Imperial College London



A lesson in bacterial variability:

The *C. difficile* cell wall protein CwpV

Catherine B. Reynolds

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Cellular and Molecular Biology

Imperial College London

January 2011

Declaration

I declare that this thesis entitled “A lesson in bacterial variability: The *C. difficile* cell wall protein CwpV” is original and has not been submitted for a degree or any other qualification at any institution.

All data presented within this thesis are the result of my own ideas and work, with contributions from colleagues appropriately acknowledged.



Signed……

Date……28th January 2011……

Abstract

Clostridium difficile is the main cause of antibiotic-associated diarrhea, leading to significant morbidity and mortality, and putting considerable economic pressure on healthcare systems. Current knowledge of the molecular basis of pathogenesis is limited primarily to the activities and regulation of two major toxins. In contrast, little is known of the mechanisms used to colonise the enteric system. *C. difficile* expresses a proteinaceous array on its cell surface known as the S-layer, consisting primarily of SlpA and a family of homologues, the cell wall protein (CWP) family. CwpV is the largest member of this family. CwpV is expressed in a phase-variable manner controlled by an invertible DNA switch, the *cwpV* switch. The novel mechanism controlling this phase variation has been charaterised using enzymatic reporter assays. A site-specific recombinase (RecV) catalyzing the inversion of the *cwpV* switch has been identified. Knocking out this recombinase has enabled isolation of *cwpV* switch locked ON and locked OFF strains of *C. difficile,* indicating that *cwpV* switch orientation is the primary determinant of CwpV expression. CwpV is post-translationally cleaved and expressed on the cell surface as two proteins that form a stable complex, with one subunit responsible for the non-covalent cell wall anchoring of the other large repetitive subunit. Due to the significant heterogeneity of *C. difficile* strains the characteristics of CwpV across a panel of strains were investigated. The *cwpV* switch and *recV* are conserved across diverse strains and all strains tested express CwpV in a phase variable manner. The N-terminus of CwpV is well conserved, however the C-terminal repetitive domain of CwpV varies markedly. Five different types have been identified and shown to be antigenically distinct. All types of CwpV repeats promote aggregation of *C. difficile* cells, which may be an important function during infection. These findings suggest a complex evolutionary history for CwpV.

Acknowledgements

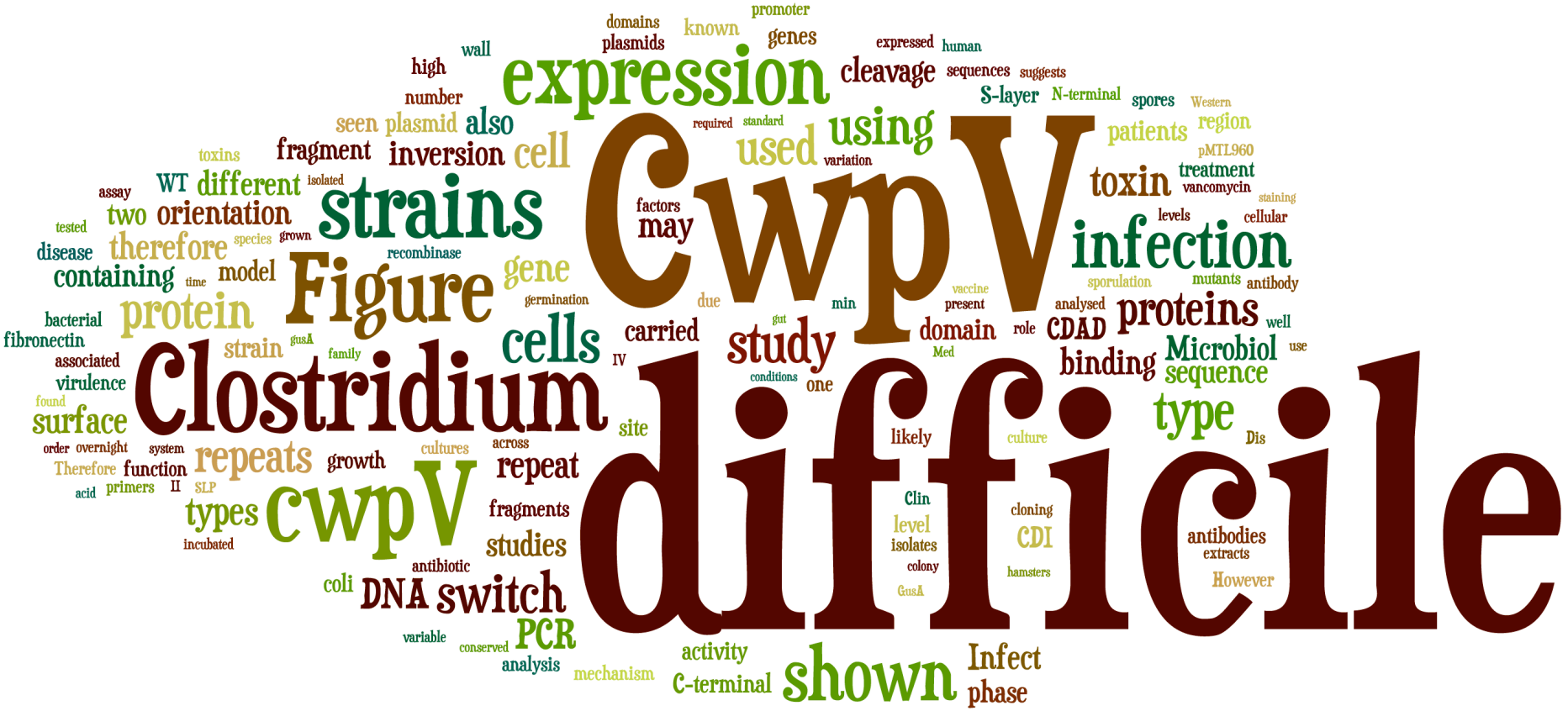
First and foremost, I would like to thank my supervisor Neil Fairweather who has been constantly supportive, a source of inspiration and a role model. I have been so lucky with my colleagues from Team Fairweather, especially Robert Fagan, who taught me (nearly) everything I needed to know for the lab, David Albesa-Jove who made me laugh, Lucia de la Riva-Perez who's positivity and optimism helped me maintain mine, Jenny Emerson for sharing such an interesting project with me, Allie Shaw for her calming influence and Zoe Seager for her sense of fun. I must also acknowledge all the students who worked with me on the project: Allie Shaw, Natalie Welsh, Stina Linden, Marcin Dembek and Tom Bell; I found teaching all of you made me even more enthusiastic about my PhD project and many of the most interesting results came from the work we discussed and did together. Thank you to the whole CMMI for providing a stimulating community to be part of.

Thank you to the Wellcome Trust for their financial support of my PhD and the preceding MRes, which has been an excellent training program to have had the privilege to be part of.

Thank you to my friends Kieran, Henry, Andrej, Alex, Ana, Dan and Graham who made lunch time, coffee time and even triathlon training time fun! You have been a hugely valuable and positive influence on my years at Imperial. I hope we all stay in touch, and I wish you all the best for your futures.

My house mates have been like my family in London, I would therefore like to thank all of mine from over the years: Alex, Alex, Neil, Charlie, Miranda, Rob, Thomas and Abdou who helped keep me sane and enjoy the time spent in our London homes. There are too many other friends to name you all, but I think you know who you are - Monkeys, Magdalen legends, Bedford girls, Knitting club and Nick. Thank you so much for the fun times and I look forward to many more.

Finally, I must thank my real family for their support; I wouldn't be where I am without you. Special thanks must go to the parents. I have the best mum in the world, thank you for always being so caring and thoughtful, but remember to chase your own dreams. Dad, it has always been inspiring to be a scientist's daughter, thank you for all our big discussions and I look forward to many more.


difficile
Clostridium
cwpV
expression
Figure
strains
infection
protein
cells
study
shown
cwpV
repeats
switch
PCR
DNA
toxin
proteins
Microbiol
binding
sequence
type
surface
repeat
studies
Infect
phase
activity
C-terminal
mechanism
analysis
variable
conserved
coli
antibiotic
incubated
types
fragments
level
isolates
colony
GusA
Clin
CDI
cloning
hamsters
antibodies
extracts
However
likely
culture
Dis
site
across
cultures
growth
function
Therefore
primers
acid
II
SLP
order
overnight
system
found
surface
virulence
associated
fibronectin
bacterial
strain
gusA
family
protein
containing
time
model
disease
therefore
species
grown
tested
two
orientation
WT
different
isolated
assay
toxins
fragment
inversion
cell
seen
plasmid
also
number
high
wall
plasmids
domains
known
promoter
genes
expressed
human
cleavage
sequences
suggests
S-layer
N-terminal
spores
Western
region
patients
pMTL960
treatment
vancomycin
staining
cellular
levels
factors
may
used
using
variation
required
standard
recombinase
gene
germination
due
min
vaccine
present
analysed
sporulation
mutants
antibody
well
use
role
CDAD
domain
Med
carried
IV
gut
conditions
one
Infect
phase
shown
activity
C-terminal
mechanism
analysis
variable
conserved
coli
antibiotic
DNA
switch
PCR
incubated
types
fragments
studies
repeat
across
cultures
growth
function
repeats
Therefore
primers
acid
II
SLP
order
overnight
system
found
surface
virulence
associated
fibronectin
bacterial
strain
gusA
family
protein
containing
time
model
disease
therefore
species
grown
tested
two
orientation
WT
different
isolated
assay
toxins
fragment
inversion
cell
seen
plasmid
also
number
high
wall
plasmids
domains
known
promoter
genes
expressed
human
cleavage
sequences
suggests
S-layer
N-terminal
spores
Western
region
patients
pMTL960
treatment
vancomycin
staining
cellular
levels
factors
may
used
using
variation
required
standard
recombinase
gene
germination
due
min
vaccine
present
analysed
sporulation
mutants
antibody
well
use
role
CDAD
domain
Med
carried
IV
gut
conditions
one
likely
culture
Dis
site
Clin
level
isolates
colony
GusA
CDI
cloning
hamsters
antibodies
extracts
However

Contents

1 INTRODUCTION 13

1.1 History of *C. difficile* infection 13

1.2 Current understanding of *C. difficile* infection 16

1.3 *Clostridium difficile* diagnosis 22

1.4 Prevention of *C. difficile* infection 26

1.5 Treatment of CDAD 35

1.6 Economics of *C. difficile* infection 47

1.7 *C. difficile* evolution and population biology 51

1.8 Studying the molecular basis of *C. difficile* pathogenesis 65

1.9 *C. difficile* virulence factors 72

1.10 The *C. difficile* cell wall protein CwpV 88

1.11 Project Aims 95

2 MATERIALS AND METHODS 96

2.1 Bacterial strains and culture 96

2.2 Bioinformatics and computer programs 99

2.3 DNA manipulation 101

2.4 Protein manipulation 105

2.5 Immunodetection 111

2.6 Reporter Assays 113

2.7 Phenotypic assays 115

3 PRODUCTION OF RECOMBINANT PLASMIDS 132

3.1 Reporter Plasmids 132

3.2 Recombinase expression in *E. coli* 136

3.3 RecV (CD1167) expression in *C. difficile* 137

3.4 Expression of CwpV repeat fragments in *E.coli* 138

3.5 Expression of CwpV in *C. difficile* 139

4 REGULATION OF CWPV EXPRESSION 142

4.1 Introduction 142

4.2 Results 144

4.3 Discussion 172

5 CwpV POST-TRANSLATIONAL PROCESSING AND EXPRESSION IN THE *C. DIFFICILE* S-LAYER: CONSERVATION AND DIVERSITY ACROSS STRAINS 176

5.1 Introduction 176

5.2 Results 179

5.3 Discussion 206

6 INVESTIGATING THE FUNCTION OF CwpV 213

6.1 Introduction **213**

6.2 Results **218**

6.3 Discussion **248**

7 DISCUSSION **254**

7.1 CwpV phase variation **254**

7.2 *C. difficile* strain diversity **255**

7.3 CwpV and the *C. difficile* S-layer **257**

7.4 CwpV and *C. difficile* pathogenicity **258**

7.5 Overall conclusions **259**

8 REFERENCES **260**

Figures

Figure 1.1 Endoscopic visualization of pseudomembranous colitis. 14

Figure 1.2 Diversity of microbial communities in an *in vitro* gut model before and after treatment with antimicrobials. 41

Figure 1.3 REA types of *C. difficile* isolates recovered from a worldwide clinical trial. 60

Figure 1.4 Diagrammatic representation of the *C. difficile* pathogenicity locus (PaLoc). 73

Figure 1.5 Model of the cell wall of *C. difficile*. 83

Figure 1.6 Diagrammatic representation of the *C. difficile* CWP family. 86

Figure 1.7 The CwpV protein in *C. difficile* strain 630. 89

Figure 1.8 Sequence variation of CwpV across *C. difficile* strains. 90

Figure 1.9 Detection of CwpV phase variable expression 91

Figure 1.10 The *cwpV* switch. 93

Figure 3.1 pCBR023 134

Figure 3.2 pCBR034 135

Figure 3.3 pCBR045 136

Figure 3.4 pCBR113 137

Figure 3.5 pCBR068 138

Figure 3.6 pCBR044 139

Figure 3.7 pCBR080 140

Figure 4.1 Surface expression of CwpV 146

Figure 4.2 Expression of CwpV at the cellular level. 147

Figure 4.3 Development of a robust *C. difficile* reporter assay. 150

Figure 4.4 Analysis of transcription from the CwpV promoter in *C. difficile* using a GusA reporter. 154

Figure 4.5 RNA structure and free energy predictions of the *cwpV* switch. 156

Figure 4.6 Testing the model of the *cwpV* switch mechanism. 158

Figure 4.7 Analysis of transcription from the CwpV promoter in *Clostridium perfringens.* 160

Figure 4.8 DNA inversion is specific to *C. difficile* and does not occur in C. *perfringens* or *E. coli*..... 162

Figure 4.9 Identification of a C.*difficile* site-specific recombinase mediating DNA inversion of the *cwpV* switch. 164

Figure 4.10 Isolation of *ΔrecV* mutants. 167

Figure 4.11 CwpV expression in Δ*recV* mutants. 169

Figure 4.12 CwpV expression at the cellular level in Δ*recV* mutants. 171

Figure 5.1 The CwpV protein from 630. 180

Figure 5.2 Determination of the 630 CwpV cleavage site. 182

Figure 5.3 The CwpV N-terminal fragment mediates cell wall binding. 183

Figure 5.4 The two CwpV fragments form a complex. 184

Figure 5.5 Expression and purification of StrepII-tagged CwpV from *C. difficile* surface extracts. 186

Figure 5.6 Production of CwpV Cleavage Domain Mutants. 187

Figure 5.7 Analysis of CwpV Cleavage Domain Mutants. 189

Figure 5.8 Overall Model for CwpV processing and expression. 191

Figure 5.9 CwpV C-terminal repeat types II-V. 194

Figure 5.10 Different CwpV types have distinct amino acid sequences..... 196

Figure 5.11 Raising CwpV repeat-specific antibodies. 198

Figure 5.12 Different CwpV repeat types are antigenically distinct. 200

Figure 5.13 There is a conserved post-translational cleavage site for all CwpV proteins..... 202

Figure 5.14 All CwpV proteins form a complex of N- and C-terminal fragments. 203

Figure 5.15 Phase variable regulation of CwpV expression by *C. difficile* strains of all CwpV repeat type. 205

Figure 6.1 Growth curves of *C. difficile* strains..... 219

Figure 6.2 *C. difficile* strains exhibit colony morphology differences. 221

Figure 6.3 *C. difficile* strains exhibit colony morphology differences. 222

Figure 6.4 Microscopic images of the edges of *C. difficile* colony growth. .. 225

Figure 6.5 CwpV is an Aggregation Promoting Factor. 227

Figure 6.6 CwpV expression does not affect FliC expression or swimming motility. 229

Figure 6.7 Screening for interactions between CwpV repeats and extracellular matrix proteins by ELISA. 231

Figure 6.8 Confirmation of interaction between CwpV repeats types I and IV with fibronectin..... 232

Figure 6.9 Comparison of cellular and plasma fibronectin interactions with CwpV repeats. 233

Figure 6.10 Cellular fibronectin binding activity of 2 and 9 CwpV type I repeat domains. 234

Figure 6.11 *C. difficile* binding to cellular fibronectin. 236

Figure 6.12 Effect of CwpV expression on *C. difficile* binding to Caco-2 BBE1 cells. 239

Figure 6.13 CwpV and *C. difficile* sporulation. 243

Figure 6.14 CwpV and *C. difficile* spore germination. 247

Tables

Table 2.1 Description of recombinant antigens used to raise antibodies..... 110

Table 2.2 Primers used in this study 121

Table 2.3 Plasmids used in this study 127

Table 5.1 CwpV characteristics of *C. difficile* strains examined 212

Table 6.1 *C. difficile* strains used in Chapter 6 218

1 Introduction

1.1 History of *C. difficile* infection

The spore-forming rod-shaped bacterium that we now know as *C. difficile* was first described in 1935. It was isolated from the stool of infants and was named *Bacillus difficilis* due to the difficulties that had to be overcome to culture the organism (Hall & O'Toole, 1935). It was shown that this bacterium could produce a potent toxin (Snyder, 1937), however no consistent link with pathogenic infection could be found. Therefore for a number of years this organism was not intensively studied, but it was shown to be a member of the Clostridial species, renamed *Clostridium difficile* and some small studies were carried out (Smith & King, 1962, Hafiz & Oakley, 1976).

