortier and Moineau, 2007; Mayer *et al*., 2008; Horgan *et al*., 2010; Nale *et al*., 2012; Shan *et al.*, 2012).

The fact that we were able to recover temperate phages from clinical isolates that were PCR negative suggest that new bacteriophages await discovery and that the burden of past phage infection is likely to extend beyond the phages that have been identified to date. In addition, these new phages, once characterised, could represent a new source of endolysins.

The possibility of employing temperate phages to treat *C.difficile* infection has been described in the literature. Ramesh and colleagues reported that *C.difficile*-specific phage 140 was able to prevent *C.difficile*-induced ileocecitis in hamsters (Ramesh *et al.*,1999), while φCD27 has been successfully shown to target *C.difficile* and inhibit toxin production in a laboratory model of a human colon system (Meader *et al.*, 2010). While these results are promising, there are many problems associated with the use of this type of phage as a therapeutic agent. They include:

- The phages have a narrow host range.
- They are susceptible to gastric acidity.
- Phage treatment can increase the sporulation rate in *C.difficile*, which may increase the risk of transmission and relapse infection.
- Treatment with these phages could result in their integration into the genome, which could render the bacteria resistant to further infection with related phages and could increase the risk of gene transfer.
- These phages are not stable (Clokie *et al.*, 2011)

For these reasons, alternative approaches are currently being investigated, such as exploiting the potential of genomics. Access to an increasing number of *C.difficile*-specific temperate phage genomes has enabled researchers to identify genes that play a key role in the biology of the virus. Indeed, an alternative approach to employing the entire phage is to investigate the therapeutic potential of individual phage proteins such as the endolysins. These enzymes are responsible for lysing the vegetative cell wall of the bacterial cells and allowing it to escape from its prey to infect other neighbouring bacteria.

Endolysin genes have been cloned and expressed as recombinant proteins and have been shown to lyse the target bacterial cells when applied externally (Fischetti, 2005; Loessner *et al*., 1998, Meyer *et al*., 2008). One advantage of employing endolysins over the whole phage is the fact that treatment with endolysins is unlikely to lead to the development of resistance even following repeated exposure (Borysowski *et al*., 2006; Loeffler *et al*., 2001; Schuch *et al*., 2002). This is thought to be due to the fact that the endolysins targets molecules in the cell wall that are important in bacterial viability (Borysowski *et al*., 2006).

As part of this study, we identify and expressed as recombinant proteins five *C.difficile* phage encoded endolysins from φCD119, φCD2 ,φCD27 φCD6356 and φCD38-2. They all shared the same modular structure, comprising a single N-terminal and C-terminal domain, and were all class 3 amidase endolysins. However, despite the fact that they shared a high percentage of amino acid identity, they differed in their lytic activity, possibly due to differences in the binding affinity of their C terminal domains (CBD).

Interestingly, the lytic activity of the endolysins derived from *Siphoviridae* phages (φCD6356 and φCD38-2) differed markedly from that seen using endolysins derived from *Myoviridae* family phages (φCD27, φC2, and φCD119). In addition, the failure to observe synergistic activity between combinations of these endolysins supports the hypothesis that they act against the same target, as one might expect, given that they are all class 3 amidases (Proença *et al*., 2012).

The recombinant endolysins produced in this study were found to be biologically active against a wide range of *C.difficile* isolates representing a diversity of ribotypes. They were also found to be specific for *C.difficile* in that they showed little or no activity against the other bacterial species included in this study, thus highlighting their potential as novel, specific microbial agents. Indeed, our characterization of endolysin LysCD27 and its variant, LyCD227, provides a starting point from which to investigate the ability of genetic engineering to improve the activity of these endolysins. As we demonstrated in Chapter 4, subtle change in endolysin sequences seems to have a major impact on the endolysin activity.

While these endolysins showed impressive activity under *in vitro* conditions, further studies need to be undertaken under *in vivo* conditions to assess their efficacy in the complex biological environment that is the gut. Moreover, studies need to be undertaken to determine the most efficient method of formulating and delivering endolysins to their target bacteria in the gut. A number of approaches could be employed to achieve this aim, such as encapsulation (Zhong and Jin, 2009; Liu *et al.*, 2010) or the expression of a *C.difficile*-specific phage lysin from a strain capable of colonizing the gut, such as *Lactococcus Lactis* (Mayer *et al*., 2008) .

7.3 *C. difficile* genome-encoded lysins as potential therapeutics

In addition to examining the potential of bacteriophage-derived endolysins, we also demonstrated the feasibility of employing a bioinformatics-based approach to identify potential endolysins located in the genome of *C.difficile*. Among the sequences identified within the genome of *C.difficile* 630 was a putative endolysin gene (CD630, YP_001088405). A copy of this gene, as determined by PCR, was detected in all of the *C.difficile* isolates examined in this study, suggesting that this gene could represent an evolutionary relic of past phage infection of an early ancestor of *C.difficile*. Furthermore, a recombinant version of this gene was indeed able to lyse the *C.difficile* cell wall in a manner similar to that of the phage-derived endolysins, suggesting that it could be a strong candidate for development as a future therapeutic agent. Furthermore, its universal and conserved nature suggests that it may have potential as a diagnostic marker.

7.4 Germination-specific cortex-lytic enzymes of *C. difficile* as potential therapeutics

The mechanism of spore germination in *C.difficile* is unclear; however, it has been shown that bile salts play an important role in mediating germination (Wilson, *et al.*, 1982; Wilson, 1983; Sorg and Sonenshein, 2008; Giel, *et al.*, 2010).

While it has been proposed that *C.difficile* spores have a germination receptor which responds to taurocholate and glycine, to date, no homologs to known germination receptors have been identified in the *C.difficile* genome (Ramirez, *et al.*, 2010). While studies have failed to identify germination receptors, they have revealed the presence of genes with homology to previously described spore cortex lysins, which in the case of *C.difficile* CD630 comprise four genes, CD0551 (*sleC*), CD0552 (*sleB*), CD0106 (*cwID*) and CD3563, which is thought to encode a putative cortex lytic enzyme (Sebaihia *et al*., 2006; David *et al*., 2009).

Our bioinformatics analysis of these genes revealed that CD0551, which encodes for a homolog of sleC, was highly conserved (98% homology or higher) in the nine *C.difficile* strains sequenced to date and shared 53% amino acid identity with the SleC gene of *C. perfringens* (David *et al.*, 2009). In this strain, SleC plays an essential role in germination, as it lyses the peptidoglycan in the spore cortex (Paredes-Sabja, *et al.*, 2009a). Subsequent gene knock-out studies by David and colleagues confirmed that the SleC of *C.difficile* plays a similar role in the germination of *C.difficile* (David *et al*., 2009).

Further confirmation of the similarity of the SleC of the two bacteria was provided by the observation that full length recombinant SleC (C_{50}) derived from *C.perfringens* and *C.difficile* when digested by GSP isolated from spores of *C.difficile* DS1813 generated identical sized fragments of approximately 34 kDa and 16 kDa. The ability of the GSP from *C.difficile* to process the SleC of *C.perfringens* lends support to the idea that the two bacteria employ the same process. However, our results showed that even though these enzymes play the same role, they only have lytic activity against their own bacteria.

Another similarity was also seen when the lytic activity of these enzymes was tested against coated and decoated spores. These enzymes were shown to be active against coatless spores only, showing that spore coat is a major barrier in both strains and these enzymes are not permeable. Furthermore, due to this property, it would be expected that the enzyme is located in the spore coat, as in the case of *C.perfringens* SleC.

This study confirmed that there are fundamental differences in the structure of vegetative and spore derived peptidoglycan. We found that the recombinant SleC of *C.difficile* failed to lyse peptidoglycan extracted from vegetative bacteria, while phage derived endolysins had no activity against decoated spores of *C.difficile*. We also found that lysozyme failed to lyse the PG of vegetative cells but was active against decoated spores. This finding is in agreement with the previous studies by Peltier and colleagues, who linked lysosome resistance to the presence of a high proportion of non-acetylated amino sugars in the peptidoglycan of vegetative cells (Peltier *et al*., 2011), and another study by Escobar-Cortés and colleagues, who showed that lysozyme was active against coatless spores of *C.difficile* (Escobar-Cortés *et al*., 2013).

As far as we are aware, this is the first report of such differences in *C.difficile,* although other bacteria such as *B.subtilis* (Atrih and Foster, 1999) and *C.perfringens* (Marquis and Bender, 1990; Makino and Moriyama, 2002) are known to differ in a similar manner.

7.5 Suggestions for future studies

- Investigate the potential of the gene encoding the CD630 lysin as a diagnostic marker of *C.difficile.*
- Isolate further bacterial lysins and determine their ability to target vegetative *C.difficile* - seek to identify a lysin with glucosaminidase activity which could be used in combination with the lysins with class 3 amidase activity that we have already isolated.
- Determine the activity of each endolysin in combination with antibiotics with a view to identifying synergistic formulations such as the endolysin (Cpl-1) and penicillin, which has been shown to have synergistic activity against a penicillin-resistant strain of *S. pneumoniae* (Djurkovic *et al.*,2005) .
- Studies should be undertaken to examine the administration of the endolysin in combination with therapies that neutralise the activity of the toxins, such as antibodies and tolevamer (Babcock *et al*., 2006; Barker *et al*., 2006).
- Investigate the feasibility of employing molecular biology to enhance the activity of recombinant lysins by modifying lysin structure, as has been demonstrated by Meyer *et al*. (2010).
- Examine means of delivering the lysins to the gut. Approaches that could be employed include the use of genetically-modified lactic acid strain such as *Lactococcus Lactis*, in order to overcome the challenge of delivering a bacteriophage endolysin to the GI tract.

- The failure of SleC to lyse the cortex of whole spores when applied externally represents an obstacle to its use as a therapeutic agent. Studies are required to identify spore permabilising agents that could be used to enable SleC to access its target site.
- Further studies should be undertaken to characterise the processing of SleC in *C.difficile* spores following the induction of germination.

7.6 Conclusions

Our attempts to isolate lytic bacteriophages which could form the basis of a treatment for *C.difficile* were unsuccessful. This failure may in part be due to repeated episodes of phage infection over time which have resulted in the emergence of "phage resistant" species mediated by systems such as CRISPR. For this reason, we investigated the antibacterial potential of phage-encoded lytic enzymes and were able to demonstrate their activity against multiple clinical isolates of the vegetative form of the pathogen.

We also found that a genome associated lysin (CD630, YP_001088405) was able to lyse vegetative *C.difficile* and that this gene was present in all of the clinical isolates of *C.difficile* examined in this study. Given its apparent universal nature, this gene has the potential to be developed as a diagnostic marker for this pathogen.

In addition to examining bacteriophage-associated endolysins, we cloned and expressed a *C.difficile* spore cortex enzyme (SleC), which was able to lyse the peptidoglycan of the spore when applied externally, but only once the outer layers of the spore had been removed, suggesting that they acted as a permeability barrier. The sequence similarity between the SleC genes of *C.perfringens* and *C.difficile* and the fact that their expressed proteins are processed in a similar manner suggest that these bacteria share similar germination pathways.

The finding that phage lysins were unable to lyse spore peptidoglycan and that SleC was equally ineffective against vegetative peptidoglycan supports the supposition that the peptidoglycan compositions of these two structures differ markedly. Thus, any lysin-based therapeutic agent which seeks to target all phases of the pathogen's life cycle should include phage and spore derived lysins.

References

Abhilash, M., Vidya, A. , and Jagadevi ,T. (2009) Bacteriophage Therapy: A War Against Antibiotic Resistant Bacteria. *The Internet Journal of Alternative Medicine* 7(1).

Aboudola, S., Kotloff, K.L. , Kyne, L. , Warny, M. , Kelly, E.C. , Sougioultzis, S. , Giannasca, P.J. , Monath, T.P. , and Kelly, C.P. (2003). *Clostridium difficile* vaccine and serum immunoglobulin G antibody response to toxin A. *Infection and Immunity* 71(3), pp. 1608-1610.

Ackermann, H.W. (1996). Frequency of morphological phage descriptions in 1995: Brief Review. *Archives of Virology* 141(2), pp. 209-218.

Ackermann, H. W. (2001). Frequency of morphological phage descriptions in the year 2000. *Archives of Virology* 146(5), pp. 843-857.

Adams, M. H. (1959). Bacteriophages. New York: Interscience publishers.

Ahn, S.J., and Burne, R.A. (2007). Effects of oxygen on biofilm formation and the AtlA autolysin of *Streptococcus mutans*. *Journal of Bacteriology* 189(17), pp. 6293-6302.

Alisky, J., Iczkowski, K., Rapoport, A. , and Troitsky, N. (1998). Bacteriophages show promise as antimicrobial agents. *Journal of Infection* 36(1), pp. 5-15.

Alonso , R., Muñoz , C. ,Gros , S. ,García de Viedma ,D. ,Peláez, T., and Bouza, E. (1999). Rapid detection of toxigenic *Clostridium difficile* from stool samples by a nested PCR of toxin B gene. *Journal of Hospital Infection* 41(2), pp. 145-149.

Altschul, S. F., Madden, T.L., Schäffer, 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 (Review). *Nucleic Acids Research* 25(17), pp. 3389-3402.

Ando, Y. (1979). Spore lytic enzyme released from *Clostridium perfringens* spores during germination. *Journal of Bacteriology* 140(1), pp. 59-64.

Arroyo, L. G., Kruth, S.A. , Willey, B.M. , Staempfli, H.R. , Low, D.E. , and Weese, J.S. (2005). PCR ribotyping of *Clostridium difficile* isolates originating from human and animal sources. *Journal of Medical Microbiology* 54(2), pp. 163-166.

Aspinall, S. T., and Hutchinson, D.N (1992). New selective medium for isolating *Clostridium difficile* from faeces. *Journal of Clinical Pathology* 45(9), pp. 812-814.

Atrih, A., and Foster, S.J. (2001). Analysis of the role of bacterial endospore cortex structure in resistance properties and demonstration of its conservation amongst species. *Journal of Applied Microbiology* 91(2), pp. 364-372.

Atrih, A., and Foster, S.J. (2002). Bacterial endospores the ultimate survivors .*International Dairy Journal* 12(2-3), pp. 217-223.

Babcock, G. J., Broering, T.J. , Hernandez, H.J. , Mandell, R.B. , Donahue, K. , Boatright, N. , Stack, A.M. , Lowy, I. , Graziano, R. , Molrine, D. , Ambrosino, D.M. , and Thomas Jr., W.D. (2006). Human monoclonal antibodies directed against toxins A and B prevent *Clostridium difficile*-induced mortality in hamsters. *Infection and Immunity* 74(11), pp. 6339-6347.

Bagyan, I., and Setlow, P (2002). Localization of the cortex lytic enzyme *CwlJ* in spores of *Bacillus subtilis*. *Journal of Bacteriology* 184(4), pp. 1219-1224.

Baines, S. D., Freeman, J., and Wilcox, M.H. (2009). Tolevamer is not efficacious in the neutralization of cytotoxin in a human gut model of *Clostridium difficile* infection. *Antimicrobial Agents and Chemotherapy* 53(5), pp. 2202-2204.

Baker, J. R., Liu, C. , Dong, S. , and Pritchard, D.G. (2006) . Endopeptidase and glycosidase activities of the bacteriophage B30 lysin. *Applied and Environmental Microbiology* 72(10), pp. 6825-6828.

Barbut, F., Richard, A., Hamadi, K., Chomette, V., Burghoffer, B., and Petit, J.C. (2000). Epidemiology of recurrences or reinfections of *Clostridium difficile*- associated diarrhea. *Journal of Clinical Microbiology* 38(6), pp. 2386-2388.

Barer, M. R. (2007). Morphology and nature of microorganisms. In: Greenwood D, S.R., Peutherer JF and Barer MR. ed. *Medical Microbiology: A Guide to Microbial Infections: Pathogenesis, Immunity, Laboratory Diagnosis and Control* .17th ed. London: Churchill Livingstone pp. 8-23.

Barker Jr., R. H., Dagher, R., Davidson, D.M. , and Marquis, J.K. (2006). Tolevamer, a novel toxin-binding polymer: overview of preclinical pharmacology and physiochemical properties. *Alimentary Pharmacology and Therapeutics* 24(11-12), pp. 1525-1534.

Baross, J. A., Liston, J., and Morita, R.Y.(1978). Incidence of *Vibrio parahaemolyticus* bacteriophages and other Vibrio bacteriophages in marine samples. *Applied and Environmental Microbiology* 36(3), pp. 492-499.

Barrangou, R., Fremaux, C. , Deveau, H. , Richards, M. , Boyaval, P. , Moineau, S. , Romero, D.A. , and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819), pp. 1709-1712.

Barrangou, R., and Horvath, P. (2009). The CRISPR system protects microbes against phages, plasmids. *Microbe* 4(5), pp. 224-230.

Bartlett, J. G. (2002). Clinical practice. Antibiotic-associated diarrhea. *New England journal of medicine* 346(5), pp. 334-339.

Bartlett, J. G. (2006). Narrative review: The new epidemic of *Clostridium difficile* -Associated enteric disease. *Annals of Internal Medicine* 145(10), pp. 758-764.

Bauer, M. P., and van Dissel, J.T. (2009). Alternative strategies for *Clostridium difficile* infection. *International Journal of Antimicrobial Agents* 33(S1), pp. S51-S56 .

Bayne-Jones, S., and Petrilli, A. (1933). Cytological changes during the formation of the endospore in *Bacillus megaterium*. *Journal of Bacteriology* 25(3), pp. 261-275.

Bennett, G. C. J., Allen, E. , and Millard, P. H. (1984). *Clostridium difficile* diarrhea: A highly infectious organism. *Age and Ageing* 13(6), pp. 363-366.

Benson, D. A., Karsch-Mizrachi, I., Clark, K., Lipman, D.J., Ostell, J., and Sayers, E.W. (2012). GenBank. *Nucleic Acids Research* 40(D1), pp. D48-D53.

Bernhardt, T. G., Wang, I.N., Struck, D.K., and Young, R. (2002). Breaking free:"Protein antibiotics" and phage lysis. *Research in Microbiology* 153(8), pp. 493-501.

Bienvenut, W. V., Deon, C., Pasquarello, C., Campbell, J.M. Sanchez, J.C., Vestal, M.L. and Hochstrasser, D.F. (2002). Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. *Proteomics*. 2(7), pp. 868-876.

Bignardi , G. (1998). Risk factors for *Clostridium difficile* infection. *Journal of Hospital Infection*, 40(1), pp. 1-15.

Blackman, S. A., Smith, T.J., and Foster, S.J. (1998). The role of autolysins during vegetative growth of *Bacillus subtilis* 168. *Microbiology* 144(1), pp. 73-82.

Boland, F. M., Atrih, A., Chirakkal, H., Foster, S.J. , and Moir, A. (2000). Complete spore-cortex hydrolysis during germination of *Bacillus subtilis* 168 requires *SleB* and *YpeB*. *Microbiology* 146(1), pp. 57-64.

Bolotin, A., Quinquis, B., Sorokin, A., and Dusko Ehrlich, S.(2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151(8), pp. 2551-2561.

Borriello, S. P., Davies, H.A., and Barclay, F.E. (1988). Detection of fimbriae amongst strains of *Clostridium difficile*. *FEMS Microbiology Letters* 49(1), pp. 65-67.

Borriello , S. P. (1998). Pathogenesis of *Clostridium difficile* infection. *Journal of Antimicrobial Chemotherapy* 41(SC), pp. 13–19.

Borysowski, J., Weber-Dąbrowska, B. , and Górski, A.(2006). Bacteriophage endolysins as a novel class of antibacterial agents. *Experimental Biology and Medicine* 231(4), pp. 366-377.

Bradley, D. E. (1967). Ultrastructure of bacteriophage and bacteriocins. *Bacteriological reviews* 31(4), pp. 230-314.

Brazier, J. S., and Borriello, S.P. (2000). Microbiology,Epidemiology and Diagnosis of *Clostridium difficile* Infection. In: Aktories, K., and Wilkins, T.D. ed. *Current Topics in Microbiology and Immunology*. Vol. 250. Springer, pp. 1-33.

Brennan, P., Shore, A.M., Clement, M., Hewamana, S., Jones, C.M., Giles, P., Fegan, C., Pepper, C. and Brewis, I.A. (2009). Quantitative nuclear proteomics reveals new phenotypes altered in lymphoblastoid cells. *Proteomics - Clinical Applications* 3(3), pp. 359-369.

Briers, Y., Volckaert, G. , Cornelissen, A. , Lagaert, S. , Michiels, C.W. , Hertveldt, K. , Lavigne, R. (2007). Muralytic activity and modular structure of the endolysins of *Pseudomonas aeruginosa* bacteriophages φKZ and EL. *Molecular Microbiology* 65(5), pp. 1334-1344.

Brock, T. (1975). Milestones in Microbiology. Washington: American Society for Microbiology.

Brüssow, H., Canchaya, C. , and Hardt, W.-D. (2004). Phages and the evolution of bacterial pathogens: From genomic rearrangements to lysogenic conversion. *Microbiology and Molecular Biology Reviews* 68(3), pp. 560-602.

Bruttin, A., and Brussow, H. (2005). Human volunteers receiving *E. coli* phage T4 orally: a safety test of phage therapy. *Antimicrobial Agents and Chemotherapy* 49(7), pp. 2874-2878.

Buggy, B. P., Wilson, K.H., and Fekety, R. (1983). Comparison of methods for recovery of *Clostridium difficile* from an environmental surface. *Journal of Clinical Microbiology* 18(2), pp. 348-352.

Bult, C. J., White, O. , Olsen, G.J. , Zhou, L. , Fleischmann, R.D. , Sutton, G.G. , Blake, J.A., FitzGerald, L.M. , Clayton, R.A. , Gocayne, J.D. , Kerlavage, A.R. , Dougherty, B.A. , Tomb, J.-F. , Adams, M.D. , Reich, C.I. , Overbeek, R. , Kirkness, E.F. , Weinstock, K.G. , Merrick, J.M. , Glodek, A. , Scott, J.L. , Geoghagen, N.S.M. , Weidman, J.F. , Fuhrmann, J.L. , Nguyen, D. , Utterback, T.R. , Kelley, J.M. , Peterson, J.D., Sadow, P.W. , Hanna, M.C. , Cotton, M.D. , Roberts, K.M. , Hurst, M.A. , Kaine, B.P. , Borodovsky, M. , Klenk, H.-P. , Fraser, C.M. , Smith, H.O. , Woese, C.R. , and Venter, J.C.(1996). Complete genome sequence of the *Methanogenic archaeon, Methanococcus jannaschii*. *Science* 273(5278), pp. 1058-1073.

Burns, D. A., Heap, J.T., and Minton, N.P. (2010). SleC is essential for germination of *Clostridium difficile* spores in nutrient-rich medium supplemented with the bile salt taurocholate. *Journal of Bacteriology* 192(3), pp. 657-664.

Castagliuolo, I., Keates, A.C. , Wang, C.C. , Pasha, A. , Valenick, L. , Kelly, C.P. , Nikulasson, S.T. , LaMont, J.T. , and Pothoulakis, C. (1998). *Clostridium difficile* toxin A stimulates macrophage-inflammatory protein-2 production in rat intestinal epithelial cells. *The Journal of Immunology* 160(12), pp. 6039-6045.

Chen, F., and Suttle, C.A. (1996). Evolutionary relationships among large double-stranded DNA viruses that infect microalgae and other organisms as inferred from DNA polymerase genes. *Virology* 219(1), pp. 170-178.

Chen, F., Suttle, C.A., and Short, S.M. (1996). Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. *Applied and Environmental Microbiology* 62(8), pp. 2869-2874.

Chen, Y., Miyata, S., Makino, S., and Moriyama, R.(1997). Molecular characterization of a germination-specific muramidase from *Clostridium perfringens* S40 spores and nucleotide sequence of the corresponding gene. *Journal of Bacteriology* 179(10), pp. 3181-3187.

Cheng, Q., and Fischetti, V.A. (2007). Mutagenesis of a bacteriophage lytic enzyme PlyGBS significantly increases its antibacterial activity against group B *streptococci*. *Applied Microbiology and Biotechnology* 74(6), pp. 1284-1291.

Cheng, Q., Nelson, D., Zhu, S., and Fischetti, V.A. (2005). Removal of group B *streptococci* colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. *Antimicrobial Agents and Chemotherapy* 49(1), pp. 111-117.

Clokie MR, M. A., Letarov AV, and Heaphy S. (2011). Phages in nature. *Bacteriophage* 1(1), pp. 31-45.

Collins, M. D., Lawson, P.A., Willems, A., Cordoba, J.J., Fernandez-Garayzabal, J., Garcia, P., Cai, J., Hippe, H., and Farrow, J.A.E. (1994). The phylogeny of the genus Clostridium: Proposal of five new genera and eleven new species combinations. *International Journal of Systematic Bacteriology* 44(4), pp. 812-826.

Courchesne, N. M., Parisien, A. , and Lan, C.Q. (2009). Production and application of bacteriophage and bacteriophage-encoded lysins. *Recent Patents on Biotechnology*. 3(1), pp. 37-45.

Cudmore, S. L., Delgaty, K.L., Hayward-McClelland, S.F., Petrin, D.P., and Garber, G.E.(2004). Treatment of infections caused by metronidazole-resistant *Trichomonas vaginalis*. *Clinical Microbiology Reviews* 17(4), pp. 783-793.