Pseudomembraneous lesions in the colon, known as pseudomembranous colitis (PMC) were first reported in 1893 and there are a number of other reports of PMC in the pre-antibiotic era. PMC refers to visible changes on the inner surface of the lining of the colon. Characteristically, the colon has patches of inflammatory membrane that consist of red and white blood cells, fibrin and bacteria. Images of PMC visualized by endoscopy are shown in Figure 1.1. Once the use of antibiotics became commonplace so did the reports of PMC. *Staphylococcus aureus* was isolated from stool samples and became the accepted pathogen thought to be responsible for PMC on the basis of animal studies, consistent isolation from patients with PMC and already being recognized as an important nosocomial pathogen. Oral

vancomycin became the standard treatment for PMC and clinicians had substantial success with it (Bartlett, 2009a).

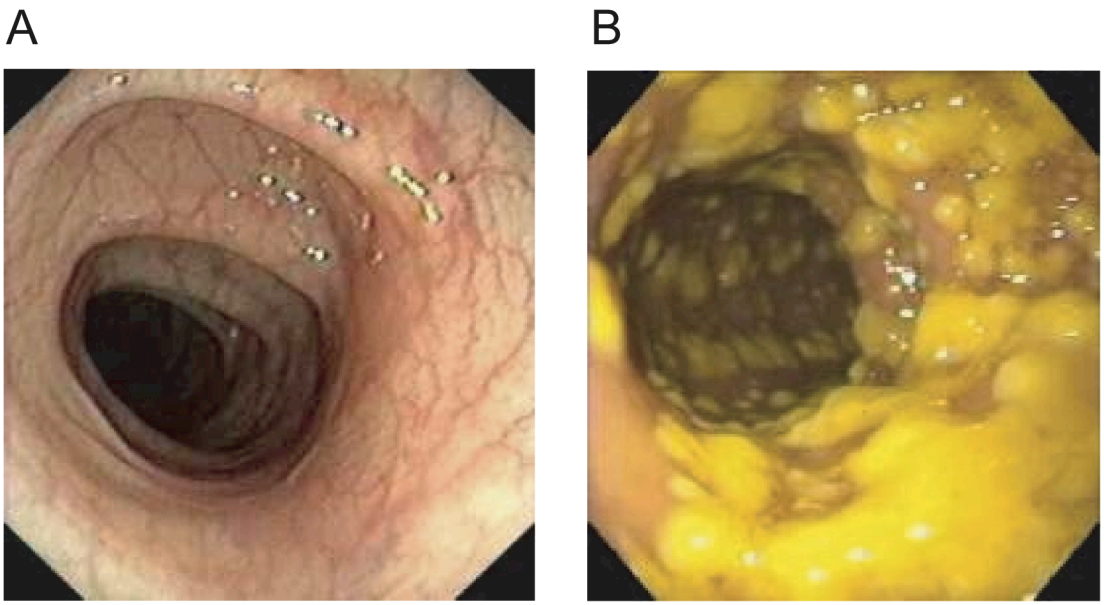


Figure 1.1 Endoscopic visualization of pseudomembranous colitis.

A. Image of a healthy colon. **B.** Image from a CDAD patient. Classic pseudomembranes are visible as raised yellow plaques, which range from 2-10 mm in diameter and are scattered over the colorectal mucosa. Adapted from (Yassin, 2009).

It was not until the early 1970s that this disease state came under closer scrutiny. Clindamycin introduction as a treatment for anaerobic bacterial infections was accompanied by a significant number of patients with severe diarrhoea. On behalf of the company responsible for the production of clindamycin an in depth study into this unfortunate side effect was carried out. 21% of patients receiving clindamycin developed diarrhoea and when examined by endoscopy half of these patients had PMC. The fact that 10% of patients receiving clindamycin developed PMC meant that the FDA required this serious side effect of clindamycin to be reported to all physicians using the drug and PMC became known as clindamycin colitis. Importantly in this study *S. aureus* could not be isolated from stool samples, indicating that PMC

is not caused by this pathogen (Tedesco et al., 1974). Therefore the hunt for the pathogen responsible began.

It was known that antibiotic use in rodents often caused a lethal complication characterised by diarrhoea known as wet tail followed by death a few days after challenge with necropsy showing a large, dilated hemorrhagic cecitis. This process was therefore studied as a model for clindamycin colitis in humans. One of the first clues that these symptoms were caused by a gram positive bacterial pathogen was that challenge with clindamycin followed by oral vancomycin provided protection (Bartlett *et al.*, 1977a). Then it was shown that the symptoms were caused by a toxin, which could be transferred from an infected hamster to a healthy hamster causing the expected symptoms. This toxin could be neutralized by *Clostridium sordellii* antitoxin, implicating a member on the *Clostridium* family as the pathogen at work. Subsequent screening of Clostridial species present in these animals led to the identification of *C. difficile* as the causative agent of antibiotic-associated disease in hamsters (Bartlett *et al.*, 1977b). Following this work in the hamster it could be shown the PMC in humans was also caused by *C. difficile* (Bartlett et al., 1978).

1.2 Current understanding of *C. difficile* infection

1.2.1 Nosocomial infection

C. difficile is the main cause of infectious diarrhoea that develops in patients after hospitalization and antibiotic treatment. The association with antibiotic usage is due the interruption of the normal intestinal flora by the broad-spectrum activity of the antibiotic. Antibiotic associated diarrhoea occurs in 5-35% of patients who take broad-spectrum antibiotics (Wistrom et al., 2001). It is unclear why *C. difficile* is particularly effective at colonising this niche compared to other organisms, but what is clear is that *C. difficile* is the most common pathogen responsible and it is estimated that 20-30% of cases of antibiotic-associated diarrhoea are caused by *C. difficile* (McFarland, 2008). In England in 2008 there were 36,095 reported cases of *C. difficile* infection (NHS, 2010).

C. difficile infection has been linked to use of almost every known antibiotic, with fluoroquinolones indicated as particularly predisposing (Delaney, 2007). The particularly high association between antibiotic usage and CDAD when patients are in hospital is probably due to contamination of the hospital environment with infectious spores. Contamination of the hands of hospital personnel and the ability of spores to persist for up to 20 weeks in the environment after seeding was first shown in 1981 (Fekety et al., 1981, Kim et al., 1981) and since then many studies have indicated high levels of contamination of the hospital environment with infectious *C. difficile* spores (Weber et al., 2010).

The clinical outcomes of *C. difficile* infection (CDI) range from asymptomatic colonisation to mild diarrhoea and more severe disease syndromes, including abdominal pain, fever and leukocytosis. Fulminant or severe complicated CDI is characterised by inflammatory lesions and the formation of pseudomembranes in the colon (PMC). In the most severe cases toxic megacolon or bowel perforation can then lead to septic shock and death. The mortality of CDI infection has recently been estimated to be approximately 6% within three months of diagnosis (Karas et al., 2010). In England in 2008 *C. difficile* was stated to be the underlying cause of death in 2,298 death certificates (Carter, 2009), equating to 6.4% of reported *C. difficile* cases. The large range of clinical symptoms caused by CDI is collectively referred to as *C. difficile* associated disease (CDAD).

1.2.2 Community-acquired infection

Although *C. difficile* is often referred to as a nosocomial pathogen infections do also occur in the community. As expected the most clearly associated risk factor is use of antibiotics, with other risk factors including gastrointestinal disorders and recent hospitalisation (Pituch, 2009). The risk factors generally overlap with those for nosocomial infection (Kutty et al., 2010), although there may be some differences in predisposing factors, for example one study showed that community-acquired CDI occurrs on average in younger patients than nosocomial infection (Naggie et al., 2010). As yet there has not been much work focused on community acquired CDI, and this is an area where future research will shed light on many of the currently unknown factors such as sources of infection (Otten et al., 2010, Wilcox et al., 2008).

1.2.3 Infection of infants

As previously mentioned *C. difficile* was first isolated from newborn infants (Hall & O'Toole, 1935) and it is now accepted that a high proportion (estimates of 30-40%) of infants in their first year are colonised with *C. difficile* (Jangi & Lamont, 2010, Bolton *et al.*, 1984). Despite the presence of toxin produced by *C. difficile* infants do not develop CDAD (Kotloff *et al.*, 1988, Bolton et al., 1984). The mechanisms behind this tolerance are not yet fully understood. Maternal antibodies acquired through breast-feeding may play a role in protection from toxin action (Kim et al., 1984). During months 12-24 the majority of infants lose *C. difficile* as a member of their intestinal flora and during this time they are developing a more adult intestinal flora composition (Palmer et al., 2007). Interestingly these findings of asymptomatic infant colonisation and subsequent clearance of infection are mimicked in hamsters (Rolfe & Iaconis, 1983, Kim *et al.*, 1987) and dogs (Perrin *et al.*, 1993a). Studies in rabbits have also shown neonatal tolerance of *C. difficile* toxin, and this has been linked this to low expression of toxin receptor in neonatal intestine compared to adult intestine (Eglow et al., 1992).

It is not yet clear how the immune system of infants responds to CDI and how this affects their subsequent susceptibility to CDI and CDAD. It has been shown that the majority (60-70%) of children and adults do have antibodies to *C. difficile* toxins (Viscidi et al., 1983), which may be the result of infantile colonisation, and it is known that such antibodies reduce the risk of CDAD (Kyne et al., 2000). Therefore infantile colonisation may lead to a reduction in the risk of developing CDAD later in life, but this has not been shown directly.

1.2.4 Asymptomatic infection

C. difficile infection is obviously a pre-requisite to the development of CDAD, however it has become clear that adults are commonly infected with *C. difficile* but are asymptomatic, therefore *C. difficile* infection alone does not mean that a patient will develop CDAD. As with CDAD, the prevalence of asymptomatic CDI has been shown to be associated with antibiotic usage (Walker *et al.*, 1993, Ryan *et al.*, 2010), although this association is not always seen (Rivera & Woods, 2003). Low levels of anti-toxin A IgG can predispose to symptomatic infection (Kyne et al., 2000) and development of CDAD may be linked to a reduction in IgA producing cells and macrophages in the gut (Johal et al., 2004), however more research in this area is required to fully understand why some patients develop CDAD and some do not.

Asymptomatic carriers are possible sources of *C. difficile* infection, and there is evidence that in the nosocomial environment this is a major source of strains that are transmitted to non-colonised patients (Clabots et al., 1992). For this reason the feasibility of treating asymptomatic carriers with metranidozole or vancomycin to eliminate this source of infection has been investigated. However, metranidazole was ineffective and vancomycin was only temporally effective with a subsequent higher rate of *C. difficile* carriage than prior to vancomycin treatment (Johnson et al., 1992). This reflects the fact that these treatments are themselves broad-spectrum antibiotics and therefore show some propensity to induce CDI. In the future it may be possible to treat asymptomatic *C. difficile* carriers to eliminate these sources

of *C. difficile* from the nosocomial environment but this is not recommended with the treatments currently available.

1.2.5 Animal infection

As previously mentioned the link between *C. difficile* and antibiotic-associated diarrhoea was first discovered in hamsters (Bartlett et al., 1977b) and since this time *C. difficile* has been shown to cause infection in a large number of different animal species. Gnotobiotic mice (those without an intestinal flora) were shown to be susceptible to *C. difficile* infection in 1980, and *C. difficile* was later shown to be present in normal laboratory mouse colonies, where use of ampicillin increased the prevalence of *C. difficile* infection detected (Itoh et al., 1986). In 1981 antibiotic usage in guinea pigs (Rehg & Pakes, 1981) and rabbits (Rehg & Lu, 1981) was shown to induce CDAD.

Some attention therefore was turned from laboratory animals to domestic animals, as these are a potential source of *C. difficile* infection for humans. Dogs, cats and birds were shown to be transiently infected with both toxgenic and non-toxigenic *C. difficile* strains, and there was limited association with antibiotic usage (Borriello et al., 1983). Subsequent studies have confirmed the significant prevalence of *C. difficile* in cats and dogs (Berry & Levett, 1986, Riley *et al.*, 1991, Lefebvre *et al.*, 2006a, Lefebvre *et al.*, 2006b, Lefebvre & Weese, 2009) and the role as a potential source for human infection has been highlighted although not directly shown. *C. difficile* also causes colitis in horses, and has a strong association with antibiotic usage (Perrin *et al.*, 1993b, Ruby *et al.*, 2009). A diverse set of strains have also been isolated

from healthy horses suggesting that *C. difficile* is a common member of the gut flora in horses (Ossiprandi et al., 2009).

The third major group of animals that has been assessed for *C. difficile* infection are animals used in food production, again due to the importance as a possible source of infection transmission to humans. *C. difficile* was first isolated from pigs in 1983 (Jones & Hunter, 1983) and has more recently been shown to be an important cause of diarrhoea in piglets (Yaeger et al., 2002) and is considered to be the most important cause of diarrhoea in piglets that is currently uncontrolled (Songer & Anderson, 2006). Recently one ribotype of *C. difficile* (078) commonly found in pigs has emerged in humans and zoonotic transfer is therefore suggested, but has not been proven (Goorhuis *et al.*, 2008a).

Calves have been shown to be colonised with *C. difficile* although it appears it does not cause disease in calves (Rodriguez-Palacios et al., 2007). The majority of strains isolated from calves are also known to cause human disease (Rodriguez-Palacios et al., 2006) so the possibility of *C. difficile* having transferred between calves and humans is evident. Farm-raised deer have also been shown to have a significant level of *C. difficile* colonisation (French et al., 2010). One study has shown the presence of a diverse set of *C. difficile* strains at poultry farms (Zidaric et al., 2008).