De Hoon, M. J. L., Eichenberger, P. , and Vitkup, D. (2010). Hierarchical evolution of the bacterial sporulation network. *Current Biology* 20(17), pp. R735-R745.

Dei, R. (1989). Observations on phage-typing of *Clostridium difficile*: Preliminary evaluation of a phage panel. *European Journal of Epidemiology* 5(3), pp. 351-354.

Denève, C., Janoir, C. , Poilane, I. , Fantinato, C. , and Collignon, A. (2009). New trends in *Clostridium difficile* virulence and pathogenesis. *International Journal of Antimicrobial Agents* 33(S1), pp. S24-S28.

Desrosier, J. P., and Lara, J.C. (1984). Synthesis of the exosporium during sporulation of *Bacillus cereus*. *Journal of General Microbiology* 130(4), pp. 935-940.

Deveau, H., Barrangou, R. , Garneau, J.E. , Labonté, J. , Fremaux, C. , Boyaval, P. , Romero, D.A. , Horvath, P. , and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *Journal of Bacteriology* 190(4), pp. 1390-1400.

Dhalluin, A., Bourgeois, I., Pestel-Caron, M., Camiade, E., Raux, G., Courtin, P., Chapot-Chartier, M.-P., and Pons, J.-L. (2005). *Acd*, a peptidoglycan hydrolase of *Clostridium difficile* with N-acetylglucosaminidase activity. *Microbiology* 151(7), pp. 2343-2351.

Dial, S., Delaney, J.A.C. , Barkun, A.N. , and Suissa, S. (2005). Use of gastric acid-suppressive agents and the risk of community-acquired *Clostridium difficile*-associated disease. *Journal of the American Medical Association* 294(23), pp. 2989-2995.

Djurkovic, S., Loeffler, J.M., and Fischetti, V.A. (2005). Synergistic killing of *Streptococcus pneumoniae* with the bacteriophage lytic enzyme Cpl-1 and penicillin or gentamicin depends on the level of penicillin resistance. *Antimicrobial Agents and Chemotherapy* 49(3), pp. 1225-1228.

Doern, G. V., Coughlin, R.T., and Wu, L. (1992). Laboratory diagnosis of *Clostridium difficile*-associated gastrointestinal disease: Comparison of a monoclonal antibody enzyme immunoassay for toxins A and B with a monoclonal antibody enzyme immunoassay for toxin A only and two cytotoxicity assays. *Journal of Clinical Microbiology* 30(8), pp. 2042-2046.
Douglas, J. (1975). Bacteriophages. London: Chapman and Hall.

Dowd, M. M., Orsburn, B., and Popham, D.L. (2008). Cortex peptidoglycan lytic activity in germinating *Bacillus anthracis* spores. *Journal of Bacteriology* 190(13), pp. 4541-4548.

Edwards, D. I. (1993). Nitroimidazole drugs - Action and resistance mechanisms I. Mechanisms of action. *Journal of Antimicrobial Chemotherapy* 31(1), pp. 9-20.

Errington, J. (2003). Regulation of endospore formation in *Bacillus subtilis*. *Nature reviews. Microbiology* 1(2), pp. 117-126.

Escobar-Cortés, K., Barra-Carrasco, J., and Paredes-Sabja, D. (2013). Proteases and sonication specifically remove the exosporium layer of spores of *Clostridium difficile* strain 630. *Journal of Microbiological Methods* 93(1), pp. 25-31.

Fekety, R., Silva, J., Kauffman, C., Buggy, B., and Gunner Deery, H. (1989). Treatment of antibiotic-associated *Clostridium difficile* colitis with oral vancomycin: Comparison of two dosage regimens. *American Journal of Medicine* 86(1), pp. 15-19.

Fekety, R. (1997). Guidelines for the diagnosis and management of *Clostridium difficile*-associated diarrhea and colitis. *American Journal of Gastroenterology* 92(5), pp. 739-750.

Fernandes, S., Proença, D., Cantante, C., Silva, F.A., Leandro, C., Lourenço, S., Milheiriço, C., De Lencastre, H., Cavaco-Silva, P., Pimentel, M., and São-José, C. (2012). Novel chimerical endolysins with broad antimicrobial activity against methicillin-resistant *Staphylococcus aureus*. *Microbial Drug Resistance* 18(3), pp. 333-343.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., Gavin, O.L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E.L.L., Eddy, S.R., and Bateman, A. (2010). The Pfam protein families database. *Nucleic Acids Research* 38(D211–D222).

Fischetti, V. A. (2008). Bacteriophage lysins as effective antibacterials. *Current Opinion in Microbiology* 11(5), pp. 393-400.

Florea, N. R., Kuti, J.L., Nightingale, C.H., and Nicolau, D.P. (2003). Treatment of *Clostridium difficile*-associated disease (CDAD). *Connecticut Medicine* 67(3), pp. 153-155.

Fortier, L.-C., and Moineau, S. (2007). Morphological and genetic diversity of temperate phages in *Clostridium difficile*. *Applied and Environmental Microbiology* 73(22), pp. 7358-7366.

Fortier, L. C., and Moineau, S. (2009). Phage production and maintenance of stocks, including expected stock lifetimes. *Methods in molecular biology* 501, pp. 203-219.

Fortuna, W., Międzybrodzki, R. , Weber-Dąbrowska, B. , and Górski, A.(2008). Bacteriophage therapy in children: Facts and prospects. *Medical Science Monitor* 14(8), pp. RA126-RA132.

Fuller, N. J., Wilson, W.H. , Joint, I.R., and Mann, N.H. (1998) .Occurrence of a sequence in marine cyanophages similar to that of T4 g20 and its application to PCR-based detection and quantification techniques. *Applied and Environmental Microbiology* 64(6), pp. 2051-2060.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* 31(13), pp. 3784-3788.

Gebhart, D., Williams, S.R., Bishop-Lilly, K.A., Govoni, G.R., Willner, K.M., Butani, A., Sozhamannan, S., Martin, D., Fortier, L.-C., and Scholl, D. (2012). Novel high-molecular-weight, R-type bacteriocins of *Clostridium difficile*. *Journal of Bacteriology* 194(22), pp. 6240-6247.

George, W. L., Sutter, V.L., Citron, D., and Finegold, S.M. (1979). Selective and differential medium for isolation of *Clostridium difficile*. *Journal of Clinical Microbiology* 9(2), pp. 214-219.

Gerding, D. N., Johnson, S., Peterson, L.R., Mulligan, M.E., and Silva Jr., J. (1995). *Clostridium difficile*-associated diarrhea and colitis. *Infection control and hospital epidemiology* 16(8), pp. 459-477.

Gerding, D. N. (2000). Treatment of Clostridium associated disease and colitis. In: Aktories K, .W.T. ed. *Clostridium difficile.* Berlin: Springe, pp. 127-140.

Gerhardt, P., and Ribi, E. (1964). Ultrastructure of the exosporium enveloping spores of *Bacillus cereus*. *Journal of Bacteriology* 88(6), pp. 1774-1789.

Giel, J. L., Sorg, J.A., Sonenshein, A.L.b, and Zhu, J. (2010). Metabolism of bile salts in mice influences spore germination in *Clostridium difficile*. *PLoS ONE 5* (1), p. Article number e8740.

Gill, J. J., and Hyman, P. (2010). Phage choice, isolation, and preparation for phage therapy. *Current Pharmaceutical Biotechnology* 11(1), pp. 2-14.

Glückmann, M., Fella, K., Waidelich, D., Merkel, D., Kruft, V., Kramer, P.J., Walter, Y., Hellmann, J., Karas, M. and Kroger, M. (2007). Prevalidation of potential protein biomarkers in toxicology using iTRAQ reagent technology. *Proteomics* 7(10), pp. 1564-1574.

Gnezda-Meijer, K., Mahne, I., Poljšak-Prijatelj, M., and Stopar, D. (2006). Host physiological status determines phage-like particle distribution in the lysate. *FEMS Microbiology Ecology* 55(1), pp. 136-145.

Godde, J. S., and Bickerton, A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *Journal of Molecular Evolution* 62(6), pp. 718-729.

Goh, S., Ong, P.F. , Song, K.P. , Rily, T.V. , and Chang, B.J. (2007). The complete genome sequence of Clostridium difficile phage φC2 and comparisons to φCD119 and inducible prophages of CD630. *Microbiology* 153(3), pp. 676-685.

Goh, S., Chang, B.J., and Riley, T.V. (2005a). Effect of phage infection on toxin production by *Clostridium difficile*. *Journal of Medical Microbiology* 54(2), pp. 129-135.

Goh, S., Riley, T.V. , and Chang, B.J. (2005b). Isolation and characterization of temperate bacteriophages of *Clostridium difficile*. *Applied and Environmental Microbiology* 71 (2), pp. 1079-1083.

Golshahi, L., Seed, K.D. , Dennis, J.J. , and Finlay, W.H. (2008). Toward Modern Inhalational Bacteriophage Therapy: Nebulization of Bacteriophages of *Burkholderia cepacia* Complex. *Journal of Aerosol Medicine and Pulmonary Drug Delivery* 21(4), pp. 351-360.

Gombas, D. E., Labbe, R.G. (1981). Extraction of spore-lytic enzyme from *Clostridium perfringens* spores. *Journal of General Microbiology* 126(1), pp. 37-44.

Gomi, M., Sonoyama, M., and Mitaku, S. (2004). High performance system for signal peptide prediction: SOSUIsignal. *Chem-Bio Informatics Journal* 4(4), pp. 142-147.

Goorhuis, A., Debast, S.B., Van Leengoed, L.A.M.G., Harmanus, C., Notermans, D.W., Bergwerff, A.A., and Kuijper, E.J. (2008). *Clostridium difficile* PCR ribotype 078: An emerging strain in humans and in pigs? .*Journal of Clinical Microbiology* 46(3), p. 1157.

Gorbach, S. L., Barza, M., Giuliano, M., and Jacobus, N.V. (1988). Colonization resistance of the human intestinal microflora: Testing the hypothesis in normal volunteers. *European Journal of Clinical Microbiology* 7(1), pp. 98-102.

Gottesman, S., Halpern, E., and Trisler, P. (1981). Role of *sulA* and *sulB* in filamentation by lon mutants of *Escherichia coli* K-12. *Journal of Bacteriology* 148(1), pp. 265-273.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research* 38(Issue SUPPL. 2), pp. W695-W699.

Gould, G. W., and Hitchins, A.D. (1963). Sensitization of bacterial spores to lysozyme and hydrogen peroxide with reagents which rupture disuphide bonds. *Nature* 197(4867), p. 622.

Gould , G. W. (1970). Symposium on bacterial spores: IV. Germination and the problem of dormancy. *Journal of Applied Bacteriology* 33(1), pp. 34-49.

Govind, R., Vediyappan, G. , Rolfe, R.D. , Dupuy, B. , and Fralick, J.A. (2009). Bacteriophage-mediated toxin gene regulation in *Clostridium difficile*. *Journal of Virology* 83(23), pp. 12037-12045.

Govind, R., Fralick, J.A. , and Rolfe, R.D. (2006). Genomic organization and molecular characterization of *Clostridium difficile* bacteriophage φCD119. *Journal of Bacteriology* 188(7), pp. 2568–2577.

Goyal, M. S. (1987). Methods in Phage Ecology. In: Goyal, S.M., Gerba, G.P., and Bitton, G ed. *Phage Ecology*. New York: John Wiley & Sons, pp. 267-284.

Grabow, W. O. K., and Coubrough, P. (1986). Practical direct plaque assay for coliphages in 100-ml samples of drinking water. *Applied and Environmental Microbiology* 52(3), pp. 430-433.

Grabow, W. O. K. (2001). Bacteriophages: Update on application as models for viruses in water. *Water SA* 27(2), pp. 251-268.

Greenstein, A. J., Byrn, J.C.,Zhang, L.P.,Swedish, K.A.,Jahn, A.E., and Divino, C.M. (2008). Risk factors for the development of fulminant *Clostridium difficile* colitis. *Surgery* 143(5), pp. 623-629.

Gründling, A., Manson, M.D. , and Young, R. (2001). Holins kill without warning. *Proceedings of the National Academy of Sciences of the United States of America* 98(16), pp. 9348-9352.

Gu, J., Xu, W., Lei, L., Huang, J., Feng, X., Sun, C., Du, C., Zuo, J., Li, Y., Du, T., Li, N., and Han, W. (2011). LysGH15, a novel bacteriophage lysin, protects a murine bacteremia model efficiently against lethal methicillin-resistant *Staphylococcus aureus* infection. *Journal of Clinical Microbiology* 49(1), pp. 111-117.

Guerrant, R. L., Van Gilder, T., Steiner, T.S., Thielman, N.M., Slutsker, L., Tauxe, R.V., Hennessy, T., Griffin, P.M., DuPont, H., Sack, R.B., Tarr, P., Neill, M., Nachamkin, I., Reller, L.B., Osterholm, M.T., Bennish, M.L., and Pickering, L.K. (2001). Practice guidelines for the management of infectious diarrhea. *Clinical Infectious Diseases* 32(3), pp. 331-351.