Zoo animals are also at risk of developing *C. difficile* disease as exemplified by an outbreak in elephants, thought to have been triggered by diet-related alteration of the normal gut flora (Bojesen et al., 2006).

From these studies it is clear that *C. difficile* can colonise a diverse set of mammals and birds, therefore zoonotic transfer of infection is a real possibility, although transfer has not yet been shown to occur directly.

1.3 *Clostridium difficile* diagnosis

Accurate diagnosis early in the disease course is important to the successful management of CDI. In the clinical setting this will ensure that patients receive timely, appropriate treatment and that correct infection control measures are put in place to reduce the chances of nosocomial transmission of infection. A false-positive test result may lead to unnecessary treatment and isolation, while the true cause of the patient's diarrhoea may also not be further investigated. In hospitals where *C. difficile* infection patients are isolated together in large wards, due to lack of availability of single-room isolation facilities, a false-positive result could also lead to a *C. difficile*-free patient being at increased risk of acquiring infection from patients who really do have *C. difficile* infections. Conversely, false-negative results may lead to delays in patients receiving appropriate treatment. In many cases over-treatment with empirical antibiotics may occur and this can further exacerbate the CDAD. As these patients would not be isolated the risk of spreading infection are also increased.

Suspicion of CDI, is often based on diarrhoea that has a typical foul-smelling odour (Bartlett & Gerding, 2008). Odour assessment for the likelihood of CDI might be useful when assessing patients with diarrhoea to determine priority for isolation in settings where isolation capacity is limited, especially if rapid laboratory diagnosis is not available (Wilcox, 2007a). Laboratory testing for *C. difficile* is recommended for adults and children over 1 year of age who have otherwise unexplained diarrhoea associated with antibiotic use. The most sensitive assay available for diagnosis of CDAD is the cytotoxin assay or cell cytotoxicity assay (CYT), which was first described in 1978 shortly after the discovery of the association between *C. difficile* and clindamycin colitis (Chang et al., 1978). This assay tests the supernatants of patient faeces for toxic activity on cell monolayers. Toxicity attributable to *C. difficile* toxins is determined by assessing the ability of anti-*C. difficile* toxin antibodies to neutralize the cellular toxicity. *C. difficile* toxins can be detected at picogram levels using this assay. This assay is commonly used as the gold standard for comparing alternative toxin detection assays to. The major disadvantages of the CYT are that it is technically demanding and has a relatively long turnaround time of 24-48 hours. Both the sensitivity and specificity of CYT are very high (98-99%), therefore if this test is carried out additional tests provide little or no new information from the point of view of basic diagnosis, and are not usually warranted.

Cytotoxigenic culture (CYTGC) is where testing is performed on supernatants from a bacterial culture inoculated from faeces, rather than from the faecal sample itself (Bouza et al., 2001). The culturing step adds further time to the

assay, 72 hours being a typical wait for results. A further caveat to CYTGC for the clinician is that it is not specific for *in vivo* production of toxins. If a toxigenic *C. difficile* strain can be cultured it may produce detectable toxin *in vitro* even though it is not producing toxin nor causing CDAD in that particular patient. Therefore this assay is seldom used for routine clinical diagnosis of CDI. However, CYTGC is necessary for downstream determination of antibiotic sensitivity and molecular typing of organisms, and therefore is of high importance to epidemiological studies.

Enzyme immunoassays (EIAs) have been developed to detect *C. difficile* toxins and in the UK there are currently six commercially available EIAs that detect toxins A and B. There are also three commercial lateral flow/immunochromatography assays that also use anti-toxin antibodies to detect toxins A and B (Eastwood et al., 2009). These toxin detection kits are attractive because they remove the need for laboratories to maintain the cell lines necessary for CYT testing. They are also technically easier to carry out, and most importantly give results much more quickly. Typically EIAs take around 2 hours, and immunochromatography takes less than an hour. Although tests were originally designed to detect either toxin A or toxin B, the kits currently available detect both toxins as it is now recognized that toxin A-negative, toxin B-positive strains are virulent and therefore clinically important (Drudy *et al.*, 2006a). In addition to these commercial toxin-detecting assays there is a PCR-based assay available that detects the presence of the toxin B gene. This does not give an indication of whether or not there is a high level of toxin being expressed *in vivo*, but does indicate presence of *C. difficile*. The

final test that is currently commercially available is an EIA that detects the *C. difficile* enzyme glutamate dehydrogenase (GDH). Again this indicates whether or not the organism is present but does not determine toxin expression level.

A recent study was carried out to compare all these rapid tests against CYT and CYTGC for sensitivity and specificity (Eastwood et al., 2009). The main finding was that all rapid tests do not have high enough sensitivity, thereby generating an unacceptably high number of false negatives to allow them to be used as the only test for CDI with sufficient confidence. The specificity of these tests is much better, however some false positives are generated. As previously stated, both false-positive and false-negative results have serious implications for patient care. In this comparative study the PCR assay and GDH assay showed the best sensitivity, but less specificity than the rapid toxin detection kits. This has led to proposals that the PCR and/or GDH assay could be used as a first round of testing, or negative screening tool to give an initial indication of whether *C. difficile* is likely to be present. CYT or CYTGC could then be carried out as a second step of testing on PCR/GDH-positive samples to identify patients with toxins in their faeces (Eastwood et al., 2009, Wren *et al.*, 2009a, Reller *et al.*, 2010). The use of an additional test for lactoferrin, which is a marker of intestinal inflammation has also been suggested to aid identification of patients who are asymptomatically colonised and do not have intestinal inflammation but would test positive for *C. difficile* by PCR or GDH assay (Wren *et al.*, 2009b). Asympotomatically colonised patients may then be isolated as an infection control measure but not treated

for CDAD.

Multiple step testing is likely to increase laboratory costs, but these costs might be offset by reduced total healthcare costs of managing *C. difficile* infection if diagnosis is more accurate. The length of time until a final result is known would be longer than a single test with such a protocol, however if the interim results were made available and used to guide clinicians while waiting for CYT or CYTGC results the overall treatment and isolation outcomes may be improved. Further work is required to assess which assays would provide the optimal multi-step laboratory diagnosis of *C. difficile* infection, with recent studies reporting high levels of sensitivity and specificity for multi-step procedures (Sharp et al., 2009, Swindells et al., 2010). Development of new diagnostic kits is an area of active research and new technology and/or biological targets may improve the options available for diagnosis of *C. difficile* infection.

1.4 Prevention of *C. difficile* infection

Infection control measures to prevent CDI can be put into two groups:

1. Those that attempt to prevent patient contact with *C. difficile* spores.
2. Those that reduce the risk of patients becoming colonised with *C. difficile* upon contact with infectious spores.

1.4.1 Preventing contact with spores

In the hospital environment a number of measures have been tested for their ability to reduce *C. difficile* transmission. The use of gloves and gowns by

healthcare workers are effective methods for reducing the spread of *C. difficile* spores between patients (Johnson et al., 1990). If gloves are not used hand hygiene is critical to preventing the spread of spores, although alcohol hand rubs do not remove or kill *C. difficile* spores, so hand washing with soap and water is preferred (Jabbar et al., 2010). Studies have shown that another effective way to reduce transmission of *C. difficile* is by using disposable thermometers (Brooks et al., 1998, Jernigan et al., 1998). It has been shown that contamination of rooms is closely correlated with the *C. difficile* status of the patient. Patients with CDAD have the highest contamination rate, followed by asymptomatic CDI patients, and as would be expected the lowest contamination level is seen in rooms of patients with no CDI (McFarland et al., 1989). Therefore isolation of patients with CDI in private rooms can reduce contamination of the general environment by these patients, thereby reducing the risk of transmission of infection to other patients. The hands of health care workers become contaminated in proportion to the contamination of the environment (Samore et al., 1996), therefore by having patients with CDI isolated and encouraging the changing of gloves and gowns between rooms and patients the risk of transmission can be reduced. Reducing contamination of the environment by cleaning is also important. However, traditional cleaning products are not sporicidal so are not the most effective means for reducing contamination levels. Hypochlorite solution is sporicidal and has been shown to significantly reduce spore contamination of the environment and therefore *C. difficile* infection rates (Mayfield et al., 2000, Wilcox et al., 2003). Ward overcrowding has been shown to correlate with increased CDI rates, likely

due to sub-optimal implementation of the infection control measures discussed (Kaier et al., 2010).

For community associated CDI the sources of infection are less well understood. A recent study developed a model for transmission of infection outside of the hospital setting. Possible sources of infection were grouped into four broad categories: consumption, person-person contact, animal-person contact and environment-person (Otten et al., 2010).

Consumption of contaminated food and water could lead to CDI. One of the largest studies into environmental sources of *C. difficile* was carried out in South Wales and found that 5.5% of tap water and 2.3% of raw vegetables tested positive for *C. difficile* (al Saif & Brazier, 1996). A number of more recent studies have found *C. difficile* present at a significant level (6-42% of samples) in commercial meat products (Songer et al., 2009, Von Abercron et al., 2009, Weese et al., 2009, Weese et al., 2010). One study has shown presence of virulent *C. difficile* isolates in ready to eat salads (Bakri et al., 2009). Food and drink may therefore be a reservoir for CDI but transmission has not yet been directly proven. One study has shown that *C. difficile* spores can survive the heating recommendations for some meat products, and therefore current food preparation guidelines are not sufficient to prevent *C. difficile* transmission (Rodriguez-Palacios et al., 2010). Contamination of food during preparation by the persons preparing food is also a realistic means for *C. difficile* transmission but this has not been investigated. Direct person-person transmission of infection is likely to occur in a similar way to that in

which it occurs in the nosocomial environment. A couple of studies have been carried out to try to address transmission patterns among university students and employees (Ozaki et al., 2004) and in a family setting (Kim et al., 1981). Individuals in both settings were found to be infected with common strains but it is not possible to distinguish in these studies between direct person-person transmission and a common environmental source.

Person-person transmission should also be considered as a means of *C. difficile* being transmitted from environments with high levels of *C. difficile* contamination into the community by people who work in these environments. Healthcare workers are the obvious example and one study showed the 19% of staff uniforms tested positive for *C. difficile* after a shift, and that these were taken home for laundering (Perry et al., 2001). This may be a mechanism by which healthcare workers act as a reservoir for community infection. Employees and veterinary workers on farms where animal CDI is common may similarly spread *C. difficile* spores out from their work environment. In these cases the transmission from person-person would be preceeded by animal-person transmission.

Environment-person contamination is likely to be a significant transmission mechanism. In one study testing the hospital environment samples from diseased and colonised patients' rooms ranged from 0% to 53.3% positive with an average of 32.5%, while control rooms ranged from 0% to 7% with an average of 1.3% (Mulligan et al., 1980). In a number of similar studies the level of environmental contamination with *C. difficile* has been shown to

increase with proximity to colonised and diseased persons and animals. Therefore environmental sources of infection in the community are also likely to be produced by the presence of infected humans and animals. The only way to counteract this is by cleaning such areas with sporicidal products, as occurs in hospitals, but as the importance and relevance of such sources for infection have not yet been determined the possible benefits of this are not known. Environmental sources of infection less directly linked to human or animal infection are soil and natural water sources (lakes, rivers, sea). *C. difficile* can be isolated from such sources (al Saif & Brazier, 1996) and activities such as gardening and water-sports may provide a route for transmission.

Until more is understood about the relative contributions of different sources of infection to community associated CDI rates little is likely to be done to reduce the risks of transmission. In many ways nosocomial and community associated infection should not be considered as completely distinct infection types, as the nosocomial environment is directly linked in many ways to the wider environment. A better understanding of the dynamics of infection transmisson between these environments and the relative contributions of different infection sources may facilitate reduction of CDI incidence in both nosocomial and community settings.

1.4.2 Reducing patient susceptibility to *C. difficile* infection

Altering antibiotic prescribing practice is an effective method of reducing the rates of CDI. This is based on the knowledge that some antibiotics are less likely to induce CDI than others. Therefore if high risk antibiotic prescribing

(for example clindamycin and cephalosporins) is reduced in favour of prescription of low-risk antibiotics (for example vancomycin and penicillin) the overall risk of inducing CDI is reduced, and infection rates have been shown to fall (Wilcox *et al.*, 2004, Davey *et al.*, 2005, Davey *et al.*, 2006). In one study major changes in antibiotic prescribing were used to control a large outbreak of CDI in combination with optimal hygiene practices, and the importance of using a combination of both of these methods of CDI prevention is emphasized (Valiquette et al., 2007).

1.4.3 Novel methods for preventing CDAD

A number of studies have been carried out using probiotics to try to reduce CDI. The theory behind this prevention measure is that probiotics may partially counteract the floral disruption caused by antibiotic treatment. There is currently insufficient clinical evidence to support routine use of probiotics in prevention of CDI (Dendukuri et al., 2005, Graul et al., 2009). In the future improved trials with existing probiotics or new probiotic preparations may be shown to be efficacious in prevention of CDI (Gao et al., 2010). It should also be remembered that the possibility of inducing bacteraemia or fungaemia when using probiotics is a risk associated with this strategy for *C. difficile* prevention.

A twist on the strategy of using probiotics to prevent CDI is to use non-toxic *C. difficile* colonisation to prevent infection with toxigenic strains of *C. difficile*. This strategy was suggested on the basis of evidence showing that patients who are colonised with *C. difficile* prior to hospitalization are at decreased risk of developing diarrhoea (Shim et al., 1998). Studies in the hamster have

shown colonisation with non-toxic strains to effectively prevent CDAD in hamsters upon challenge with toxigenic strains (Sambol *et al.*, 2002, Merrigan *et al.*, 2009). Clinical trials are currently underway (Dale Gerding, personal communication).

Vaccination is an obvious strategy for prevention of infectious disease in general, and development of a vaccine to prevent *C. difficile* is an area of active research. As the *C. difficile* toxins are known to be the key virulence factors in CDAD most vaccine research so far has focused on vaccinating against these toxins. Existing effective toxoid vaccines for tetanus (caused by *Clostridium tetani*) and diphtheria (caused by *Corynebacterium diphtheriae*) demonstrate the potential of antitoxin immune responses in preventing toxin-mediated disease and provide encouraging precedent for a toxoid vaccine.

It has long been known that vaccination of hamsters with toxoids containing *C. difficile* toxins A and B can protect them from CDAD (Libby *et al.*, 1982, Fernie *et al.*, 1983). *C. difficile* culture filtrate and killed whole cells from a highly toxigenic strain have also been successfully used as antigens to prevent disease in hamsters (Torres et al., 1995). More recently a DNA vaccine encoding a domain of toxin A has been shown to protect mice from death when challenged with toxin A (Gardiner et al., 2009).

High concentrations of serum antibodies against toxin A in humans have been shown to correlate with reduced risk of developing recurrent CDAD (Kyne et al., 2001) and more recent evidence suggests that concentrations of anti-toxin

B antibody correlate in a similar way (Leav et al., 2010). The caveat to such studies is that a demonstration of antibody levels against a particular antigen correlating with risk of CDAD does not necessarily mean that it is antibodies against this antigen that are directly protective. It may be that some patients have a generally higher immune reponse to infection than others, which reduces their risk of developing CDAD. If this were the case antibody levels to many *C. difficile* antigens would be higher in these patients. In the same study that evaluated anti-toxin A antibodies it was also shown that patients with recurrent CDAD has significantly lower IgM antibodies against a crude preparation of *C. difficile* non-toxin antigens (Kyne et al., 2001). From this type of studies alone it is not clear which of the antibodies are protective against colonisation or disease.

Toxoid vaccination in healthy humans against toxin A can provide anti-toxin antibody titres exceeding that shown to correlate with protection against recurrent disease in CDAD patients (Aboudola et al., 2003). This means that testing the hypothesis that anti-toxin immunization may protect against CDAD is feasible with a toxoid vaccine. In a small study with 3 patients suffering from recurrent CDAD a toxin A and B toxoid vaccine successfully boosted levels of anti-toxin A and B antibodies and this was associated with resolution of recurrent diarrhoea (Sougioultzis et al., 2005). Larger trials are required to validate this vaccine strategy.

In terms of desired outcome for a *C. difficile* vaccine the main aim is to reduce CDAD. However, the ideal vaccine would also prevent colonisation with *C.*

difficile as this would reduce transmission of infection and herd immunity would be a positive outcome of vaccine usage. Anti-toxin vaccination alone would likely not provide protection against *C. difficile* colonisation, as naturally occurring strains without toxins can colonise. Therefore aside from anti-toxin vaccine strategies, the main avenue that has been pursued is use of surface antigens for vaccination with the aim of reducing the ability of *C. difficile* to colonise the host. Antibodies against surface proteins have been detected in CDAD patients (Pechine *et al.*, 2005, Wright *et al.*, 2008a) suggesting that these antigens are immunogenic and may play a role in host response to infection. It has been shown that high IgM antibody levels to *C. difficile* surface layer proteins are associated with a reduced risk of developing recurrent CDAD (Drudy et al., 2004). Passive immunization of hamsters with anti-surface layer protein (SLP) antibodies has been shown to prolong survival in the hamster model of infection (O'Brien et al., 2005). Active immunization of mice with surface proteins has also been shown to reduce colonistion levels during CDI (Pechine et al., 2007). Therefore, vaccination against surface antigens may have some efficacy in preventing colonisation but much more work is required to determine the best targets. It will also be important to consider strain coverage with surface antigen strategies, as the level of conservation of surface antigens is in general much lower than that of the toxins across *C. difficile* strains.

1.5 Treatment of CDAD

1.5.1 Current standard treatment

Current standard treatment for CDAD is to discontinue any antibiotic therapy that may have induced the CDI, then commence treatment with metranidazole or vancomycin. Metronidazole is the drug of choice for an initial episode of CDAD with mild-moderate symptoms, the recommended dosage being 500mg orally 3 times a day for 10-14 days. The preference for metronidazole over vancomycin is largely due to recommendations that vancomycin use in hospitals be reduced to reduce the selection pressure for the emergence of vancomycin-resistant enterococci (HICPAC, 1995). There is some evidence that vancomycin may be generally more effective in treating CDAD than metronidazole (Pepin et al., 2004), however more recent data does not support this claim (Pepin et al., 2007). For patients with severe disease symptoms vancomycin has been shown to be superior to metronidazole (Zar et al., 2007). Therefore vancomycin treatment is recommended for severe symptoms, at a dosage of 125mg orally 4 times per day for 10-14 days. In very severe cases oral vancomycin and intravenous metronidazole can be administered together and at higher dosages. In the most severe, complicated cases surgical colectomy is required, and can be life-saving (Lamontagne et al., 2007). High serum lactate level and peripheral white blood cell counts can help guide the decision to operate, as these are markers of severe CDAD (Cohen et al., 2010).

Vancomycin and metronidazole do both effectively kill *C. difficile*, but relapses

following withdrawal of the antibiotic are common. The relapse rate after initial therapy has been shown to range from 6 to 25% (Olson *et al.*, 1994, Bartlett, 1992). Patients who are retreated for relapse have a 40% probability of another relapse. Subsequent relapses occur at the alarmingly high rate of 60% or more. Studies of relapses show that about 60% are true relapses, meaning the same strain of *C. difficile* is the causitive agent. The assumption in these cases is that antibiotic treatment fails to fully eradicate the pathogen. However, it is impossible to distinguish between this and reinfection with spores of the same strain from the environment, which is likely to be contaminated. The remaining 40% of relapses must represent new infections with *C. difficile*, because they involve a different strain (Barbut et al., 2000, Johnson et al., 1989).

Treatment for relapses is usually determined using the same recommendations as for the initial episode (Cohen et al., 2010). However, for multiple relapses metronidazole is not recommended as there is the potential for cumulative neurotoxicity when using this drug long-term (Kapoor et al., 1999). In these cases vancomycin treatment is preferred and a tapered or pulse dosing regimen is recommended to try to allow recovery of some of the native flora before full removal of the antibiotic, which is suppressing *C. difficile*. The relapsing nature of CDAD reflects the fact that vancomycin and metronidazole are broad-spectrum antibiotics themselves and are therefore also CDI-inducing agents. Thus, the clinician is currently faced with the paradox that the standard drugs used to treat *C. difficile* infection also cause it. Reduction in relapse rate can be achieved in part by good infection control

measures, but one of the main aims of new therapy development is to prevent or overcome the CDAD relapses seen with vancomycin and metronidazole.

1.5.2 Resistance to current antibiotics

1.5.2.1 Metronidazole

Historically resistance of *C. difficile* to metronidazole has rarely been seen. Minimal inhibitory concentrations (MICs) of nearly all strains tested have been less than 2 mg/L. More recently slightly higher MICs have been reported. In Spain 6.3% of isolates were shown to have MICs of 32 mg/L or higher (Pelaez et al., 2005). In the UK recent isolates of a specific ribotype (001) were shown to have a 6-fold increase in MIC compared to historic isolates (Baines et al., 2008). Metronidazole does not achieve very high concentrations in the gut after oral administration due to >85% of the drug being absorbed. The mean concentration in watery stool has been shown to be 9.3 mg/L (Bolton & Culshaw, 1986). Therefore even a modest increase in susceptibility might affect clinical outcomes and continued metronidazole susceptibility monitoring will be necessary.

1.5.2.2 Vancomycin

The accepted MIC of vancomycin against *C. difficile* is 1-2 mg/L, with the highest ever reported MIC being 16 mg/L (Cohen et al., 2010). Vancomycin achieves very high faecal concentrations when administered orally, due to very limited absorption. Faecal levels of up to 880 mg/L have been measured (Keighley et al., 1978). Therefore development of resistance is considered less of a concern than with metronidazole. Concerns that vancomycin usage may promote the emergence of vancomycin-resistant enterococci (HICPAC,

1995) are more of a driving force to reduce vancomycin usage, than concerns of *C. difficile* itself becoming resistant to vancomycin.

1.5.3 Alternative antibiotics

Alternative antibiotics are being developed and tested with the aim of having agents available for use in the case of resistance emerging to metronidazole and/or vancomycin, but also as antibiotics with a more limited spectrum of activity would disturb the normal faecal flora to a lesser extent, thus reducing the risk of relapse.

1.5.3.1 Rifamycins

Rifampin, rifaximin and rifalazil are three members of the rifamycin class of antibiotic that have all been used to treat *C. difficile* (Garey et al., 2008). They all show good *in vitro* activity against *C. difficile* and high levels of the drug in the colon can be achieved. There is some evidence from pilot studies that there is a lower risk of relapse than with vancomycin therapy (Johnson *et al.*, 2007, Garey et al., 2008), but larger controlled trials are necessary to confirm this. The main worry with this class is natural and emerging resistance, which has been noted for many bacterial species, including *C. difficile* (Johnson et al., 2007, Curry *et al.*, 2009).

1.5.3.2 Ramoplanin

Ramoplanin is a new lipoglycodepsipeptide, which has been shown to have activity comparable to vancomycin for treatment of CDAD (Bartlett, 2009b). It has been fast-tracked for approval by the FDA, as it does not have the same drawback as vancomycin with concerns about vancomycin-resistant enterococci (Shah et al., 2010).

1.5.3.3 Fidaxomicin (OPT-80, PAR101, difimicin)

Fidaxomicin is a new macrocyclic antibiotic that has potent *in vitro* activity against *C. difficile*. Importantly it has limited activity against normal faecal flora (Louie et al., 2009). It achieves high concentrations in the colon as it is minimally absorbed from the gut. It has been tested against oral vancomycin in clinical studies of CDI therapy and has demonstrated efficacy equivalent to vancomycin for curing acute CDI. However, fidaxomicin therapy shows an improvement over vancomycin therapy in that it is associated with significantly lower risk of CDI recurrence in the 28 days following treatment. Symptom resolution is also slightly quicker for some patients (Miller, 2010). Thererfore this treatment shows promise in becoming a routine treatment for CDAD.

1.5.3.4 Nitazoxanide

Nitazoxanide was developed for treatment of protozoal infections of the gut, but was shown to have *in vitro* activity against *C. difficile*. Testing in the hamster model of infection showed that it was comparable to vancomycin and metronidazole in preventing lethality, but importantly unlike these drugs it did not induce CDI. It was therefore tested in humans in a randomized, double blind trial of metronidazole and was shown to be equally effective (Musher et al., 2006). A subsequent trial tested responsiveness to nitazoxanide in patients who were not responding to metronidazole. 74% of patients responded to treatment and 27% of these relapsed (Musher et al., 2007). Nitazoxanide may therefore be a useful addition to the *C. difficile* antimicrobial armory.

1.5.3.5 Tigecycline

Tigecycline was approved by the FDA in 2005 as the first in the glycylcyclines class of antibiotics. It is a ribosomal-binding protein synthesis inhibitor and has broad-spectrum activity. Despite this broad-spectrum activity, and therefore the disruption of the gut flora that tigecycline use causes, it is associated with low risk of CDI (Wilcox, 2007b). A large panel of *C. difficile* isolates have been shown to be susceptible to tigecycline *in vitro* (Nagy & Dowzicky, 2010) and recent case reports have shown that it is also effective at treating CDAD without causing relapses (Herpers et al., 2009, Lu et al., 2010). Larger trials are therefore warranted to assess the usage of tigecyline to treat CDAD.

1.5.3.6 Thuricin CD

A recent study focusing on the concept that the native bacterial flora is in some way protective against *C. difficile,* screened gut bacteria for anti-*C. difficile* activity and isolated an antimicrobial peptide from the human gut bacterium *Bacillus thuringiensis.* Thuricin CD kills *C. difficile,* some other *Clostridium* sp. and *Listeria monocytogenes*, but has no impact on most other genera tested (Rea *et al.*, 2010b). This narrow-spectrum bacteriocin was compared to vancomycin, metronidazole and a broad-spectrum bacteriocin in its propensity to affect the gut microbiota in an *in vitro* distal colon model. High-throughput pyrosequencing of 16s rRNA genes was used to determine the compostition of the microbiota before and after antimicrobial usage. Results of this 16sRNA sequencing are summarized in Figure 1.2. This showed that thuricin CD is as effective in killing *C. difficile* as the other agents but had no significant impact on the rest of the microbiota (Rea *et al.*, 2010a).

Therefore this bacteriocin or a strain producing it could be used as a therapeutic, which is unlikely to lead to relapses, due to its lack of effect the protective native flora.

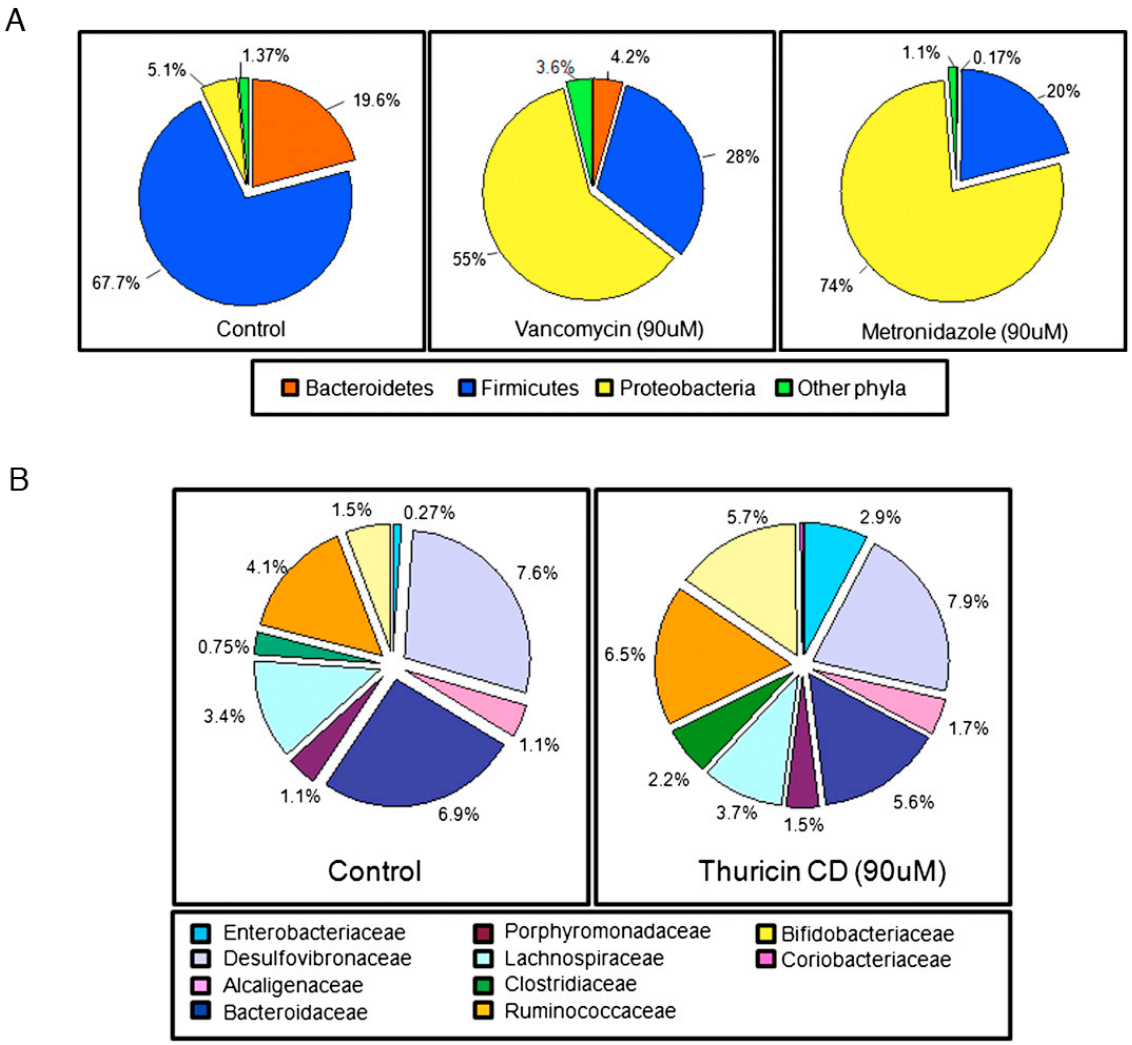


Figure 1.2 Diversity of microbial communities in an *in vitro* gut model before and after treatment with antimicrobials.

Determined by high-throughput 16s rRNA gene pyrosequencing. **A.** Phylum level diversity in an *in vitro* distal colon model. Control with no antimicrobial, is compared to vancomycin and metronidazole treatment. **B.** Family-level taxonomic distribution of the microbial communities present in an *in vitro* distal colon model with or without thuricin CD treatment. Values are expressed as percentages of total assignable sequences. Adapted from (Rea et al., 2010a).

It would be useful to compare all the antimicrobials discussed in this section for impact on the gut microbiota by 16s rRNA sequencing. It has been shown that different antibiotics have different effects on the gut microbial community (Antonopoulos et al., 2009). This information combined with their association with CDI relapse would determine what kind of microbiota disruption is most associated with predisposing to CDI and risk of relapse. This would provide a rationale for selection of antibiotics to treat *C. difficile* without causing relapse, and to treat other conditions without making patients susceptible to CDI. It may also lead to identification of the microbiota species most strongly associated with protection from CDI and/or relapse, and this could be used as a rational basis for development of therapies that specifically introduce or promote the growth of these species.