Gumerlock, P. H., Tang, Y.J., Meyers, F.J., and Silva Jr., J.(1991). Use of the polymerase chain reaction for the specific and direct detection of *Clostridium difficile* in human feces. *Reviews of Infectious Diseases* 13(6), pp. 1053-1060.

Gustafsson, A., Berstad, A., Lund-Tønnesen, S. , Midtvedt, T. , and Norin, E. (1999). The effect of faecal enema on five microflora-associated characteristics in patients with antibiotic-associated diarrhoea. *Scandinavian Journal of Gastroenterology* 34(6), pp. 580-586.

Guttman, B., Raya, R. and Kutter, E. (2005). Basic phage biology. In: Kutter, E., and Sulakvelidze, A. ed. *Bacteriophages* :Biology and Applications. Boca Raton, FL: CRC Press pp. 29-66.

Hafiz, S., and Oakley, C.L. (1976). *Clostridium difficile* isolation and characteristics. *Journal of Medical Microbiology* 9 (2), pp. 129-136.

Haft, D. H., Selengut, J., Mongodin, E.F., and Nelson, K.E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Computational Biology* 1(6), pp. 0474-0483**.**

Hall, J. C., and O'Toole,E. (1935). Intestinal flora in newborn infants with description of a new pathogenic anaerobe,*Bacillus difficile*. *American Journal of Disease in Children* 49(1935), pp. 390–402.

Hanlon, G. W. (2007). Bacteriophages: an appraisal of their role in the treatment of bacterial infections. *International Journal of Antimicrobial Agents* 30 (2), pp. 118-128.

Hayes, W. (1968). The Genetics of Bacteria and their Viruses. 2nd ed. London: Blackwell Scientific Publications Ltd.

He, M., Sebaihia, M., Lawley, T.D., Stabler, R.A., Dawson, L.F., Martin, M.J., Holt, K.E., Seth-Smith, H.M.B., Quail, M.A., Rance, R., Brooks, K., Churcher, C., Harris, D., Bentley, S.D., Burrows, C., Clark, L., Corton, C., Murray, V., Rose, G., Thurston, S., Van Tonder, A., Walker, D., Wren, B.W., Dougan, G., and Parkhill, J. (2010). Evolutionary dynamics of *Clostridium difficile* over short and long time scales. *Proceedings of the National Academy of Sciences of the United States of America* 107(16), pp. 7527-7532.

Health Protection Agency (2010-2011) Annual Report. *Clostridium difficile* Ribotyping Network (CDRN) for England and Northern Ireland .

Heffron, J. D., Sherry, N., and Popham, D.L. (2011). In vitro studies of peptidoglycan binding and hydrolysis by the *Bacillus anthracis* germination-specific lytic enzyme SleB. *Journal of Bacteriology* 193(1), pp. 125-131.

Hermans, P. W. M., Van Soolingen, D., Bik, E.M., De Haas, P.E.W., Dale, J.W., and Van Embden, J.D.A. (1991). Insertion element IS987 from *Mycobacterium bovis* BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium tuberculosis* complex strains. *Infection and Immunity* 59(8), pp. 2695-2705.

Hermoso, J. A., Monterroso, B., Albert, A., Galán, B., Ahrazem, O., García, P., Martínez-Ripoll, M., García, J.L., and Menéndez, M. (2003). Structural basis for selective recognition of pneumococcal cell wall by modular endolysin from phage Cp-1. *Structure* 11 (10), pp. 1239-1249.

Hickson, M., D'Souza, A.L. , Muthu, N. , Rogers, T.R., Want, S. , Rajkumar, C. , and Bulpitt, C.J. (2007). Use of probiotic *Lactobacillus* preparation to prevent diarrhoea associated with antibiotics: Randomised double blind placebo controlled trial. *British Medical Journal* 335(7610), pp. 80-83.

Holt, S. C., and Leadbetter, E.R (1969). Comparative ultrastructure of selected aerobic spore-forming bacteria: a freeze-etching study. *Bacteriological reviews* 33(2), pp. 346-378.

Horgan, M., O'Sullivan, O., Coffey, A. , Fitzgerald, G.F. , van Sinderen, D. , McAuliffe, O. , and Ross, R.P. (2010). Genome analysis of the *Clostridium difficile* phage φCD6356, a temperate phage of the *Siphoviridae* family. *Gene* 462(1-2), pp. 34-43.

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of Bacteria and Archaea. *Science* 327(5962), pp. 167-170.

Ingrey, K. T., Ren, J., and Prescott, J.F. (2003). A fluoroquinolone induces a novel mitogen-encoding bacteriophage in *Streptococcus canis*. *Infection and Immunity* 71(6), pp. 3028-3033.

Innis, M. A., and Gelfand, D .H. (1990). Optimization of PCR's. I. In: Innis, M.A., D.H. Gelfend, J.J. Sninsky and T.J. White ed. : *A Guide to Methods and Applications*. New York: Academic Press, pp. 3-12.

Ishikawa, S., Yamane, K., and Sekiguchi, J. (1998). Regulation and characterization of a newly deduced cell wall hydrolase gene (*cwlJ*) which affects germination of *Bacillus subtilis* spores. *Journal of Bacteriology* 180(6), pp. 1375-1380.

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

Janion, C. (2001). Some aspects of the SOS response system - A critical survey. *Acta Biochimica Polonica* 48 (3), pp. 599-610.

Jank, T., Giesemann, T., and Aktories, K. (2007). Rho-glucosylating *Clostridium difficile* toxins A and B: New insights into structure and function (Review). *Glycobiology* 17(4), pp. 15R-22R.

Jansen, R., Van Embden, J.D.A. , Gaastra, Wim. , and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology* 43(6), pp. 1565-1575 .

Jensen, E. C., Schrader, H.S. , Rieland, B. , Thompson, T.L. , Lee, K.W. , Nickerson, K.W. , and Kokjohn, T.A. (1998). Prevalence of broad-host-range lytic bacteriophages of *Sphaerotilus natans*, *Escherichia coli*, and *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology* 64(2), pp. 575-580.

Johnson, L. P., Tomai, M.A., and Schlievert, P.M. (1986). Bacteriophage involvement in the group A streptococcal pyrogenic exotoxin A production. *Journal of Bacteriology* 166(2), pp. 623-627.

Jonathan, C. (2009). Bacteriophage therapy: old treatment, new focus? *Microbiologist* 10(2), pp. 20-23.

J**oshi, L. T., Phillips, D.S., Williams, C.F., Alyousef, A., and Baillie, L. (2012)**. Contribution of spores to the ability of *Clostridium difficile* to adhere to surfaces. *Applied and Environmental Microbiology* 78(21), pp. 7671-7679.

Jump, R. L. P., Pultz, M.J. ,and Donskey, C.J. (2007). Vegetative *Clostridium difficile* survives in room air on moist surfaces and in gastric contents with reduced acidity: A potential mechanism to explain the association between proton pump inhibitors and *C. difficile*-associated diarrhea?. *Antimicrobial Agents and Chemotherapy* . 51(8), pp. 2883-2887.

Kakikawa, M., Yokoi, K.-J. , Kimoto, H. , Nakano, M. , Kawasaki, K.-I. , Taketo, A. , and Kodaira, K.-I. (2002). Molecular analysis of the lysis protein Lys encoded by *Lactobacillus plantarum* phage φg1e. *Gene* 299(1-2), pp. 227-234.

Kamerbeek, J., Schouls, L. , Kolk, A. , Van Agterveld, M. , Van Soolingen, D. , Kuijper, S. , Bunschoten, A. , Molhuizen, H. , Shaw, R. , Goyal, M. , and Van Embden, J. (1997). Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *Journal of Clinical Microbiology* 35(4), pp. 907-914.

Karginov, F. V., and Hannon, G.J. (20100. The CRISPR System: Small RNA-Guided Defense in Bacteria and Archaea. *Molecular Cell* 37(1), pp. 7-19.

Kato, N., Ou, C.-Y., Kato, H., Bartley, S.L., Brown, V.K., Dowell Jr., V.R., and Ueno, K. (1991). Identification of toxigenic *Clostridium difficile* by the polymerase chain reaction. *Journal of Clinical Microbiology* 29(1), pp. 33-37.

Kelly, C. P., Pothoulakis, C., and LaMont, J.T. (1994). *Clostridium difficile* colitis. *New England Journal of Medicine* 330(4), pp. 257-262.

Kelly, C. P., and LaMont, J. (1998). *Clostridium difficile* infection. *Annual Review of Medicine* 49, pp. 375-390.

Kikkawa, H., Fujinami, Y., Suzuki, S.-i., and Yasuda, J. (2007). Identification of the amino acid residues critical for specific binding of the bacteriolytic enzyme of γ-phage, PlyG, to *Bacillus anthracis*. *Biochemical and Biophysical Research Communications* 363(3), pp. 531-535.

Kotloff, K. L., Wasserman, S.S., Losonsky, G.A., Thomas W., Jr., Nichols, R., Edelman, R., Bridwell, M., and Monath, T.P. (2001). Safety and immunogenicity of increasing doses of a *Clostridium difficile* toxoid vaccine administered to healthy adults. *Infection and Immunity* 69 (2), pp. 988-995.

Kuipers, E. J., and Surawicz, C. M (2008). *Clostridium difficile* infection. *The Lancet* 371 (9623), pp. 1486-1488.

Kumazawa, T., Masayama, A., Fukuoka, S., Makino, S., Yoshimura, T., and Moriyama, R. (2007). Mode of action of a germination-specific cortex-lytic enzyme, SleC, of *Clostridium perfringens* S40. *Bioscience, Biotechnology and Biochemistry* 71(4), pp. 884-892.

Kunin, V., Sorek, R., and Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biology* 8, R61 (4).

Kutateladze, M., and Adamia, R. (2008). Phage therapy experience at the Eliava Institute. *Medecine et Maladies Infectieuses* 38(8), pp. 426-430.

Kuzminov, A. (1999). Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage λ . *Microbiology and Molecular Biology Reviews* 63(4), pp. 751-813.

Labbe, R. G. (2005). Sporulation (Morphology) of Clostridia. In: Dürre, P. ed. *Handbook on Clostridia*. CRC Press pp. 647-665.

Lamont, I., Brumby, A.M., and Egan, J.B. (1989). UV induction of coliphage 186: Prophage induction as an SOS function. *Proceedings of the National Academy of Sciences of the United States of America* 86(14), pp. 5492-5496.

Larkin, M. A., Blackshields, G., Brown, N.P., Chenna, R., Mcgettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21), pp. 2947-2948.

Larson , H. E., Parry , J. V. , Price,A. B., Davies, D. R. , Dolby, J. , and Tyrrell, D. A. (1977). Undescribed toxin in pseudomembranous colitis. *British Medical Journal* 1(6071), pp. 1246-1248.

Lawley, T. D., Croucher, N.J. , Yu, L. , Clare, S. , Sebaihia, M. , Goulding, D. , Pickard, D.J. , Parkhill, J. , Choudhary, J. , and Dougan, G. (2009). Proteomic and genomic characterization of highly infectious *Clostridium difficile* 630 spores. *Journal of Bacteriology* 191(17), pp. 5377-5386.

Leighton, T. J., and Doi, R.H.(1971). The stability of messenger ribonucleic acid during sporulation in *Bacillus subtilis*. *Journal of Biological Chemistry* 246(10), pp. 3189-3195.

Letarov, A., and Kulikov, E. (2009). The bacteriophages in human- and animal body-associated microbial communities. *Journal of Applied Microbiology* 107(1), pp. 1-13.

Liang, M. E. K. (2003). *Clostridium difficile* infection. *Reviews in Clinical Gerontology* 13(1), pp. 1-9.

Lillestøl, R. K., Redder, P., Garrett, R.A., and Brügger, K. (2006). A putative viral defence mechanism in archaeal cells. *Archaea* 2(1), pp. 59-72.

Little, J. W. (1991). Mechanism of specific LexA cleavage: autodigestion and the role of RecA coprotease. *Biochimie* 73(4), pp. 411-421.

Little, J. W. (2005). lysogen, prophage induction, and lysogenic conversion In: Matthew, K.W., David, I. F., and Sankar, L. A. ed. *Phages: Their Role in Bacterial Pathogenesis and Biotechnology*. USA: ASM press, pp. 37-54.

Liu, D., Guo, Y. , Wang, Z., and Yuan, J. (2010). Exogenous lysozyme influences *Clostridium perfringens* colonization and intestinal barrier function in broiler chickens. *Avian Pathology* 39(1), pp. 17-24.

Loeffler, J. M., Djurkovic, S., and Fischetti, V.A. (2003). Phage Lytic Enzyme Cpl-1 as a Novel Antimicrobial for Pneumococcal Bacteremia. *Infection and Immunity* 71(11), pp. 6199-6204.