1.5.4 Faecal transplant/faecal biotherapy

Based on the rationale that disruption of the indigenous gut flora is a major risk for both primary infection with *C. difficile* and recurrence of disease, repopulation of the gut flora using stool from a healthy donor has been tested as a therapeutic strategy. A high degree of success has been reported in several uncontrolled case series (Gustafsson *et al.*, 1998, Aas *et al.*, 2003, Yoon & Brandt, 2010, Silverman *et al.*, 2010). A recent systematic review of the literature, covering a total of 150 patients, estimated that faecal therapy has exhibited a 91% success rate (van Nood et al., 2009). A randomised trial of faecal therapy is currently underway in the Netherlands (van Nood et al., 2009). 16S rRNA sequencing has been used to show directly how a recipient's intestinal microbiota was strikingly similar to a donor's after 14 days post-therapy, and it is this restoration of the microbial community that is

thought to be protective against *C. difficile* (Khoruts et al., 2010).

However, the availability of this treatment is limited. For safety reasons the donor should be screened for transmissible agents. The logistics of the treatment are relatively complex. The timing, collection and processing of the specimen from the donor followed by preparation of the recipient, and the route and means of administration are all key aspects of the treatment for which there is as yet no standard recommendation. Despite the efficacy of such a treatment it seems unlikely that it will be adopted as standard practice, unless in the next few years no other therapies emerge to successfully address the problem of *C. difficile* recurrence. As mentioned in the previous section, a better understanding of the specific species responsible for protection against CDAD would allow more palatable, defined and logistically easier treatments to be developed.

1.5.5 Passive immunotherapy

Intravenous human immunoglobulin has been used to treat a small number of patients with severe disease in a few uncontrolled studies (Salcedo *et al.*, 1997, McPherson *et al.*, 2006, Wilcox, 2004). Intravenous immunoglobulin (IVIG) is the pooled IgG extracted from the plasma of thousands of blood donors, which has been FDA-approved for a number of indications including chronic lymphocytic leukemia and pediatric HIV. The prevalence of serum antibodies against *C. difficile* toxins (and therefore likely also against other *C. difficile* antigens) in healthy populations, combined with the finding that recurrent disease is more likely to occur in patients with limited immune response to infection (Kyne et al., 2001) prompted the investigation into using

IVIG to treat CDAD. All studies with CDAD have demonstrated some success with this treatment in patients who were not responding to standard therapy. However, controlled studies are required to determine the true potential of this treatment. Unfortunately the prospects for controlled studies of intravenous immunoglobulin therapy are not good. Immunoglobulin manufacturers in the UK have not taken up any offer of being part of a comparative study (Wilcox, 2004). IVIG is also expensive and there is limited availability, therefore the use of IVIG may be considered more as a proof of concept for passive immunotherapy. More specific passive immunotherapy strategies are being explored that are more likely to become standard therapy for CDAD in the future.

Human monoclonal IgG antibodies with neutralizing activity against toxins A and B have been shown to be protective in the hamster model of infection (Babcock et al., 2006). Recently a randomized, double-blind, placebo-controlled phase 2 study was carried out with 200 patients, who were given anti-toxin A/B antibodies or placebo alongside the standard treatment of vancomycin or metronidazole. Those patients receiving antibodies had a significantly lower risk of recurrence of CDI, although there was no effect on acute infection (severity of diarrhoeal illness, duration of hospitalization or time to resolution of diarrhoea) (Lowy et al., 2010). Monoclonal antibodies against *C. difficile* antigens other than the toxins may also be efficacious in preventing or treating CDI, but our current understanding of virulence factors other than the toxins is limited and therefore is holding back development of such antibodies.

An alternative passive immunotherapy strategy in development involves immunization of cows with formaldehyde-killed whole *C. difficile* cells and fomaldehyde-inactivated *C. difficile* culture filtrate (toxoid). Protein concentrate from their milk, which contains antibodies, can then be used as an oral treatment for *C. difficile* infection. This treatment has been shown to be safe in CDAD patients (Young et al., 2007) and there is some evidence that such a strategy has efficacy that is comparable to metronidazole (Mattila et al., 2008). Further work is required to determine the efficacy of these treatments and how they compare to other passive immunotherapy strategies both in terms of efficacy and cost.

1.5.6 Active immunization

Vaccination could be used to prevent *C. difficile* infection, and was discussed previously. However, it could also be used to treat infection, as boosting the immune response to *C. difficile* may help to clear infection and prevent relapses. The only *C. difficile* vaccine trial in humans carried out so far has assessed the efficacy of a toxoid A and B vaccine in three patients with recurrent CDAD and in all cases was associated with resolution of disease (Sougioultzis et al., 2005). Therefore a vaccine for *C. difficile* could be used to treat patients who are not responding to standard treatments.

Larger trials are required to validate this vaccine strategy, and one concern is that some patients may not respond well to vaccination due to reduced immune responses in general, which may be underpinning their disease. Only clinical data will resolve this issue, but patients who have limited response to

active immunization at a critical time in their illness could be suited to passive immunization instead. Having both active and passive immunotherapies available may provide clinicians with a much better armoury with which to fight *C. difficile* infection.

1.5.7 Probiotics

As discussed in the prevention of CDI section, probiotics have shown some efficacy in trials both to prevent CDI and reduce the risk of recurrence. However, current evidence is not sufficient for endorsement of probiotic use to treat CDI and the increased risk of bacteremia or fungemia associated with probiotic use may outweigh the benefits (Pillai & Nelson, 2008, Cohen et al., 2010). The general strategy of using probiotics to promote the gut microbiota may prove to be efficacious for treating CDAD, but further research and development of treatments is required.

1.5.8 Anion exchange resins

C. difficile toxins A and B bind to anion exchange resins, and are therefore candidates for treating CDAD. Cholestyramine, a bile acid sequestrant that is used to treat hypercholesterolemia and to prevent diarrhoea in Crohn's disease patients, was used to treat clindamycin colitis before *C. difficile* was known to be the cause, and appeared to have some efficacy (Burbige & Milligan, 1975). There is a strong caveat with cholestyramine, in that it binds vancomycin. Therefore it can interfere with vancomycin efficacy. A later study using a different bile acid sequestrant colestipol did not show efficacy (Mogg et al., 1982), and ion exchange resins lost favour as a therapeutic strategy. Recently tolevamer was developed specifically to treat CDAD. It is a large

anionic polymer that binds *C. difficile* toxins, but has no antibiotic activity so will not disrupt the native flora (Barker et al., 2006). It has been tested against vancomycin and metronidazole in two large phase III studies and was shown to be inferior to both antibiotics in curing acute disease, although the relapse rate was lower (Weiss, 2009). Therefore development was halted. It seems unlikely that such a treatment with no direct antibiotic activity will be used alone to treat CDAD, although perhaps it could be useful in combination with antibiotic therapy or immunization strategies.

1.6 Economics of *C. difficile* infection

Determining the cost of *C. difficile* infection to healthcare systems is complex and likely to differ between different healthcare settings, different patient groups and different countries. A number of studies have been carried out to try to estimate the economic burden that *C. difficile* is responsible for. The conclusions of economical studies have important implications for the strategies behind developing and improving current preventative, diagnostic and therapeutic options.

There have been two excellent systematic reviews carried out in the past two years that have aimed to summarise and critically examine the available literature on the economic impact of CDI (Dubberke & Wertheimer, 2009, Ghantoji *et al.*, 2010). These reviews analysed studies carried out from 1983-2008 and converted currencies and costs to 2008 USD equivalents to allow for comparisons between studies. These converted costs are used in this section to allow for ease of comparison.

Costs attributed to CDI can be broadly grouped into three categories; costs of medication, laboratory costs for diagnosis and general hospital costs associated with length of stay. In a UK-based study carried out by Wilcox et al in 1996 CDI patients had an average increase in length of stay of 21.3 days over control patients. The overall cost per patient attributable to CDI was calculated to be $8995, and 94% of this cost was due to the increased length of stay, with only 6% due to medicine and diagnostic costs (Wilcox et al., 1996). A study in Northern Ireland in 2000 also calculated the major cost of CDI to be increased length of stay, accounting for 93.8% of the total attributable cost (Al-Eidan et al., 2000).

Two US-based studies have aimed to estimate the total cost to the US healthcare system attributable to CDI. In 2002 a prospective study on patients from a single hospital receiving antibiotic treatment and found that those who developed CDAD cost an extra $4486 per patient, of which most was accounted for by an increased length of stay of 3.6 days (Kyne *et al.*, 2002). Assuming 239,000 cases per year in the US this cost could be extrapolated to costing the US healthcare system a total of $1.3 billion per year. This study did not take into account subsequent relapses of infection. One study has estimated the cost of recurrences of disease, and shown that each episode is associated with an average 8.8 day length of stay, translating to a cost per episode of $4,096 (McFarland et al., 1999). Give the high rate of relapse seen with current therapies, this would add a significant amount to the total cost estimated by Kyne et al. In 2007 O'Brien et al. used retrospective data from the whole of Massachusetts to estimate the costs of CDI in the US (O'Brien et

al., 2007). Their total cost estimate was $3.4 billion per year, and this significantly higher estimate than Kyne et al. was due to increased estimates of cost per patient of $10,000 for a patient admitted for CDI and $14,507 for a patient with underlying disease with the secondary problem of CDI. Again these costs could be primarily attributed to increased length of stay.

Some studies have addressed specific patient populations. It has been estimated that CDI infection in inflammatory bowel disease patients costs $11,346 per patient, primarily due to an average increased length of stay of 3 days (Ananthakrishnan et al., 2008). A similar 3-day length of stay increase has been shown to occur in the intensive care unit (ICU) for patients with CDAD (Lawrence et al., 2007). In surgical inpatients length of stay has been shown to increase by an average of 16 days due to CDI, and this leads to an average increased cost of $90,664 per patient attributable to CDI (Zerey et al., 2007). This estimate shows how much the cost consequence of increased length of stay can vary between patient groups.

Important gaps in knowledge exist that should be highlighted. There is a lack of studies addressing the economic cost of CDI in non-hospital settings, such as long-term care facilities. Also studies so far ignore indirect costs associated with CDI such as missed work, reduced productivity and increased childcare costs due to increased length of hospital stay of patients. Recent changes in the epidemiology of *C. difficile* (discussed in the next section) appear to have altered infection rates, risk of relapse and severity of symptoms, and therefore

new studies are required to assess the impact of these changes on the economics of *C. difficile* infection.

Given the significant costs associated with CDI, it is clear that investment in improving prevention and treatment of CDAD makes economical sense, as well as improving the quality of life of patients and reducing morbidity and mortality. An example of spending money on prevention leading to overall cost saving is a study carried out in the US in 1998, where hospital-wide clindamycin usage restriction cost on average $269 per patient due to enforced usage of more expensive antibiotics. However, this led to a reduction of cases of CDI from 11.5 to 3.3 per month translating to an overall cost saving for the hospital (Climo et al., 1998). This exemplifies how cost-effective infection prevention measures can be, and in the more costly patient groups such as surgical inpatients prevention measures have the potential to save even more money, and therefore increased spending on prevention measures is justified.

A recent study used an economic computer simulation model to determine the potential economic value of a *C. difficile* vaccine (Lee et al., 2010). Using inputs for costs based on studies similar to those discussed so far, and probabilities of patients developing CDI and subsequent grades of CDAD symptoms based on clinical studies, the cost effectiveness of a vaccine could be simulated. Use of a vaccine to prevent initial CDI or to prevent recurrence of disease in CDAD patients were considered. A number of realistic scenarios with respect to vaccine cost, vaccine efficacy and risk of infection were shown

to be cost-effective. This suggests that as well as preventing morbidity and mortality a *C. difficile* vaccine could in fact save healthcare systems money, due to the substantial cost associated with CDAD management. Interestingly, using a vaccine to prevent recurrence of CDAD was shown to be likely to be more cost effective than a vaccine used to prevent initial episodes of CDAD, as the vaccine would only be administered to the target population and would prevent the high costs associated with recurrence. However, this model did not incorporate effects that a vaccine could have on *C. difficile* transmission. If a vaccine reduced colonisation rates this would lead to lower levels of contamination of the nosocomial environment and thus a lower risk of infection. This would make a vaccine strategy aimed at preventing initial CDI more cost-effective than this model suggests. From a strategic perspective, it is important to consider this aspect of vaccination, as the design of the vaccine (antigens used) would determine whether or not it is likely to reduce colonisation or just prevent CDAD symptoms. It would be interesting to compare by modelling the cost-effectiveness of a vaccine that only prevents CDAD symptoms (toxin-targeted) to a vaccine that also prevents colonisation, as this would help to guide vaccine development strategies.

1.7 *C. difficile* evolution and population biology

1.7.1 Typing strains

Since the discovery of *C. difficile* as a human pathogen a number of typing methods have been used to allow epidemiological studies of *C. difficile* to be carried out. Early methods relied on testing phenotypes such as antibiotic

resistance or serotype, but were quickly superceded by genotype-based methods once the technology became available. A brief description of the main typing methods currently in use, and their various merits are discussed in this section.

1.7.1.1 PCR-ribotyping

This method exploits the differences in length of the spacer regions of the 16S and 23S ribosomal RNA genes that occur in different strains. Specific primers are used that amplify a small number of fragments of DNA from this region and the banding pattern visualized by gel electrophoresis corresponds to a particular ribotype (Stubbs *et al.*, 1999, Bidet *et al.*, 1999). This has become a well-accepted typing system for *C. difficile*, particularly in Europe, exhibiting a useful level of discrimination for studying the macroevolution of *C. difficile*.

1.7.1.2 Restriction endonuclease analysis (REA)

Restriction enzymes are used to cut the bacterial genome at a relatively high frequency resulting in a large number of small DNA fragments that are separated on an agarose gel (Clabots et al., 1993). The complex banding patterns produced allow a high degree of discrimination between strains, but can be difficult to interpret and reproduce. REA is not widely used, therefore its interlaboratory reproducibility is not known.

1.7.1.3 Pulse-field gel electrophoresis (PFGE)

Restriction enzymes are used to cut the bacterial genome at a low frequency to yield large DNA fragments. These fragments are slowly separated in a polyacrylamide gel using repeatedly switching voltage. This enables the large DNA fragments to migrate through the gel, and the banding pattern visualized

by staining indicates the strain type sometimes referred to as a pulsovar (Killgore et al., 2008).

1.7.1.4 Amplified fragment length polymorphism (AFLP)

This technique is carried out by digesting the genomic DNA with restriction enzymes, followed by ligation of DNA adapters to the end of the digested fragments. Primers specific to these adapters are used to amplify fragments by PCR, and some degree of selectivity is included in the primers to limit the number of fragments amplified. AFLP types can be distinguished by the degree of similarity of the fragments amplified, usually a cut-off is determined and in the case of *C. difficile* different AFLP types are defined as having <86% similarity (Killgore et al., 2008).

1.7.1.5 Multilocus variable number tandem repeat analysis (MLVA)

MLVA is used to count the numbers of repeat alleles present at predefined, conserved repetitive loci in the genome by PCR amplification (van den Berg et al., 2007). This method produces a numerical result (code) for each strain. Much of the technology of this technique has been developed for forensic science, as it is the basis of human DNA fingerprinting. MLVA requires expensive equipment but is highly discriminatory and facilitates cross-laboratory comparisons.

1.7.1.6 Multilocus sequence typing (MLST)

This technique relies on the sequencing of a specified number of genes amplified by PCR. For each gene a strain will have an allele type based on point mutations. The combination of the different allele types for each gene defines the sequence type. As the data produced is sequence data this

technique is ideal for data-sharing and comparison between laboratories. For *C. difficile* an MLST protocol was developed analyzing seven housekeeping genes (*aroE*, *ddl*, *dutA*, *tpi*, *recA*, *gmk*, and *sodA*) (Lemee et al., 2004). However, this protocol has not been widely used to type *C. difficile* strains. This may be partly due to the discovery that the *ddl* gene does not amplify from all stains, compromising the usefulness of this protocol. Also a curated online database was not developed for this protocol. More recently a reportedly more robust MLST protocol has been developed with an online database to facilitate comparison of data between laboratories (Griffiths et al., 2010). Seven loci were chosen based on information available from *C. difficile* genome sequences. These genes (*adk, atpA, dxr, glyA, recA, sodA,* and *tpi)* were chosen as they show a significant degree of divergence across the 500 bp, but no gaps were present and according to the 630 genome annotation they span housekeeping genes. This MLST typing system correlates well with ribotyping and it remains to be seen whether it will be widely adopted.

1.7.1.7 *slpA* sequence typing (*slpA*ST)

*slpA*ST relies on the sequencing of a PCR product amplified from the variable region of the major surface layer gene (*slpA*). The amino acid sequence can be deduced from the DNA sequence and a new type is defined as differing from existing types by more than 20 amino acid residues. Groups differing by less than 20 amino acid residues are referred to as subtypes (Kato et al., 2005). As for MLST, sharing and comparing of sequence data makes this technique well suited to comparisons of strains typed in different laboratories.

1.7.1.8 Toxinotyping

This method focuses on the 19kb long pathogenicity locus (PaLoc), which encodes *tcdA* and *tcdB* genes for the *C. difficile* toxins A and B along with three regulatory genes *tcdC, R* and *E*. In order to toxinotype two PCR products (one from *tcdA* and one from *tcdB*) are amplified and subjected to restriction digest (an RFLP-PCR method) (Rupnik et al., 1998). By reference to characterised strains the test strain can then be identified as belonging to a certain toxinotype, or if a new toxinotype is identified further analysis of the PaLoc is carried out. There are currently 27 defined toxinotypes. Some *C. difficile* strains produce a third toxin CDT, a clostridial binary toxin, which is not encoded in the PaLoC. This can be tested for by carrying out a PCR to amplify one of the binary toxin genes *cdtB* (Stubbs et al., 2000).

1.7.2 Comparison of Typing Methods

A study carried out in 2008 by all the main *C. difficile* typing laboratories worldwide compared these techniques in their ability to type and discriminate between 42 isolates of *C. difficile* (Killgore et al., 2008). The isolates were selected to include strains known to be of different subtypes by a given method but also closely related strains, to allow the discriminatory power of the techniques to be tested. All techniques were successfully able to type all the isolates and showed a high degree of concordance of groupings by all methods. The discrimination index was calculated for each technique, which is a mathematical description of the ability of the technique to discriminate between strains. The closer the value is to 1 the higher degree of discrimination it exhibits. The discrimination indices for the techniques tested ranged from 0.964 to 0.631 in the following order: MLVA, REA, PFGE,

*slpA*ST, PCR-ribotyping, MLST, and AFLP. This meant that MLVA, REA, PFGE and slpAST could identify subtypes amongst the types determined by the less discriminatory techniques. All types were sufficiently discriminatory to detect a strain that is causing an outbreak within an institution. However, if intratypic discrimination is required to study the relationship between isolates of the same type that are temporally or geographically distinct MLVA or REA should be used. The final conclusion of this study was that all of these methods, if carried out in a laboratory with appropriate technology and expertise, are sufficient for detection of an outbreak strain within an institution. However, there is currently no method in use for strain tracking across institutions that exhibits intratypic discrimination with proven reproducibility between laboratories.

It should be noted that as the cost of genome sequencing continues to fall it is likely that many in-depth evolution studies will employ genome sequencing rather than any of these typing methods. The first of such studies on *C. difficile* has been recently published (He et al., 2010), and provides phylogenetic information that agrees with the results of typing studies. As more diverse genomes are sequenced, and we improve our understanding of the diversity of the species it is likely that sequence-based typing methods (eg. MLST, slpAST) can be developed or improved to provide the optimum level of discrimination for typing studies. Provided these sequence-based typing methods are sufficiently validated they would provide a much more transferable and inter-lab reproducible typing method than the other methods. This would facilitate worldwide study of *C. difficile* epidemiology.

1.7.3 Patterns of disease transmission

Many studies have used typing to characterise the strains present during outbreaks of *C. difficile* in hospitals (Kato et al., 2001, Muto et al., 2007). These studies provide information about transmission of strains within this environment and evidence for the need to improve or increase infection control measures in order to get the outbreak under control.

Large-scale typing studies across regions or countries are also carried out with the aim of monitoring the general prevalence of different *C. difficile* types. These studies provide information about the transmission of *C. difficile* over larger geographical areas. Some studies include data from community-acquired cases and animal isolates. Temporal information about the types causing disease can be ascertained if samples are collected over a long period of time. In the past 10 years these studies have highlighted that the prevalence of different *C. difficile* types causing disease has changed and is continuing to change. This has been accompanied by an increase in *C. difficile* incidence and changing clinical aspects of CDI, and therefore studies are also trying to address possible associations between *C. difficile* types and these changes.

The most significant recent change is the emergence of outbreak-causing ribotype 027 strains. It was reported in 2005 that outbreaks that occurred between 2000 and 2003 in eight different health care facilities in the United States were all associated with this type of strain (REA group B1, PFGE type NAP1, ribotype 027), and for five of the facilities the majority of strains were

027 (McDonald et al., 2005). This finding was of significance because historically 027 was a rare type. Since the original reported ribotype 027 strains have been isolated from outbreaks in 40 US states, Canada, most of Europe, Hong Kong and Japan (Kato *et al.*, 2007, Smith, 2005, Cheng *et al.*, 2009, Kuijper *et al.*, 2007, Loo *et al.*, 2005, O'Connor *et al.*, 2009). In English hospitals this ribotype became the most common type of *C. difficile* strains associated with hospital outbreaks, accounting for 41.3% of isolates collected nationwide in 2007-2008 (Brazier et al., 2008). A number of studies have found evidence that cases caused by 027 strains are associated with more severe disease than cases caused by other ribotypes, although not all studies have been able to show this (Freeman et al., 2010). There are many possible confounding factors such as unrecognized independent risks, temporal changes in the patient population, and infection control measures that make studies of this nature very hard to compare with confidence. It is clear that ribotype 027 *C. difficile* has spread quickly around the globe causing outbreaks of CDAD, but the evidence for its role in increased severity of CDAD is still under scrutiny.

Cases of CDAD caused by ribotype 078 *C. difficile* strains also seem to be on the increase. In the Netherlands between 2005 and 2007 the proportion of cases associated with 078 ribotypes increased from 3 to 13% (Goorhuis et al., 2008a). Disease severity was reported to be similar to the 027 ribotype, but interestingly 078 cases were more likely to be community-acquired and affected younger patients. As this ribotype is commonly found in pigs and

calves, the possible zoonotic transmission of this type of *C. difficile* has been recognized and is under investigation (Goorhuis *et al.*, 2008b).

Another ribotype, which has emerged causing outbreaks of CDAD in the past 10 years is ribotype 017 (Drudy *et al.*, 2007b, Kim *et al.*, 2008). These strains are of particular interest as they are toxin A-negative, toxin B-positive strains. The prevalence of these strains had immediate significance to diagnostic practices, as toxin A-specific assays were common, but would obviously not detect these strain types. Any therapeutic strategies targeting toxin A alone are therefore also undesirable.

Despite the presence of *C. difficile* types that are associated with outbreaks of *C. difficile* and the intriguing changing patterns of incidence and disease severity that may be associated with the evolution of particular types, a clear and underemphasized conclusion from *C. difficile* typing studies is that the types of strains causing CDAD at any one time and in any one place are remarkably diverse. A recent large study typed samples (REA and ribotype) from 1,111 CDAD cases from 187 sites in 16 countries that were acquired during a clinical trial for tolevamer (Cheknis et al., 2009). Isolates were REA-typed and compared to 7 REA types known to have caused hospital outbreaks. In North America the most common type was REA B1 (ribotype 027), accounting for around 40% of isolates. Of the other 60% of isolates, half were one of the other 6 types known to cause outbreaks and half were a non-specific REA type. In Europe the most common REA type was type J (ribotype 001), accounting for 19% of isolates, 37% fell into the other 6

common REA groups and 44% were a non-specific REA type. 24 samples were from Australia and the majority of these types were not of the well-known outbreak types, indicating differences in the geographical distribution of *C. difficile* types. Therefore the overall diversity of *C. difficile* strains responsible for CDAD is remarkably high, and the prevalence of known outbreak types and emerging types is not as high as may be suggested by the number of studies focusing on these emerging types. The results of this study are summarized in Figure 1.3

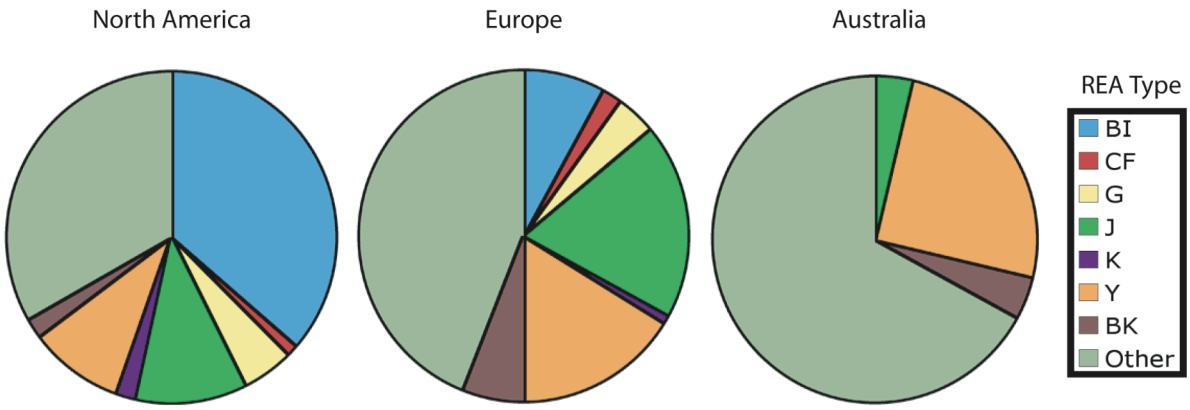


Figure 1.3 REA types of *C. difficile* isolates recovered from a worldwide clinical trial.

Data taken from (Cheknis et al., 2009). The diversity of REA types is clear, with a large number of isolates (Other) not falling into any of the outbreak types. Geographical differences between incidences of REA types are also clear.

1.7.4 Genomic studies

The first *C. difficile* genome sequence was published in 2006 (Sebaihia et al., 2006). This strain is known as 630, is a PCR-ribotype 012 strain with multi-drug resistance and was isolated from a patient with severe pseudomembranous colitis in Zurich. It has a circular 4.29 Mb chromosome

and a 7.9 kb plasmid, together encoding 3679 predicted protein sequences. A large proportion (11%) of the genome consists of mobile genetic elements. The presence of these elements implies that *C. difficile* may have acquired an extensive array of its genes involved in antimicrobial resistance, virulence and host interaction by horizontal gene transfer.

Having this reference *C. difficile* genome sequence enabled a PCR spotted microarray to be constructed, including probes for all the known genes from this strain 630. A broad range of 75 strains from widespread geographical locations isolated at different times was then tested using this microarray to determine levels of gene conservation across the *C. difficile* strain population and to give information about the phylogeny of strains (Stabler *et al.*, 2006b). This comparative phylogenomics study revealed four distinct clades. PCR ribotype 027 strains fall into one clade, toxin A-negative toxin B-positive ribotype 017 strains fall into another clade, and there were two further clades containing a mixture of animal and human isolates. Ribotype 078 strains were all shown to fall into one of these clades, with both animal and human isolates clustering together, further suggesting the possibility of zoonotic transfer. Closely related ribotype 027 and 017 strains were from diverse geographical origins, providing further evidence for the transcontinental spread of these lineages.

This microarray study also provided evidence of evolution within the 027 strains, providing clues as to why this lineage suddenly became such a common CDAD-associated isolate and spread globally (Stabler et al., 2006b).

To follow up this finding two ribotype 027 strains were sequenced, a historic isolate and an isolate from a recent 027 outbreak. Three way genomic comparisons were carried out between these two strains and 630. Genes were therefore identified that were specific to the 027 lineage and that were specific to the recent 027 isolate, suggesting that they have been recently acquired and may have contributed to the increased incidence of human CDAD caused by this lineage (Stabler et al., 2009). Future work will aim to experimentally determine the impact of these identified differences on the virulence of *C. difficile* isolates.

From this original microarray study only 19.7% of genes were highly conserved by all 75 strains, highlighting how diverse the *C. difficile* species is (Stabler et al., 2006b). A later similar comparative microarray study was carried out using 73 isolates from humans and animals (Janvilisri et al., 2009). Again the isolates could be grouped into 4 clades, with 3 of the 4 groups containing both human and animal isolates. From this study 16% of 630 genes were highly conserved among all strains. This study also contained probes for genes from a sequenced 027 strain, to define a core gene set for this clade. Interestingly this study showed that some divergent genomic regions were conserved between strains isolated from the same host species. These regions may be associated with host adaptation. In both studies the genes common to all isolates tested (core genes) were not only house-keeping genes, some appeared to have more complex functions and may represent genes that are key for infectivity and virulence of *C. difficile*.

A very recent comparative array study using *in situ* inkjet oligonucelotide synthesis to generate the probes allowed a more comprehensive set of probes to be used in an array (Marsden et al., 2010). This microarray consisted of over 41,000 probes (~10-fold more probes than previous arrays, which had approximately one probe per gene), mostly based on the 630 genome sequence with an extra 631 probes covering genes present in sequenced 027 strains and not in 630. 94 strains were chosen for comparative hybridization based on selection of the most prevalent ribotypes from clinical isolates in the UK and Europe. This array found that 32% of the 630 probes were conserved, higher than the previous two array studies. The variety of strains tested was wider, but as the density of probes used was much higher, variable genes such as *slpA* could be detected by probes located in their conserved regions. Therefore it seems likely that the 32% level of conservation estimated from this denser array is a more accurate estimate of conservation of genes across *C. difficile* strains than estimated by the PCR-arrays. One of the aims of the study was to see if any sequence markers could be found that are in the 027 strains and no others. No such markers were found, in fact no single probe was found to be representative of just one ribotype. However, some genes appear to be most conserved amongst known outbreak-causing ribotypes (027, 001, 020, 106 and 078) and therefore may be associated with increased transmission and/or virulence (Marsden et al., 2010).

Further detail regarding the evolutionary relationship between different *C. difficile* isolates has come from new genome sequencing projects. A recent

study reported comparing the genome sequences of thirty *C. difficile* isolates (He et al., 2010). There were two aims of this study, to investigate the macroevolution of *C. difficile* as a diverse species and to investigate the microevolution of the ribotype 027 clade. Therefore 8 diverse isolates from human and animal origin were sequenced and 21 isolates from the 027 clade. The genome sequences of the diverse isolates were used to construct a phylogenetic tree (containing 630, and ribotypes 027, 001, 017, 078). Each ribotype was confirmed to be a distinct lineage, indicating that ribotyping is a good reflection of *C. difficile* genetic diversity. Interestingly the 078 lineage appeared to be highly divergent from the other lineages. Given the significant diversity of the *C. difficile* species this study concluded that it is an ancient bacterial species, with an estimated age of 1.1-85 million years.