Loeffler, J. M., Nelson, D., and Fischetti, V.A (2001) Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. *Science* 294(5549), pp. 2170-2172.

Loessner, M. J., Wendlinger, G., and Scherer, S. (1995). Heterogeneous endolysins in *Listeria monocytogenes* bacteriophages: A new class of enzymes and evidence for conserved holin genes within the siphoviral lysis cassettes. *Molecular Microbiology* 16(6), pp. 1231-1241.

Loessner, M. J., Maier, S.K. , Daubek-Puza, H. , Wendlinger, G. , and Scherer, S. (1997). Three *Bacillus cereus* bacteriophage endolysins are unrelated but reveal high homology to cell wall hydrolases from different bacilli. *Journal of Bacteriology* 179(9), pp. 2845-2851.

Loessner, M. J., Gaeng, S., and Scherer, S. (1999). Evidence for a holin-like protein gene fully embedded out of frame in the endolysin gene of *Staphylococcus aureus* bacteriophage 187. *Journal of Bacteriology* 181 (15), pp. 4452-4460.

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. *Molecular Microbiology* 44(2), pp. 335-349.

Loessner, M. J. (2005). Bacteriophage endolysins - Current state of research and applications. *Current Opinion in Microbiology* 8 (4), pp. 480-487.

Louie, T. J., Peppe, J., Watt, C.K., Johnson, D., Mohammed, R., Dow, G., Weiss, K., Simon, S., John, J.F. , Garber, G., Chasan-Taber, S., Davidson, D.M. (2006). Tolevamer, a novel nonantibiotic polymer, compared with vancomycin in the treatment of mild to moderately severe *Clostridium difficile*- associated diarrhea. *Clinical Infectious Diseases* 43(4), pp. 411-420.

Low, L. Y., Yang, C., Perego, M., Osterman, A., and Liddington, R.C. (2005). Structure and lytic activity of a *Bacillus anthracis* prophage endolysin. *Journal of Biological Chemistry* 280(42), pp. 35433-35439.

Lu, T. K., and Collins, J.J. (2007). Dispersing biofilms with engineered enzymatic bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America* 104(27), pp. 11197-11202.

Lund, B. M., Gee, J.M., King, N.R., Horne , R. W. and Harnden, J. M. (1978). The structure of the exosporium of a pigmented Clostridium. *Journal of General Microbiology* 105(2), pp. 165-174 .

Macgregor, H. E. (2006). Latest food additive: Viruses. Los Angeles Times. Sect. F-1. August 28, 2006.

Mahony, D. E., Bell, P.D., and Easterbrook, K.B. (1985). Two bacteriophages of *Clostridium difficile* . *Journal of Clinical Microbiology* 21(2), pp. 251-254.

Mahony, D. E., Clow, J., Atkinson, L., Vakharia, N., and Schlech, W.F. (1991). Development and application of a multiple typing system for *Clostridium difficile* .*Applied and Environmental Microbiology* 57(7), pp. 1873-1879.

Makarova, K. S., Aravind, L. , Grishin, N.V. , Rogozin, I.B. , and Koonin, E.V. (2002). A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Research* 30(2), pp. 482-496.

Makino, S., and Moriyama, R. (2002). Hydrolysis of cortex peptidoglycan during bacterial spore germination. *Medical Science Monitor* 8(6), pp. RA119-127.

Makoff, A. J., Oxer, M.D., Romanos, M.A., Fairweather, N.F., and Ballantine, S. (1989). Expression of tetanus toxin fragment C in *E. coli*: High level expression by removing rare codons. *Nucleic Acids Research* 17(24), pp. 10191-10202.

Marquis, R. E., and Bender, G.R. (1990). Compact structure of cortical peptidoglycans from bacterial spores. *Canadian Journal of Microbiology* 36(6), pp. 426-429.

Marraffini, L. A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in *staphylococci* by targeting DNA. *Science* 322(5909), pp. 1843-1845.

Marraffini, L. A., and Sontheimer, E.J. (2009). Invasive DNA, Chopped and in the CRISPR. *Structure* 17(6), pp. 786-788.

Marston, M. F., and Sallee, J.L. (2003). Genetic diversity and temporal variation in the cyanophage community infecting marine *Synechococcus* species in Rhode Island's Coastal Waters. *Applied and Environmental Microbiology* 69(8), pp. 4639-4647.

Marza, J. A. S., Soothill, J.S. , Boydell, P. , and Collyns, T.A. (2006). Multiplication of therapeutically administered bacteriophages in *Pseudomonas aeruginosa* infected patients. *Burns* 32(5), pp. 644-646.

Masayama, A., Fukuoka, H., Kato, S., Yoshimura, T., Moriyama, M., Moriyama, R. (2006). Subcellular localization of a germiantion-specific cortex-lytic enzyme, SleB, of Bacilli during sporulation. *Genes and Genetic Systems* 81(3), pp. 163-169.

Matsushiro, A., Sato, K. , Miyamoto, H. , Yamamura, T. , and Honda, T. (1999). Induction of prophages of enterohemorrhagic *Escherichia coli* O157:H7 with norfloxacin. *Journal of Bacteriology* 181(7), pp. 2257-2260.

Matthews, B. W., and Remington, S.J. (1974). The three dimensional structure of the lysozyme from bacteriophage T4. *Proceedings of the National Academy of Sciences of the United States of America* 71(10), pp. 4178-4182.

Mattila, E., Anttila, V.-J., Broas, M., Marttila, H. , Poukka, P. , Kuusisto, K. , Pusa, L. , Sammalkorpi, K. , Dabek, J. , Koivurova, O.-P. , Vähätalo, M. , Moilanen, V. , and Widenius, T. (2008). A randomized, double-blind study comparing *Clostridium difficile* immune whey and metronidazole for recurrent *Clostridium difficile*-associated diarrhoea: Efficacy and safety data of a prematurely interrupted trial. *Scandinavian Journal of Infectious Diseases* 40(9), pp. 702-708.

Mayer, M. J., Gasson, M.J.,and Narbad, A. (2012). Genomic sequence of bacteriophage ATCC 8074-B1 and activity of its endolysin and engineered variants against *Clostridium sporogenes. Applied and Environmental Microbiology* 78(10), pp. 3685-3692.

Mayer, M. J., Narbad, A., and Gasson, M.J. (2008). Molecular characterization of a *Clostridium difficile* bacteriophage and its cloned biologically active endolysin. *Journal of Bacteriology* 190(20), pp. 6734–6740.

Mayer, M. J., Garefalaki, V., Spoerl, R., Narbad, A., and Meijers, R (2011). Structure-based modification of a *Clostridium difficile*-targeting endolysin affects activity and host range. *Journal of Bacteriology* 193(19), pp. 5477-5486.

Mazzocco, A., Waddell, T.E., Lingohr, E., and Johnson, R.P (2009). Enumeration of bacteriophages by the direct plating plaque assay. *Methods in molecular biology* 501, pp. 77-80.

McDonald, J. E., Smith, D.L., Fogg, P.C.M., McCarthy, A.J., and Allison, H.E. (2010). High-throughput method for rapid induction of prophages from lysogens and its application in the study of shiga toxin-encoding *Escherichia coli* strains. *Applied and Environmental Microbiology* 76(7), pp. 2360-2365.

McDonnel G.E (2007). Antisepsis, Disinfection, and Sterilization: Types, Action, and Resistance. Washington, DC: ASM Press.

McFarland, L. V. (1998). Epidemiology, Risk Factors and Treatments for Antibiotic-Associated Diarrhea. *Digestive Diseases* 16(5), pp. 292-307.

McNulty, C., Logan, M. , Donald, I.P. , Ennis, D., Taylor, D. , Baldwin, R.N. , Bannerjee, M. , and Cartwright, K.A.V. (1997). Successful control of *Clostridium difficile* infection in an elderly care unit through use of a restrictive antibiotic policy. *Journal of Antimicrobial Chemotherapy* 40(5), pp. 707-711.

Meader, E., Mayer, M.J. , Gasson, M.J. , Steverding, D. , Carding, S.R. , and Narbad, A. (2010). Bacteriophage treatment significantly reduces viable *Clostridium difficile* and prevents toxin production in an in vitro model system. *Anaerobe* 16(6), pp. 549-554.

Meador-Parton, J., and Popham, D.L. (2000). Structural analysis of *Bacillus subtilis* spore peptidoglycan during sporulation. *Journal of Bacteriology* 182(16), pp. 4491-4499.

Medzihradszky, K. F., Campbell, J.M., Baldwin, M.A., Falick, A.M. Juhasz, P., Vestal, M.L. and Burlingame, A.L (2000). The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Analytical Chemistry* 72(3), pp. 552-558.

Meessen-Pinard, M., Sekulovic, O., and Fortier, LC. (2012). Evidence of In Vivo Prophage Induction during *Clostridium difficile* Infection. *Applied and Environmental Microbiology* 78(21), pp. 7662-7670.

Miller, M. A. (2007). Clinical management of *Clostridium difficile*-Associated Disease. *Clinical Infectious Diseases* 45, pp. 122-128.

Miyata, S., Moriyama, R., Miyahara, N., and Makino, S. (1995b). A gene (*sleC*) encoding a spore-cortex-lytic enzyme from *Clostridium perfringens* S40 spores; cloning, sequence analysis and molecular characterization. *Microbiology* 141(10), pp. 2643-2650.

Miyata, S., Moriyama, R., Sugimoto, K., and Makino, S. (1995a). Purification and partial characterization of a spore cortex-lytic enzyme of *Clostridium perfringens* S40 spores. *Bioscience, biotechnology, and biochemistry* 59(3), pp. 514-515.

Moir, A. (2006). How do spores germinate? *Journal of Applied Microbiology* 101(3), pp. 526–530.

Mojica, F. J. M., Díez-Villaseñor, C. , Soria, E. , and Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology* 36(1), pp. 244-246.

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution* 60(2), pp. 174-182.

Monot, M., Boursaux-Eude, C., Thibonnier, M., Vallenet, D., Moszer, I., Medigue, C., Martin-Verstraete, I., and Dupuy, B. (2011). Reannotation of the genome sequence of *Clostridium difficile* strain 630. *Journal of Medical Microbiology* 60(8), pp. 1193-1199.

Mullan, W. M. A. (2002). Plaque formation by bacteriophage. [Online]. Available at: http://www.dairyscience.info/enumeration-of-lactococcal-bacteriophages/plaque-formation.html (Accessed: 25 March, 2010).

Muniesa, M., and Jofre, J. (1998). Abundance in sewage of bacteriophages that infect *Escherichia coli* O157:H7 and that carry the Shiga toxin 2 gene. *Applied and Environmental Microbiology* 64(7), pp. 2443-2448.

Murphy, F., Fauquet, C.M., Bishop, D.H.L., Ghabrial, S.A., Jarvis, A.W., Martelli, G.P., Mayo, M.A., and Summers, M.D. (1995) Virus taxonomy : classification and nomenclature of viruses : sixth report of the International Committee on Taxonomy of Viruses.Springer-Verlag, Wien.

Myers, G. S. A., Rasko, D.A., Cheung, J.K., Ravel, J., Seshadri, R., DeBoy, R.T., Ren, Q., Varga, J., Awad, M.M., Brinkac, L.M., Daugherty, S.C., Haft, D.H., Dodson, R.J., Madupu, R., Nelson, W.C., Rosovitz, M.J., Sullivan, S.A., Khouri, H., Dimitrov, G.I., Watkins, K.L., Mulligan, S., Benton, J., Radune, D., Fisher, D.J., Atkins, H.S., Hiscox, T., Jost, B.H., Billington, S.J., Songer, J.G., McClane, B.A., Titball, R.W., Rood, J.I., Melville, S.B., and Paulsen, I.T. (2006). Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*. *Genome Research* 16(8), pp. 1031-1040.

Naaber, P., Smidt, I., Štšepetova, J., Brilene, T., Annuk, H., and Mikelsaar, M. (2004). Inhibition of *Clostridium difficile* strains by intestinal *Lactobacillus* species. *Journal of Medical Microbiology* 53(6), pp. 551-554.

Nagarajan, R. (1991). Antibacterial activities and modes of action of vancomycin and related glycopeptides. *Antimicrobial Agents and Chemotherapy* 35(4), pp. 605-609.

Nagy, E., and Foldes, J. (1991). Electron microscopic investigation of lysogeny of *Clostridium difficile* strains isolated from antibiotic-associated diarrhea cases and from healthy carriers. *APMIS* 99(4), pp. 321-326.

Nale, J. Y., Shan, J., Hickenbotham, P.T., Fawley, W.N., Wilcox, M.H.,and Clokie, M.R.J (2012). Diverse temperate bacteriophage carriage in *Clostridium difficile* 027 strains. Appl Environ Microbiol 78(17), pp. 6027-6034.

Navarre, W. W., Ton-That, H. , Faull, K.F. , and Schneewind, O. (1999). Multiple enzymatic activities of the murein hydrolase from *staphylococcal* phage φ11: Identification of a D-alanyl-glycine endopeptidase activity. *Journal of Biological Chemistry* 274(22), pp. 15847-15856.

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

Nerandzic, M. M., and Donskey, C.J. (2010). Triggering germination represents a novel strategy to enhance killing of *Clostridium difficile* spores. *PLoS ONE* 5(8), p. e12285.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* 302(1), pp. 205-217.

Oakley, B. B., Talundzic, E. , Morales, C.A. , Hiett, K.L. , Siragusa, G.R. , Volozhantsev, N.V. , and Seal, B.S. (2011). Comparative genomics of four closely related *Clostridium perfringens* bacteriophages reveals variable evolution among core genes with therapeutic potential. *BMC Genomics* 12, p. Article number282.

O'Flaherty, S., Ross, R.P., and Coffey, A. (2009). Bacteriophage and their lysins for elimination of infectious bacteria. *FEMS Microbiology Reviews* 33 (4), pp. 801-819.

Ohye, D. F., and Murrell, W.G (1973). Exosporium and spore coat formation in *Bacillus cereus* T. *Journal of Bacteriology* 115(3), pp. 1179-1190 .

Okamura, S., Urakami, K., Kimata, M., Aoshima, T., Shimamoto, S., Moriyama, R., and Makino, S. (2000). The N-terminal prepeptide is required for the production of spore cortex-lytic enzyme from its inactive precursor during germination of *Clostridium perfringens* S40 spores. *Molecular Microbiology* 37(4), pp. 821-827.

Oliveira, H., Azeredo, J. , Lavigne, R. , and Kluskens, L.D. (2012). Bacteriophage endolysins as a response to emerging foodborne pathogens. *Trends in Food Science and Technology* 28(2), pp. 103-115.

Paidhungat, M., Ragkousi, K., and Setlow, P. (2001). Genetic requirements for induction of germination of spores of *Bacillus subtilis* by Ca2+-dipicolinate. *Journal of Bacteriology* 183(16), pp. 4886-4893.

Paredes-Sabja, D., Torres, J.A., Setlow, P., and Sarker, M.R. (2008). *Clostridium perfringens* spore germination: Characterization of germinants and their receptors. *Journal of Bacteriology* 190 (4), pp. 1190-1201.

Paredes-Sabja, D., Setlow, P., and Sarker, M.R. (2011). Germination of spores of Bacillales and Clostridiales species: Mechanisms and proteins involved (Review). *Trends in Microbiology* 19(2), pp. 85-94.

Paredes-Sabja, D., Setlow, P., Sarker, M.R. (2009b). The protease CspB is essential for initiation of cortex hydrolysis and dipicolinic acid (DPA) release during germination of spores of *Clostridium perfringens* type A food poisoning isolates. *Microbiology* 155(10), pp. 3464-3472.

Paredes-Sabja, D. , Setlow, P., and Sarker, M.R. (2009a). SleC is essential for cortex peptidoglycan hydrolysis during germination of spores of the pathogenic bacterium *Clostridium perfringens*. *Journal of Bacteriology* 191(8), pp. 2711-2720.

Parisien, A., Allan, B., Zhang, J., Mandeville, R. and Lan, C.Q. (2008). Novel alternatives to antibiotics: Bacteriophages, bacterial cell wall hydrolases, and antimicrobial peptides. *Journal of Applied Microbiology* 104(1), pp. 1-13.

Payne, S., Gibson, G. , Wynne, A. , Hudspith, B. , Brostoff, J. , and Tuohy, K. (2003). In vitro studies on colonization resistance of the human gut microbiota to *Candida albicans* and the effects of tetracycline and *Lactobacillus plantarum* LPK. *Current Issues in Intestinal Microbiology* 4(1), pp. 1-8.

Pecenkova, T., and Paces, V. (1999). Molecular phylogeny of φ29-like phages and their evolutionary relatedness to other protein-primed replicating phages and other phages hosted by gram-positive bacteria. *Journal of Molecular Evolution* 48(2), pp. 197-208.

Peck, M. W. (2009). Biology and Genomic Analysis of *Clostridium botulinum .Advances in Microbial Physiology* 55, pp. 183-265,320.

Peltier, J., Courtin, P., El Meouche, I., Lemée, L., Chapot-Chartier, M.-P., and Pons, J.-L. (2011). *Clostridium difficile* has an original peptidoglycan structure with a high level of N-acetylglucosamine deacetylation and mainly 3-3 cross-links. *Journal of Biological Chemistry* 286(33), pp. 29053-29062.

Perez, J., Springthorpe, V.S., and Sattar, S.A. (2005). Activity of selected oxidizing microbicides against the spores of *Clostridium difficile:* Relevance to environmental control. *American Journal of Infection Control* 33(6), pp. 320-325.

Perkins, D. N., Pappin, D.J., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20(18), pp. 3551-3567.

Petersen, T. N., Brunak, S., Von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: Discriminating signal peptides from transmembrane regions (Letter). *Nature Methods* 8(10), pp. 785-786.

Pochapin, M. (2000). The effect of probiotics on *Clostridium difficile* diarrhea. *American Journal of Gastroenterology* 95(1 SUPPL), pp. S11-S13.

Popham, D. L., Helin, J.b, Costello, C.E., and Setlow, P. (1996). Analysis of the peptidoglycan structure of *Bacillus subtilis* endospores. *Journal of Bacteriology* 178(22), pp. 6451-6458.

Popham, D. L., Sengupta, S., and Setlow, P. (1995). Heat, hydrogen peroxide, and UV resistance of *Bacillus subtilis* spores with increased core water content and with or without major DNA-binding proteins. *Applied and Environmental Microbiology* ,(10)61 pp. 3633-3638.

Pourcel, C., Salvignol, G. , and Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151(3), pp. 653-663.

Poutanen, S. M., and Simor, A.E. (2004). *Clostridium difficile*-associated diarrhea in adults. *Canadian Medical Association Journal* 171(1), pp. 51-58.

Proença, D., Fernandes, S., Leandro, C., Silva, F.A., Santos, S., Lopes, F., Mato, R., Cavaco-Silva, P., Pimentel, M., and São-José, C. (2012). Phage endolysins with broad antimicrobial activity against *Enterococcus faecalis* clinical strains. *Microbial Drug Resistance* 18(3), pp. 322-332.

Purdy, R. N., Dancer, B.N., Day, M.J., and Stickler, D.J (1984). A novel technique for the enumeration of bacteriophage from water. *FEMS Microbiology Letters* 21(1), pp. 89-92.

Ramesh, V., Fralick, J.A., and Rolfe, R.D. (1999). Prevention of *Clostridium difficile*-induced ileocecitis with bacteriophage. *Anaerobe* 5 (2), pp. 69-78.

Ramirez, N., Liggins, M., and Abel-Santos, E. (2010). Kinetic evidence for the presence of putative germination receptors in *Clostridium difficile* spores. *Journal of Bacteriology* 192(16), pp. 4215-4222.

Redmond, C., Baillie, L.W.J. , Hibbs, S. , Moir,A.J.G. ,and Moir, A. (2004). Identification of proteins in the exosporium of *Bacillus anthracis*. *Microbiology* 150(2), pp. 355-363.

Rokyta, D. R., Burch, C.L. , Caudle, S.B. , and Wichman, H.A. (2006). Horizontal gene transfer and the evolution of microvirid coliphage genomes. *Journal of Bacteriology* 188(3), pp. 1134-1142.

Romero, P., Garcia, E., and Mitchell, T.J. (2009). Development of a prophage typing system and analysis of prophage carriage in *Streptococcus pneumoniae*. *Applied and Environmental Microbiology* 75(6), pp. 1642-1649.

Russel, M. (1995). Moving through the membrane with filamentous phages. *Trends in Microbiology* 3(6), pp. 223-228.

Russell, A. D. (1990). Bacterial spores and chemical sporicidal agents. *Clinical Microbiology Reviews* 3(2), pp. 99-119.

Santos, S. B., Carvalho, C.M. , Sillankorva, S. , Nicolau, A. , Ferreira, E.C. , and Azeredo, J. (2009). The use of antibiotics to improve phage detection and enumeration by the double-layer agar technique. *BMC Microbiology* 9(148).

Sanz, J. M., Diaz, E., and Garcia, J.L. (1992). Studies on the structure and function of the N-terminal domain of the pneumococcal murein hydrolases. *Molecular Microbiology* 6(7), pp. 921-931.

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

Sass, P., and Bierbaum, G. (2007). Lytic activity of recombinant bacteriophage φ11 and φ12 endolysins on whole cells and biofilms of *Staphylococcus aureus*. *Applied and Environmental Microbiology* 73(1), pp. 347-352.

Schaeffer, A. B., and Fulton, M.D. (1933). A simplified method of staining endospores. *Science* 77(1990), p. 194.

Schleifer, K. H., and Kandler, O. (1972). Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriological reviews* 36(4), pp. 407-477.

Schmitz, J. E., Ossiprandi, M.C. , Rumah, K.R. , and Fischetti, V.A. (2011). Lytic enzyme discovery through multigenomic sequence analysis in *Clostridium perfringens*. *Applied Microbiology and Biotechnology* 89(6), pp. 1783-1795.

Schuch, R., Nelson, D., Fischetti, V.A. (2002). A bacteriolytic agent that detects and kills *Bacillus anthracis*. *Nature* 418(6900), pp. 884-889.

Sebaihia, M., Wren, B.W. , Mullany, P. , Fairweather, N.F. , Minton, N. , Stabler, R. , Thomson, N.R. , Roberts, A.P. , Cerdeño-Tárraga, A.M. , Wang, H. , Holden, M.T. , Wright, A. , Churcher, C. , Quail, M.A. , Baker, S. , Bason, N. , Brooks, K. , Chillingworth, T. , Cronin, A. , Davis, P. , Dowd, L. , Fraser, A. , Feltwell, T. , Hance, Z. , Holroyd, S. , Jagels, K. , Moule, S. , Mungall, K. , Price, C. , Rabbinowitsch, E. , Sharp, S. , Simmonds, M. , Stevens, K. , Unwin, L. , Whithead, S. , Dupuy, B. , Dougan, G. , Barrell, B. , and Parkhill, J. (2006). The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nature Genetics* 38(7), pp. 779-786.

Seeley, N. D., and Primrose, S.B. (1982). The isolation of bacteriophages from the environment. *Journal of Applied Bacteriology* 53(1), pp. 1-17.

Sekulovic, O., Meessen-Pinard, M., and Fortier, L.-C. (2011). Prophage-stimulated toxin production in *Clostridium difficile* NAP1/027 lysogens. *Journal of Bacteriology* 193(11), pp. 2726-2273.

Sell, T. L., Schaberg, D.R., and Fekety, F.R. (1983). Bacteriophage and bacteriocin typing scheme for *Clostridium difficile* . *Journal of Clinical Microbiology* 17(6), pp. 1148-1152

Setlow, P. (2007). I will survive: DNA protection in bacterial spores. *Trends in Microbiolog* 15(4), pp. 172-180.

Shan, J., Patel, K.V., Hickenbotham, P.T., Nale, J.Y., Hargreaves, K.R., and Clokie, M.R.J. (2012). Prophage carriage and diversity within clinically relevant strains of *Clostridium difficile*. *Applied and Environmental Microbiology* 78(17), pp. 6027-6034.

Sherman, M. (2008). Bacteriophages: Beyond antibiotics. *U.S. Pharmacist* 33(10), pp. 46-51.

Shevchenko, A., Wilm, M., Vorm, O. and Mann M. (1996). Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analytical Chemistry* 68(5), pp. 850-858.

Shimamoto, S., Moriyama, R., Sugimoto, K., Miyata, S., and Makino, S. (2001). Partial characterization of an enzyme fraction with protease activity which converts the spore peptidoglycan hydrolase (SleC) precursor to an active enzyme during germination of *Clostridium perfringens* S40 spores and analysis of a gene cluster involved in the activity. *Journal of Bacteriology* 183(12), pp. 3742-3751.

Shimizu, T., Ohtani, K., Hirakawa, H., Ohshima, K., Yamashita, A., Shiba, T., Ogasawara, N., Hattori, M., Kuhara, S., and Hayashi, H. (2002). Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater. *Proceedings of the National Academy of Sciences of the United States of America* 99(2), pp. 996-1001.

Short, C. M., and Suttle, C.A. (2005). Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. *Applied and Environmental Microbiology* 71(1), pp. 480-486.

Simmons, M., Donovan, D.M., Siragusa, G.R., and Seal, B.S. (2010). Recombinant expression of two bacteriophage proteins that lyse *Clostridium perfringens* and share identical sequences in the C-terminal cell wall binding domain of the molecules but are dissimilar in their N-terminal active domains. *Journal of Agricultural and Food Chemistry* 58(19), pp. 10330-10337.