The phylogentic trees constructed from all these studies have confirmed that outbreak-causing strains have emerged independently from multiple lineages, suggesting that virulent strains can emerge from across the diversity of the species, rather than being confined to a specific pathogenic lineage. This suggests that certain genetic elements common to all *C. difficile* strains underlie virulence. Given our current knowledge it is likely that the incidences of different lineages causing disease are dynamic, and we cannot predict which lineages may cause future outbreaks. Therefore it may make more sense to try to understand the common features (core genes) of *C. difficile* as an ancient species, and how these underpin virulence, rather than focusing on the differences between current strains. In order for intervention strategies to

be successful in the long-term they must target the diversity of *C. difficile* strains, rather than specific lineages.

1.8 Studying the molecular basis of *C. difficile* pathogenesis

From typing and sequencing studies of different naturally occurring *C. difficile* strains, and correlations with various phenotypic assays or clinical outcomes it has been possible to make hypotheses as to the molecular bases of *C. difficile* virulence. However, definitive studies of the molecular basis of bacterial virulence rely on the ability to manipulate the pathogen at the genetic level and test the resulting isogenic strains for virulence. In this section the tools currently available for genetic manipulation and experimental models of *C. difficile* infection are discussed.

1.8.1 Tools for the genetic manipulation of *C. difficile*

A system for clostridial gene knock-out utilising the mechanism of group II intron insertion has been developed. The rules governing integration site specificity of mobile group II introns have been elucidated (largely base-pairing between intron RNA and target site DNA) therefore the intron can be retargeted to disrupt genes of interest (Mohr et al., 2000). These constructs are referred to as Targetrons (Perutka *et al.*, 2004, Karberg *et al.*, 2001, Zhong *et al.*, 2003). A retrotranscription activiated marker (RAM) is used for positive selection of genomic integrants: a resistance gene is carried by the group II intron, but before transfer it is interrupted by a self-splicing group I intron. The process of chromosomal integration of the group II intron leads to

excision of the group I intron, generating a functional selectable marker in the group II intron integrated into the genome. This system has been adapted to incorporate a promoter and selectable markers that function in the clostridia: resulting in the ClosTron system (Heap et al., 2010, Heap et al., 2007). In proof of principle studies, ClosTron constructs were used to knock out *pyrF* and *spo0A* genes in *C. difficile,* C. *acetobutylicum* and C. *sporogenes*.

The ClosTron system generates stable mutants as there are no homologous sequences flanking the insertion site created using this technique. Therefore the insertion cannot be lost by homologous recombination to revert to the original genotype. The only other reported method for making site-specific *C. difficile* knock-outs is the use of an unstable plasmid with a targeting region of homology to the gene of interest to generate single homologous crossover insertions in the *C. difficile* chromosome (O'Connor et al., 2006). The drawback of this method is that the recombination event that generates the site-specific insertion is reversible, making mutants generated by this method inherently unstable.

An alternative to the reverse-genetics approach of making site-specific mutants is the forward-genetics approach of generating a library of random mutants and screening these for phenotypes of interest, then determining the mutations responsible. The advantage of the forward-genetics approach is that the genetic basis of a phenotype of interest can be elucidated without a prior assumption as to the genes involved. This year the first report of a system for creating a suitable library of random mutants in *C. difficile* was

published (Cartman & Minton, 2010). This system relies on the mariner-transposon to insert randomly at TA target sites, ideal for *C. difficile* with its low G-C content. The transposon and its transposase are introduced to *C. difficile* on a unstable shuttle plasmid, such that once the transposon has inserted the plasmid is lost, losing the transposase from the cell, thus immobilizing the transposon at its insertion site. This new tool for creating libraries of *C. difficile* mutants is likely to facilitate identification of genes associated with phenotypes important for virulence

Introducing genetic material to *C. difficile* via stable replicative plasmids has also only been possible in recent years. Stable plasmids based on the plasmid replicon isolated from an indigenous *C. difficile* plasmid pCD6 have been developed (Heap et al., 2009). These autonomously replicating vectors are transferred to *C. difficile* through oriT-based conjugation from *E. coli* donor strains, as there is no established method for direct transformation of *C. difficile*. These replicative plasmids allow expression of exogenous genes in *C. difficile*, over-expression of *C. difficile* genes and complementation of site-specific mutants. There is currently no reported method available for introducing genes of interest to the chromosome of *C. difficile*.

1.8.2 Experimental models of *C. difficile* infection

Now that it is possible to make isogenic mutants and random mutant libraries of *C. difficile* experimental models of infection are required to test these mutants to allow the molecular basis of *C. difficile* virulence to be determined.

1.8.2.1 Hamsters

It has long been known that hamsters are susceptible to *C. difficile* infection, exhibiting many of the symptoms observed in man, including diarrhoea, histological damage, colonisation of the large bowel and sporulation of the organism at the terminal stage of the disease (Fekety et al., 1979), and the antibiotic-compromised golden Syrian hamster *C. difficile* infection is the established model for acute CDAD. Hamsters are given a dose of clindamycin (or another antibiotic) to perturb the normal protective gut flora and subsequently challenged with *C. difficile*. Typically toxigenic *C. difficile* infection is lethal within a few days and time to death is commonly used as a measure of virulence (Razaq et al., 2007, Sambol et al., 2001). Recent refinement of the model uses telemetry chips to monitor movement and body temperature of the hamsters, allowing the animals to be culled prior to death, giving the experiments a more humane endpoint whilst also providing extra scientific data regarding the course of infection (Douce & Goulding, 2010). The lethality is due to hamsters being very sensitive to the *C. difficile* toxins. Non-toxic *C. difficile* strains asymptomatically colonise hamsters, and have in fact been used to protect against colonisation with toxic *C. difficile* strains, indicating that the symptoms observed are primarily toxin-mediated (Borriello & Barclay, 1985, Merrigan et al., 2009, Wilson & Sheagren, 1983). Therefore this model has been extensively used to model the acute phase of *C. difficile* infection where toxins are key virulence determinants.

The large number of reported studies using the hamster model can be grouped into three main categories; studies evaluating preventative or

therapeutic interventions, studies comparing different *C. difficile* isolates, and studies comparing isogenic *C. difficile* mutants to determine the role of *C. difficile* genes in virulence. This first category includes pre-clinical vaccination experiments (Torres et al., 1995, Giannasca *et al.*, 1999), evaluation of antibiotics in the induction (Merrigan *et al.*, 2003b, Larson & Borriello, 1990) and treatment of CDAD (Anton et al., 2004, Ochsner et al., 2009), evaluation of probiotics and passive immunization strategies (Kink & Williams, 1998, van Dissel *et al.*, 2005). Studies comparing naturally occurring *C. difficile* isolates have shown that there are significant differences in the virulence and pathology caused by different isolates, these sometimes correlate with clinical virulence trends, however from these studies alone it is not clear what the molecular basis are for the differences between isolates (Delmee & Avesani, 1990, Sambol et al., 2001, Razaq et al., 2007, Goulding *et al.*, 2009). Given the recent development of *C. difficile* genetic manipulation techniques there only three reported studies using the hamster model to compare isogenic mutants of interest. (Lyras et al., 2009), (Kirby et al., 2009, Kuehne et al., 2010), findings of these studies are discussed in the relevant virulence factor sections below.

The hamster model has proved very useful in studying the acute toxin-mediated stage of disease. However it is desirable to have animal models, which can be used to study other types of *C. difficile* infection such as chronic infection, asymptomatic carriage and transmission of infection. These non-acute, non-fulminant types of infection are very relevant to human infection and improving our understanding of these processes is important for

improving our ability to manage *C. difficile* infection. Additionally, there are limited resources available to study host-aspects of hamster infections such a knock-out animals and cell lines. Therefore, other animal models are being pursued.

1.8.2.2 Mouse models

It has long been known that *C. difficile* can be isolated from mice, suggesting that mice can be colonised by *C. difficile* (Itoh et al., 1986). Models of acute infection in mice have been developed either using gnotobiotic (germ-free) mice or by antibiotic treatment of conventional mice prior to *C. difficile* exposure (Wilson *et al.*, 1986, Chen *et al.*, 2008). Such models can be used in a similar way to the acute hamster model of infection. In a recently described model five different antibiotics were used to treat mice prior to *C. difficile* exposure. Disease severity was shown to vary from minimal symptoms to fulminant disease, associated with the dose of *C. difficile* used to infect the animals (Chen et al., 2008). This spectrum of disease more closely mimics human disease and may therefore be an improvement on the fulminant-only models of infection.

Another recent study showed that mice can be asymptomatically colonised with *C. difficile* establishing a carrier state that persists for months. Antibiotic treatment with clindamycin leads to a high number of spores being shed in the faeces, referred to as a super-shedder state. Although no overt symptoms are seen in these animals, they do exhibit mild inflammation of the intestine, accompanied by disruption of the tight junctions, microvillus effacement and the transient production of an inflammatory exudates in the faeces. After

cessation of antibiotic treatment the number of *C. difficile* spores in the faeces decreases, but if a subsequent dose of clindamycin is administered animals return to the super-shedder state (Lawley *et al.*, 2009a). This is reminiscent of relapsing infection in humans. This model demonstrates active colonisation of mice by *C. difficile* without acute symptoms, and is therefore a valuable model for studying the molecular mechanisms underpinning colonisation. The power of using a mouse model was further demonstrated in this study by the use of a knock-out mouse line. Myd88(-/-) mice with impaired innate immune function were shown to develop severe intestinal disease upon challenge with *C. difficile* that was often fatal (Lawley et al., 2009a). The availability of knock-out mice will allow investigation into the role of host factors in the response to *C. difficile* infection.

1.8.2.3 Piglet model

A recent study investigated the use of piglet models of infection, as piglets are known to be susceptible to *C. difficile* infection (Steele et al., 2010). Using gnotobiotic piglets *C. difficile* was shown to cause a range of symptoms from chronic diarrhoea to acute lethal diarrhoea, dependent of the age of the piglets and the dose of *C. difficile* given. This model of infection therefore also allows investigation of colonisation, providing a useful alternative to the acute hamster model.

1.8.3 *In vitro* gut model

Study of *C. difficile* based on *in vitro* growth in faecal emulsions was first used in the mid-1980s to demonstrate inhibition of *C. difficile* growth by faecal emulsions. A continuous flow culture system (Freter et al., 1983) was also

used to demonstrate the suppression of pathogens including *C. difficile* by normal hamster flora (Wilson & Freter, 1986). This system has been further developed into a three stage compound continuous flow model, with the transit time, pH, and metabolic composition of each stage mimicking the human gut (Macfarlane et al., 1998). This model has been used to investigate the effects of dietary input such as prebiotic non-digestible oligosaccharides (Hopkins & Macfarlane, 2003); the elimination of spores by antibiotic administration (Baines et al., 2005); and the effects of antibiotic administration on gut microflora composition and on *C. difficile* proliferation and toxin production (Freeman *et al.*, 2003, Baines et al., 2005, Freeman *et al.*, 2005) including *C. difficile* strain comparisons (Baines et al., 2006, Freeman et al., 2007). Overall this *in vitro* model system allows designed experimentation into the interactions of complex ecosystems under a controlled environment. However, in this model there is obviously no interaction between bacteria and host cells or the immune system.

1.9 *C. difficile* virulence factors

As previously discussed, genetic manipulation of *C. difficile* has only been possible in the last few years, and therefore most of our knowledge of *C. difficile* virulence factors does not come from definitive studies of isogenic mutants tested in animal models. However, other types of studies have yielded information about *C. difficile* virulence factors, and our current understanding of the molecular basis of *C. difficile* virulence is summarized in this section.

1.9.1 Toxins

It has been known for some time that some strains of *C. difficile* produce two related UDP-glycosylating toxins, known as toxin A and toxin B (TcdA and TcdB) (von Eichel-Streiber et al., 1992). These two toxins are encoded in a pathogenicity locus (PaLoc) typically 19.6kb in size, which also encodes protein regulators of their expression (Braun et al., 1996). A diagram of the prototypical PaLoc is shown in Figure 1.4. As mentioned previously, the PaLoc varies between strains, with some strains not having a PaLoc at all (non-toxic strains), others having only *tcdB,* as well as many other variable regions of the PaLoc (Rupnik, 2010). Given that non-toxic strains do not cause symptomatic disease and are generally regarded as being non-pathogenic, these toxins are key *C. difficile* virulence factors, and have therefore been extensively studied.

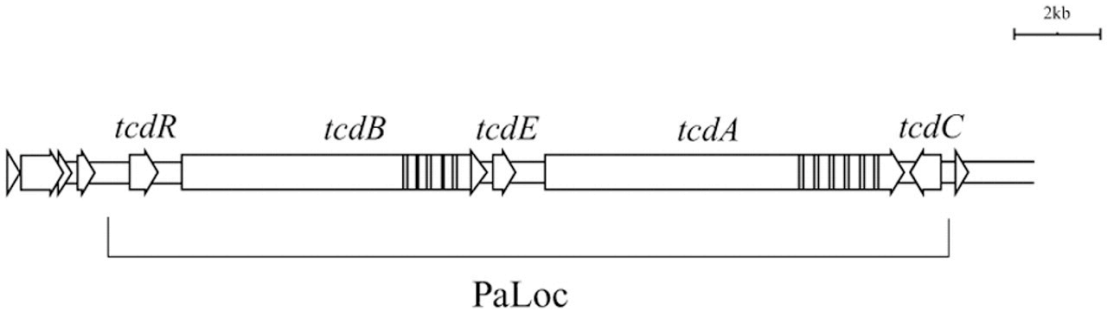


Figure 1.4 Diagrammatic representation of the *C. difficile* pathogenicity locus (PaLoc).

The five proteins encoded are the two major toxins TcdA and TcdB, the positive regulator (alternative sigma factor) TcdR, the negative regulator TcdC and the putative holin TcdE. Diagram taken from (Rupnik et al., 2005).

1.9.1.1 Toxin structure and function

TcdA and TcdB are members of the large clostridial glucosylating toxin family, which includes toxins from many other clostridial species. For example, the

hemorrhagic and lethal toxins from *Clostridium sordellii* and the alpha toxin from *Clostridium novyi*. Toxins in this family are very large, 250-308 kDa in size and have a high degree of sequence identity (ranging from 36% tp 90%). A four-domain structure of these toxins, known as the ABCD model has been suggested based on many structural and biochemical studies (Jank & Aktories, 2008). In this model the A (activity) domain at the N-terminus confers glucosyl transferase activity. Adjacent to this is the C (cutting) domain, which is a cysteine protease domain that mediates auto-cleavage of the toxin. In the middle of the protein is the D (delivery) domain, which contains a hydrophobic region and is thought to mediate translocation of the toxin across host cell membranes. At the C-terminus the B (binding) domain consists of polypeptide repeats, which mediate binding of toxin to host cells. For TcdA this binding region has been shown to bind to a trisaccharide (Krivan *et al.*, 1986, Greco *et al.*, 2006), and it is therefore thought that TcdA binds to host cells via interactions with glycosylated ligands. TcdA and TcdB act in the cytosol of host cells, glucosylating small GTPases including Rho, Rac and Cdc42. The GTPases are glucosylated at a specific threonine residue, which prevents them adopting their active GTP-bound conformation thus causing inactivation of the GTPases and blocking of signal transduction pathways. This leads to actin condensation, cell rounding and eventually cell death. In the gut this leads to impairment of tight junctions, resulting in fluid accumulation and extensive damage to the integrity of the epithelium (Voth & Ballard, 2005).