Skurnik, M. and Strauch, E. (2006). Phage therapy: facts and fiction. *International Journal of Medical Microbiology* 296(1), pp. 5-14.

Šmarda, J., and Benada, O. (2005). Phage tail-like (high-molecular-weight) bacteriocins of *Budvicia aquatica* and *Pragia fontium* (*Enterobacteriaceae*). *Applied and Environmental Microbiology* 71(12), pp. 8970-8973.

Smith, T. J., Blackman, S.A., and Foster, S.J. (2000). Autolysins of *Bacillus subtilis*: Multiple enzymes with multiple functions. *Microbiology* 146(2), pp. 249-262.

Son, B., Yun, J. , Lim, J.-A. , Shin, H. , Heu, S. , and Ryu, S. (2012). Characterization of LysB4, an endolysin from the *Bacillus cereus*-infecting bacteriophage B4. *BMC Microbiology* 12(Article number33).

Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR - A widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Reviews Microbiology* 6(3), pp. 181-186.

Sorg, J. A., and Sonenshein, A.L. (2008). Bile salts and glycine as cogerminants for *Clostridium difficile* spores. *Journal of Bacteriology* 190(7), pp. 2505-2512.

Sougioultzis, S., Kyne, L. , Drudy, D. , Keates, S. , Maroo, S. , Pothoulakis, C. , Giannasca, P.J , Lee, C.K.c , Warny, M. , Monath, T.P. , and Kelly, C.P. (2005). *Clostridium difficile* toxoid vaccine in recurrent C. difficile-associated diarrhea. *Gastroenterology* 128(3), pp. 764-770.

Sousa, J. C. F., Silva, M.T. , and Balassa, G. (1976). An exosporium-like outer layer in *Bacillus subtilis* spores. *Nature* 263(5572), pp. 53-54.

Stabler, R. A., He, M. , Dawson, L. , Martin, M. , Valiente, E. , Corton, C. , Lawley, T.D. , Sebaihia, M. , Quail, M.A. , Rose, G. , Gerding, D.N. , Gibert, M. , Popoff, M.R. , Parkhill, J. , Dougan, G. , and Wren, B.W. (2009). Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biology* 10(9), p. Article numberR102.

Stewart, F. M., and Levin, B.R. (1984). The population biology of bacterial viruses: Why be temperate. *Theoretical Population Biology* 26(1), pp. 93-117.

Sturino, J. M., and Klaenhammer, T.R. (2006). Engineered bacteriophage-defence systems in bioprocessing. *Nature Reviews Microbiology* 4(5), pp. 395-404.

Sulakvelidze, A., Alavidze, Z. and Morris, J.G. (2001). Bacteriophage therapy. *Antimicrobial Agents and Chemotherapy* 45(3), pp. 649-659.

Sulakvelidze, A., and Morris J.G., Jr. (2001). Bacteriophages as therapeutic agents. *Annals of Medicine* 33(8), pp. 507-509.

Summers, W. C. (2001). Bacteriophage therapy. *Annual Review of Microbiology* 55 pp. 437-451.

Sundar, M. M., Nagananda, G.S., Das, A., Bhattacharya, S. and Suryan, S. (2009). Isolation of host-specific bacteriophages from sewage against human pathogens. *Asian Journal of Biotechnology* 1(4), pp. 163-170.

Surawicz, C. M., McFarland, L.V., Elmer, G., and Chinn, J. (1989). Treatment of recurrent *Clostridium difficile* colitis with vancomycin and *Saccharomyces boulardii*. *American Journal of Gastroenterology* 84(10), pp. 1285-1287.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28(10), pp. 2731-2739.

Tartera, C., and Jofre, J. (1987). Bacteriophages active against *Bacteroides fragilis* in sewage-polluted waters. *Applied and Environmental Microbiology* 53(7), pp. 1632-1637.

Taylor, C. P., Tummala, S. , Molrine, D. , Davidson, L. , Farrell, R.J. , Lembo, A. , Hibberd, P.L. , Lowy, I. , and Kelly, C.P. (2008). Open-label, dose escalation phase I study in healthy volunteers to evaluate the safety and pharmacokinetics of a human monoclonal antibody to *Clostridium difficile* toxin A. *Vaccine* 26(27-28), pp. 3404-3409.

Taylor, N. S., Thorne, G. M., and Bartlett, J. G. (1981). Comparison of two toxins produced by *Clostridium difficile*. *Infection and Immunity* 34(8), pp. 1036-1043.

Tedesco, F. J., Barton, R.W., and Alpers, D.H. (1974). Clindamycin associated colitis. A prospective study. *Annals of Internal Medicine* 81(4), pp. 429-433.

Todd, S. J., Moir, A.J.G., Johnson, M.J., and Moir, A (2003). Genes of *Bacillus cereus* and *Bacillus anthracis* encoding proteins of the exosporium. *Journal of Bacteriology* 185(11), pp. 3373-3378.

Tomasz, M. (1995). Mitomycin C: Small, fast and deadly (but very selective). *Chemistry and Biology* 2(9), pp. 575-579.

Urakami, K., Miyata, S., Moriyama, R., Sugimoto, K., and Makino, S. (1999). Germination-specific cortex-lytic enzymes from *Clostridium perfringens* S40 spores: Time of synthesis, precursor structure and regulation of enzymatic activity. *FEMS Microbiology Letters* 173(2), pp. 467-473.

Van der Oost, J., Jore, M.M., Westra, E.R., Lundgren, M., and Brouns, S.J.J. (2009). CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends in Biochemical Sciences* 34(8), pp. 401-407.

Vollmer, W., Joris, B., Charlier, P., and Foster, S. (2008). Bacterial peptidoglycan (murein) hydrolases. *FEMS Microbiology Reviews* 32(2), pp. 259-286.

Vonberg, R.-P., Kuijper, E.J. , Wilcox, M.H. , Barbut, F., Tüll, P. , Gastmeier, P. , Van Den Broek, P.J. , Colville, A. , Coignard, B. , Daha, T. , Debast, S. , Duerden, B.I. , van den Hof, S., van der Kooi, T. , Maarleveld, H.J.H. , Nagy, E. , Notermans, D.W., O'Driscoll, J. , Patel, B. , Stone, S. , and Wiuff, C. (2008). Infection control measures to limit the spread of *Clostridium difficile*. *Clinical Microbiology and Infection* 14(SUPPL. 5), pp. 2-20 .

Waldor, M. K., and Mekalanos, J.J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272(5270), pp. 1910-1913.

Wang, I.-N., Deaton, J. , and Young, R. (2003). Sizing the holin lesion with an endolysin-β-galactosidase fusion. *Journal of Bacteriology* 185(3), pp. 779-787.

Warny, M., Pepin, J. , Fang, A., Killgore, G. , Thompson, A. , Brazier, J. , Frost, E. , and McDonald, L.C. (2005). Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. *Lancet* 366(9491), pp. 1079-1084.

Warth, A. D., and Strominger, J.L. (1969). Structure of the peptidoglycan of bacterial spores: occurrence of the lactam of muramic acid. *Proceedings of the National Academy of Sciences of the United States of America* 64(2), pp. 528-535.

Warth, A. D., and Strominger, J.L. (1972). Structure of the peptidoglycan from spores of *Bacillus subtilis. Biochemistry* 11(8), pp. 1389-1396.

Weaver, J., Tae, J.K. , Raines, K.W. , Cao, G.-L. , Hibbs, S. , Tsai, P. , Baillie, L. , Rosen, G.M. , and Cross, A.S. (2007). Protective role of *Bacillus anthracis* exosporium in macrophage-mediated killing by nitric oxide. *Infection and Immunity* 75(8), pp. 3894-3901.

Wenisch, C., Parschalk, B., Hasenhundl, M., Hirschl, AM., and Graninger, W. (1996). Comparison of vancomycin, teicoplanin, metronidazole, and fusidic acid for the treatment of *Clostridium difficile*-associated diarrhea. *Clinical Infectious Diseases* 22(5), pp. 813-818.

WHO (2002). Guidelines on Prevention and Control of Hospital Associated Infections. Regional Office for SE Asia New Delhi ,WHO.

Wiebe, W. J., and Liston, J. (1968). Isolation and characterization of a marine bacteriophage. *Marine Biology* 1(3), pp. 244-249.

Wilson, K. H. (1983). Efficiency of various bile salt preparations for stimulation of *Clostridium difficile* spore germination. *Journal of Clinical Microbiology* 18 (4), pp. 1017-1019.

Wilson, K. H., Kennedy, M.J., and Fekety, F.R. (1982). Use of sodium taurocholate to enhance spore recovery on a medium selective for *Clostridium difficile*. *Journal of Clinical Microbiology* 15(5), pp. 443-446.

Winn, W. J., Allen, S., Janda, W., Koneman, E., Procop, G., Schreckenberger, P., and Woods, G. (2006). Koneman's Color Atlas and Textbook of Diagnostic Microbiology. Sixth edition ed. Baltimore, MD: Lippincott, Williams and Wilkins.

Wren, M. W. D., Sivapalan, M., Kinson, R., and Shetty, N.P. (2009). Laboratory diagnosis of *Clostridium difficile* infection. An evaluation of tests for faecal toxin, glutamate dehydrogenase, lactoferrin and toxigenic culture in the diagnostic laboratory. *British Journal of Biomedical Science* 66(1), pp. 1-5.

Wu, D. Y., Ugozzoli, L., Pal, B.K., Qian, J., and Wallace, R.B. (1991). The effect of temperature and oligonucleotide primer length on the specificity and efficiency of amplification by the polymerase chain reaction. *DNA and Cell Biology* 10(3), pp. 233-238.

Xiao, Y. , Francke, C., Abee, T., and Wells-Bennik, M.H.J. (2011). Clostridial spore germination versus bacilli: Genome mining and current insights. *Food Microbiology* 28(2), pp. 266-274.

Yoong, P., Schuch, R., Nelson, D., and Fischetti, V.A. (2004). Identification of a broadly active phage lytic enzyme with lethal activity against antibiotic-resistant *Enterococcus faecalis* and *Enterococcus faecium. Journal of Bacteriology* 186(14), pp. 4808-4812.

Young, R., and Bläsi, U. (1995). Holins: Form and function in bacteriophage lysis. *FEMS Microbiology Reviews* 17(1-2), pp. 191-205.

Young, R. (2002). Bacteriophage holins: Deadly diversity. *Journal of Molecular Microbiology and Biotechnology* 4(1), pp. 21-36.

Young, R. Y., Wang, I-N., and Roof, W.D. (2000). Phages will out: strategies of host cell lysis. *Trends in Microbiology* 8(3), pp. 120-128.

Zdanovsky, A. G., and Zdanovskaia, M.V. (2000). Simple and efficient method for heterologous expression of clostridial proteins. *Applied and Environmental Microbiology* 66(8), pp. 3166-3173.

Zhong, Q., and Jin, M. (2009). Nanoscalar Structures of Spray-Dried Zein Microcapsules and in Vitro Release Kinetics of the Encapsulated Lysozyme As Affected by Formulations. *Journal of Agricultural and Food Chemistry* 57(9), pp. 3886-3894.

Zimmer, M., Vukov, N., Scherer, S., and Loessner, M.J. (2002). The murein hydrolase of the bacteriophage φ3626 dual lysis system is active against all tested *Clostridium perfringens* strains. *Applied and Environmental Microbiology* 68(11), pp. 5311-5317.

Appendices

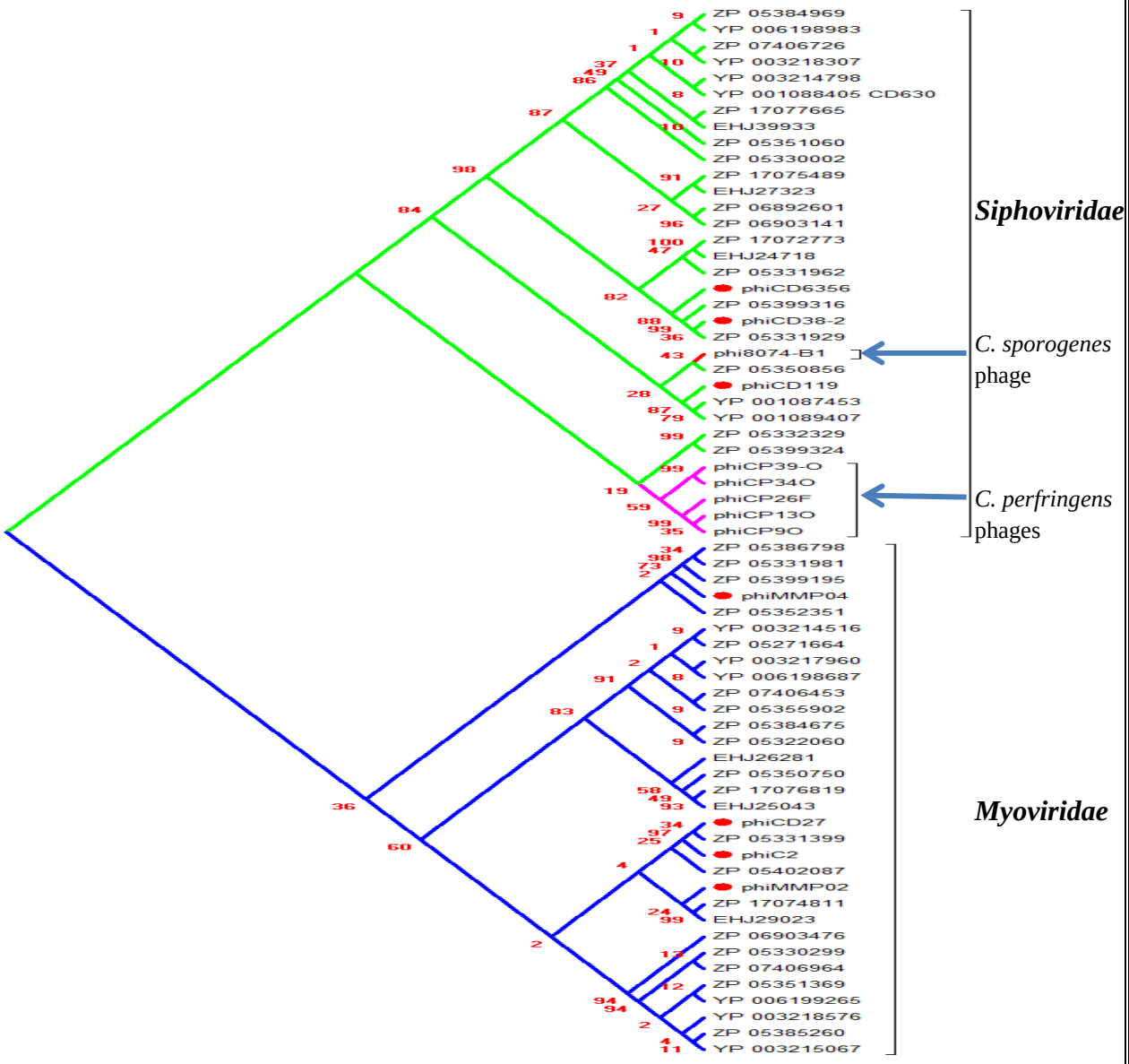
Appendix 1: The *C. difficile* isolates used in this study

Ribotype	Strain	Patient Sex	Patient Age	Origin	Year referred
001	DS1759	M	87	Maidstone	*
	DS1747	F	82	St James Leeds	*
	R8652	*	*	NCTC 11209	*
	DS1750	F	88	St James Leeds	*
	R9399	*	*	*	1996
	R12824	*	*	*	1999
	R17752	*	*	*	2002
027	DS1813	M	80	Hinchingbrooke	*
	DS1801	F	64	Leicester	*
	R20291			Stoke Mandeville	*
	DS1807	F	31	Salford	*
	R24626	*	*	Bath	2007
	R26390	*	*	Stevenage	2008
106	R10459	*	*	Dudley	*
	DS1798	M	94	Poole	*
	DS1787	M	83	Leicester	*
	DS1771	M	88	Bristol Southmead	*
	R28614	*	*	Newport	2009
014	DS1742	M	27	Bristol Frenchay	
	R30061	*	*	Stock on Trent	2010
	R22537	*	*	Surry	2006
	R31755	*	*	Waterford	2011
	R31757	*	*	Waterford	2001
002	DS1748	F	83	Leeds	*
	R25577	*	*	Oldham	2008
	R27039	*	*	Wakefiled	2008
	R31760	*	*	Waterford	2011
	R31763	*	*	Waterford	2001
005	DS1721	M	66	Leicester	*
	R27038	*	*	Margate	2008
	R25028	*	*	*	2007
	R31762	*	*	Waterford	2011
	R31778	*	*	Swansea	2011
012	DS1752	F	81	Bradford	*
	R31771	*	*	Swansea	2011
	R31080	*	*	London	2011
	R30904	*	*	London	2011
	R30851	*	*	London	2011
078	DS1723	M	71	Leicester	*
	R19058	*	*	*	2003
	R31710	*	*	Durham	2011
	R31708	*	*	Cardiff	2011
	R31777	*	*	Swansea	2011
020	DS1724	M	83	Leicester	*
	R31774	*	*	Swansea	2011
	R31550	*	*	London	2011
	R31464	*	*	London	2011
	R31502	*	*	Naas Ireland	2011

Appendix 1(continued): The *C. difficile* isolates used in this study

Ribotype	Strain	Patient Sex	Patient Age	Origin	Year referred
010	DS1684	F	92	Brighton	*
	R31750	*	*	Waterford	2011
	R31722	*	*	London	2011
	R30762	*	*	Sligo Ireland	2010
	R30327	*	*	London	2010
023	DS1665	*	*	*	*
	R15552	*	*	*	*
	R31377	*	*	Wales study	2011
	R31148	*	*	Wales study	2011
	R31314	*	*	Wales study	2011
017	R9557	*	*	*	1996
	R13695	*	*	*	2000
	R18091	*	*	*	2003
	R2139	*	*	*	1985
	R19222	*	*	*	2004
	R22660	*	*	*	*
047	R10542	*	*	*	1997
	R18045	*	*	*	2003
	R25961	*	*	*	*
	R17732	*	*	Leeds	2008
	R10543	*	*	*	*
056	R24498	*	*	*	2007
	DS2008	*	*	*	2008
	R26796	*	*	Wales study	2008
	R31312	*	*	Wales study	2011
	R31229	*	*	Wales study	2011
110	R7771	*	*	*	1994
	R17978	*	*	*	2003
	R30967	*	*	*	*
	R18040	*	*	Birmingham	2011
	R17985	*	*	*	*
046	R19168	*	*	*	2004
	R31583	*	*	London	2011
	R31263	*	*	London	2011
	R30869	*	*	London	2011
	R30870	*	*	London	2011
045	R20408	*	*	*	2004
	R30776	*	*	London	2010
	R28450	*	*	Leeds	2009
	R30072	*	*	Exeter	2010
	R27086	*	*	*	*

*Information not available



Appendix 2: Evolutionary relationship of lysin sequences encoded peptidoglycan hydrolase, N-acetylmuramoyl-L-alanine amidase from *C.difficile* genomes and phages. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The analysis involved 65 amino acid sequences including *C. difficile* phage endolysins (φCD119, φCD2, φCD27, φMMP03, φMMP04, φCD6356 and φCD38-2), *C. perfringens* phage endolysin (φ39-O, φCP13O, φCP26F, φCP34O, φCP9O) and *C. sporogenes* phage (φ8074-B1). All positions containing gaps and missing data were eliminated. There were a total of 180 positions in the final dataset. Evolutionary analyses were conducted in MEGA5

Appendix 3: The Peptide sequences identified in *C.difficile* C_{50} digestion gel (Figure 6.11 and table 6.4) match with spore cortex-lytic enzyme pre-pro-form OS=*Clostridium difficile* (strain 630) , GenBank : CAJ67384.1

Sample	Matched peptides
1	1 MQDGFLTVSI IDATNNRPIQ NAVVNIYSMS NGSQSSSTLY QNLRSNESGQ 51 VTGLVLPAPD VDYSLQPSDV RPYSQYIVEA IADGYETVVI EGTQLLATIE 101 ARQGVPMSPR TRSKRSFSRQ SELIFDIGEH TLYGTYPPKI PESNLKPLPP 151 PTGFVVLDNP VVPEFIVVHD GLPEDSSAPN YWIPFKEYIK NIASSEIYST 201 WPEQTIYANV IAIISFTLNR VFTEWYRNKG YNFTITSTTA YDHKFINNRN 251 LFEPINVVVD AIFNTFIKRP PTSRQPLLAQ YCDGQKSQCP DQMTQWGSKD 301 LGDQGYDYES ILRYFYGDEI VFERAPIVSG VPVSFPGTTL QVGSSGQYVR 351 TIQNQLNAIS NSYPAVPKVI EDGIYGTDTE NAVKIFQGIF GLPQSGVVDF 401 KTWYEISRVY VATTRIASLN PLI
2	1 MQDGFLTVSI IDATNNRPIQ NAVVNIYSMS NGSQSSSTLY QNLRSNESGQ 51 VTGLVLPAPD VDYSLQPSDV RPYSQYIVEA IADGYETVVI EGTQLLATIE 101 ARQGVPMSPR TRSKRSFSRQ SELIFDIGEH TLYGTYPPKI PESNLKPLPP 151 PTGFVVLDNP VVPEFIVVHD GLPEDSSAPN YWIPFKEYIK NIASSEIYST 201 WPEQTIYANV IAIISFTLNR VFTEWYRNKG YNFTITSTTA YDHKFINNRN 251 LFEPINVVVD AIFNTFIKRP PTSRQPLLAQ YCDGQKSQCP DQMTQWGSKD 301 LGDQGYDYES ILRYFYGDEI VFERAPIVSG VPVSFPGTTL QVGSSGQYVR 351 TIQNQLNAIS NSYPAVPKVI EDGIYGTDTE NAVKIFQGIF GLPQSGVVDF 401 KTWYEISRVY VATTRIASLN PLI

Appendix 4: The Peptide sequences identified in *C.perfringens* C50 digestion gel (Figure 6.11 and table 6.4) match with spore cortex-lytic enzyme SleC OS=*C.perfringens* CPE str , GenBank: BAA08081.1

Sample	Matched peptides
3	1 MAVGGLKVQC FSEQRRYIPI DKCKVKITPT GEDGVAVGDT IELYTDDTGS 51 TDTIELDAPP IENSNQPGTI PYSFAEVIVE REGFLPVAVN GVQIYPSRIA 101 LQNVNLPETR GYYRQEEVID IQPNRLVGNF PPKIPEAEEK ELPPPKGTVV 151 LPEPVVPEYI VVHNGRPNDN SVANYKVNYK DYIKNVACCE IFSTWSENTI 201 RANVYAIISF TLNRIYTEWY RGKGKNFDIT NSTAFDHAFS YGRNFYDNIS 251 RIVDEIFSTY MKRFNSKQPL LAQYCDGINV QCPGWMTQWG SKYLGDEGKV 301 PYDILTSFYG DDLELKSAKK VKGSPRSYPG YTLKTGYSGE PVRVIQEQLN 351 AISRAYPLIP KIAVDGKYGP KTREAVKTFQ KIFNLPQTGE VDYATWYKIS 401 DVYVAVTKIA ELRSSVEKKI FYPPTIMDRR ENVPKIIY
4	1 MAVGGLKVQC FSEQRRYIPI DKCKVKITPT GEDGVAVGDT IELYTDDTGS 51 TDTIELDAPP IENSNQPGTI PYSFAEVIVE REGFLPVAVN GVQIYPSRIA 101 LQNVNLPETR GYYRQEEVID IQPNRLVGNF PPKIPEAEEK ELPPPKGTVV 151 LPEPVVPEYI VVHNGRPNDN SVANYKVNYK DYIKNVACCE IFSTWSENTI 201 RANVYAIISF TLNRIYTEWY RGKGKNFDIT NSTAFDHAFS YGRNFYDNIS 251 RIVDEIFSTY MKRFNSKQPL LAQYCDGINV QCPGWMTQWG SKYLGDEGKV 301 PYDILTSFYG DDLELKSAKK VKGSPRSYPG YTLKTGYSGE PVRVIQEQLN 351 AISRAYPLIP KIAVDGKYGP KTREAVKTFQ KIFNLPQTGE VDYATWYKIS 401 DVYVAVTKIA ELRSSVEKKI FYPPTIMDRR ENVPKII
5	1 MAVGGLKVQC FSEQRRYIPI DKCKVKITPT GEDGVAVGDT IELYTDDTGS 51 TDTIELDAPP IENSNQPGTI PYSFAEVIVE REGFLPVAVN GVQIYPSRIA 101 LQNVNLPETR GYYRQEEVID IQPNRLVGNF PPKIPEAEEK ELPPPKGTVV 151 LPEPVVPEYI VVHNGRPNDN SVANYKVNYK DYIKNVACCE IFSTWSENTI 201 RANVYAIISF TLNRIYTEWY RGKGKNFDIT NSTAFDHAFS YGRNFYDNIS 251 RIVDEIFSTY MKRFNSKQPL LAQYCDGINV QCPGWMTQWG SKYLGDEGKV 301 PYDILTSFYG DDLELKSAKK VKGSPRSYPG YTLKTGYSGE PVRVIQEQLN 351 AISRAYPLIP KIAVDGKYGP KTREAVKTFQ KIFNLPQTGE VDYATWYKIS 401 DVYVAVTKIA ELRSSVEKKI FYPPTIMDRR ENVPKIIY