1.9.1.2 Regulation of toxin expression

Toxin production is under transcriptional regulation. Early studies showed that toxins are not produced during exponential growth phase, nor in the presence of glucose, but are expressed in stationary phase (Dupuy & Sonenshein, 1998). The TcdR protein encoded upstream of *tcdB* in the PaLoc is an alternative RNA-polymerase sigma-factor, which is required for the expression of TcdA and TcdB from their own promoters (Mani & Dupuy, 2001). Homologous proteins are found to regulate toxin expression in a number of other Clostridia, and these sigma-factors constitute their own subgroup of sigma-factors (Dupuy & Matamouros, 2006). Some of these sigma factors are conserved to a high enough level that they are functionally interchangeable (Dupuy et al., 2006). TcdR expression is responsive to the cellular growth phase and constituents of the growth medium, and it stimulates its own synthesis. It is therefore thought that toxin expression level is regulated indirectly by the regulation of TcdR expression level in response to environmental signals (Mani et al., 2002). Growth temperature has also been found to control the expression of TcdA and TcdB, dependent on TcdR (Karlsson et al., 2003). One study reported that quorum sensing may play a role in regulation of toxin production (Lee & Song, 2005), but the mechanistic basis of this regulation has not been investigated. There is evidence that the upregulation of toxin production is accompanied by the upregulation of proteins involved in alternative energy metabolism and sporulation (Karlsson et al., 2008), and it is therefore possible that TcdR plays roles in regulation of other proteins as well as the toxins. The TcdC protein encoded downstream of the toxin genes is a negative regulator of toxin expression, which destabilizes

the TcdR-containing RNA-polymerase holoenzyme, inhibiting transcription of the toxin genes (Matamouros et al., 2007). TcdC is expressed during exponential phase, ensuring toxins are not expressed during this growth phase (Dupuy et al., 2008). The presence of a positive and negative regultor of TcdA and TcdB transcription in the PaLoc is likely to allow tight regulation of toxin production during infection. It will be interesting to see the effects of mutations in these genes on virulence in animal models of infection now that the genetic technology for production of isogenic mutants is available.

1.9.1.3 Contributions of toxins to *C. difficile* virulence

Two recent studies have reported testing toxin A and toxin B negative isogenic mutants of *C. difficile* 630 in the hamster model of infection. One study showed that toxin B is essential for virulence, but toxin A is not (Lyras et al., 2009). The other study showed that either toxin A or toxin B alone can cause fulminant disease in the hamster model, with only a toxin A and toxin B double-mutant strain being avirulent (Kuehne et al., 2010). In earlier studies utilizing purified toxin preparations to challenge hamsters toxin A alone was shown to elicit all the symptoms of CDAD, and toxin B did not cause any symptoms unless there was prior intestinal damage or toxin A was co-administered (Lyerly et al., 1985). A number of possible explanations for these differences have been suggested (Carter et al., 2010). The isolation of natural isolates expressing only toxin B that cause the same spectrum of disease as toxin A and B expressing strains provides some evidence that toxin B is also key to virulence in humans (Drudy *et al.*, 2007a). The relative contribution of these two toxins to virulence has obvious implications for diagnostic and vaccine design. It is now standard diagnostic practice to detect toxin B, as

toxin A is lacking in some virulent strains. Further work using current genetic manipulation technology with diverse *C. difficile* strains in a variety of animal models is required to resolve the relative importance of these two toxins in disease.

1.9.1.4 Binary toxin

Some *C. difficile* strains produce an actin-specific ADP-ribosylating binary toxin encoded by *cdtA* and *cdtB*, which are not found in the PaLoc (Perelle et al., 1997). These binary toxin genes are tightly regulated by CdtR, which is encoded next to *cdtA/B* in a locus named CdtLoc, which is found inserted into the same position in all strains analysed so far (Carter et al., 2007). Binary toxin ADP-ribosylates G-actin and inhibits actin polymerization, which induces a redistribution of microtubules. It has been shown that this can increase the adherence of *C. difficile* to human cells *in vitro* (Schwan et al., 2009). The contribution of binary toxin to the overall virulence of *C. difficile* strains remains to be determined.

1.9.2 Sporulation factors

C. difficile produces spores, which can survive outside the colonic environment for long periods of time and are therefore responsible for long-term contamination of the environment with viable *C. difficile* organisms. Sporulation is therefore key to transmission of *C. difficile* infection, and factors important to the sporulation process can be considered virulence factors. Comparative genomics of the well-characterised sporulation processes of *Bacillus subtilis* and other Clostridial sp. (Paredes et al., 2005) allows identification of putative factors in the *C. difficile* sporulation process.

The master regulator of sporulation SpoOA is well conserved across spore forming species and the *C. difficile* gene has been shown to be essential for sporulation in knock-out studies (Heap et al., 2007, Underwood *et al.*, 2009). ClosTron knock-out technology has also been used to confirm the role of a histidine kinase (CD2492) in sporulation that was identified as playing a putative role by bioinformatic analysis (Underwood et al., 2009). Interference of the sporulation initiation pathway by gene knock-out has been shown to reduce toxin production, indicating that these two processes are linked in the *C. difficile* cellular regulation networks (Underwood et al., 2009). Such interlinking of sporulation and virulence signaling networks has been observed in other spore-forming bacterial species (Perego & Hoch, 2008). Now that it is possible to genetically manipulate *C. difficile,* future studies are likely to further unravel the role of factors involved in sporulation that can be identified by bioinformatic comparisons. Sporulation-deficient members of random mutant libraries are likely to throw up other genes involved. Such knowledge may facilitate development of therapeutics that can inhibit sporulation and perhaps also toxin production, thereby reducing the virulence of *C. difficile*.

The spore proteome has recently been described (Lawley *et al.*, 2009b). This information will allow the role of proteins known to be present in the spores to be investigated by genetic studies. As well as signaling molecules such as SpoOA, there will be other classes of proteins that are key to the production and integrity of spores. Identification of these proteins will facilitate the rational development of therapies and disinfectants.

1.9.3 Germination factors

Ingestion of *C. difficile* spores leads to germination of spores and subsequent vegetative proliferation of *C. difficile* to establish a *C. difficile* infection. Germination is therefore a pre-requisite of colonisation and factors required for germination can therefore be considered as virulence factors. Preventing germination in the host is an attractive strategy for intervention as it represents a natural bottleneck in the infection process (Nakamura et al., 1985). Also, inducing germination in the environment can aid decontamination, as it makes the organism much more susceptible to antimicrobials (Wheeldon *et al.*, 2008b), therefore germination induction could improve the efficacy of cleaning procedures.

It has long been known that bile salts sodium taurocholate and sodium cholate stimulate spore germination (Wilson *et al.*, 1982, Wilson, 1983). More recently glycine was shown to act as a co-germinant with these cholate derivatives (Wheeldon *et al.*, 2008a, Sorg & Sonenshein, 2008), and kinetic studies suggest that there are distinct germination receptors for taurocholate and glycine (Ramirez et al., 2010) although the identity of germination receptors is not yet known. Interestingly, it has also been shown that deoxycholate, a metabolite of cholate produced by the normal intestinal flora induces germination of *C. difficile* spores but inhibits growth of vegetative cells (Sorg & Sonenshein, 2008). This generated the hypothesis that the metabolism of bile salts by the normal gut flora is one mechanism by which the normal flora protects against *C. difficile* infection, and experiments with antibiotic-treated and untreated mice cecal extracts support this hypothesis

(Giel et al., 2010). Future studies using *C. difficile* mutants aim to identify the germination receptors, and characterise how germinants interact with them.

After germination has been triggered the spore cortex must be broken down, and spore cortex lytic enzymes responsible for this have been identified in *B. subtilis* and *C. perfringens*. Homologues of these lytic enzymes are present in the *C. difficile* genome and by using ClosTron mutagenesis a *C. difficile* enzyme SleC has been shown to be essential for germination of spores in response to taurocholate (Burns et al., 2010). SleC inhibition is therefore a strategy that could prove efficacious in preventing *C. difficile* infection, whilst SleC activation could aid disinfection of spore-contaminated environments. The only other germination factor that has been experimentally identified is *cspBA*, which is a homolog of a germination-specific protease known to be essential for spore germination in *C. perfringens*. The importance of this gene was determined by screening a random mutant library for sporulation/germination deficient mutants (Cartman & Minton, 2010). Many other genes of importance to sporulation and germination are likely to be identified this way now that the technology for making libraries is available, providing new targets for therapeutic intervention.

1.9.4 Colonisation factors

Successful colonisation of the gut requires interactions between *C. difficile*, the host, and the intestinal flora including functions such as adhesion to the epithelium, micro-colony formation, immune evasion, immune manipulation, inhibition of other bacterial species and antibiotic resistance. Therefore it is postulated that a wide variety of factors can be classified as colonisation

factors. The factors required for successful colonisation of the gut by *C. difficile* are as yet largely uncharacterised, but I will summarize in this section factors that have been characterised that are likely to play a role in this process.

1.9.4.1 Antibiotic resistance

The fact that *C. difficile* is resistant to a large number of antibiotics (Delaney, 2007) provides it with a competitive advantage over susceptible organisms when colonising a host undergoing antibiotic therapy. Therefore factors conferring resistance to antibiotics can be considered virulence factors. In many cases these virulence factors are transposon-mediated, such as *ermB*-conferred erythromycin resistance carried on *Tn5398* (Sebaihia et al., 2006) and chloramphenicol resistance carried on *Tn4453* (Lyras et al., 2004). However, sometimes antibiotic resistance is conferred by point mutation, such as mutations in *gyrB* conferring resistance to fluoroquinolones (Drudy *et al.*, 2006b). Fluoroquinolone resistance is spreading at an alarming rate, having been found in a large number of ribotypes (Spigaglia et al., 2008). In general it is important to monitor antibiotic resistance in *C. difficile*, as changes in resistance will need to be accompanied by appropriate changes in clinical practice to reduce the effects of antibiotic resistance on increasing *C. difficile* virulence.

1.9.4.2 Para-cresol production

It has been shown that *C. difficile* can produce para-cresol, which is a phenolic compound with bacteriostatic properties (Dawson et al., 2008). It has been suggested that following antibiotic treatment and *C. difficile* infection of

the gut, para-cresol production by *C. difficile* inhibits re-colonisation of the gut by other bacteria (Dawson et al., 2008). The importance of para-cresol production to *C. difficile* virulence remains to be tested by genetic studies.

1.9.4.3 The *C. difficile* S-layer

C. difficile produces an S-layer, a paracrystalline proteinaceous array that completely coats the bacterial cell wall. The S-layer is primarily composed of two proteins, the high molecular weight S-layer protein (HMW SLP) and the low molecular weight (LMW) SLP, which are cleavage products of the SlpA precursor (Calabi *et al.*, 2001b). Cleavage of the SlpA precursor is carried out by the cysteine protease Cwp84 (Kirby et al., 2009, Dang *et al.*, 2010). Cwp84 is essential for cleavage of SlpA, and *C. difficile* exhibits a marked growth defect *in vitro* upon knock-out of *cwp84* (Kirby et al., 2009). Interestingly, this growth defect can be compensated for by addition of trypsin to the growth medium as trypsin can cleave the SlpA precursor. Despite its clear *in vitro* phenotype the *cwp84* knock-out was found to exhibit the same level of virulence as the wild-type strain in the hamster model of infection, but the role of host proteases in compensating for the lack of *cwp84 in vivo* was not investigated (Kirby et al., 2009). Therefore it is not clear whether SlpA cleavage is essential for *C. difficile* virulence, although in the hamster model of infection Cwp84 is not essential for virulence.

The LMW and HMW SLP form a stable non-covalently associated complex, the low-resolution structure of which was recently solved by small-angle X-ray scattering (Fagan et al., 2009). The HMW SLP is thought to attach to the underlying cell wall via three copies of a putative cell wall binding domain

(Pfam domain PF04122). The LMW SLP is highly immunogenic and exposed to the environment. It is extremely variable between strains, with over 20 distinct sequences identified to date (Calabi & Fairweather, 2002b, Kato et al., 2005, Karjalainen *et al.*, 2002a, Eidhin *et al.*, 2006). A high-resolution structure of the LMW SLP shows that the residues conserved across strains form a core, likely reflecting functional conservation, with the variable regions mostly confined to loops, allowing for immune evasion (Fagan et al., 2009). Based on this knowledge of SlpA a model of *C. difficile* cell wall organization has been proposed and is illustrated in Figure 1.5.

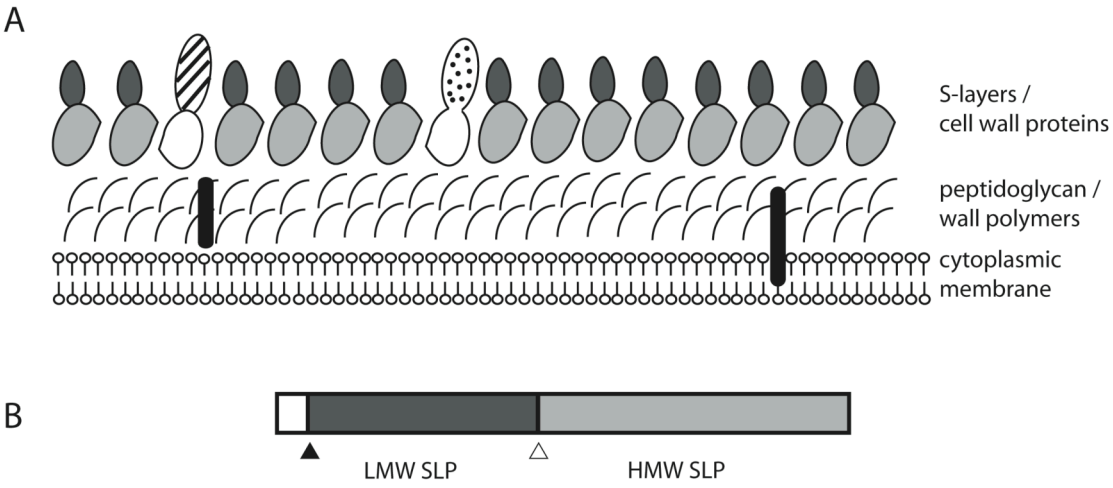


Figure 1.5 Model of the cell wall of *C. difficile*.

A. The HMW SLP (light grey) LMW SLP (dark grey) complex is shown as a layer above the peptidoglycan. Other minor cell wall proteins (CWPs) are shown as white two-lobed structures. Cell wall polymers are shown as black bars. **B.** SlpA undergoes post-translational processing to cleave off the signal peptide (black triangle) and is cleaved by Cwp84 at a specific site (open triangle) to generate the mature LMW and HMW SLP proteins. Figure taken from (Fagan et al., 2009).

Purified HMW SLP protein has been shown to bind to human and mouse intestinal epithelium and anti-HMW SLP antibodies block adherence to HEp2 cells (Calabi et al., 2002). However, our current understanding of the arrangement of the S-layer proteins suggests that the HMW SLP may not be

directly exposed to the host epithelium (Fagan et al., 2009), making the relevance of these binding studies unclear. SLPs have also been shown to induce inflammatory and regulatory cytokines in human monocytes and dendritic cells, suggesting that during infection they may interact with the immune system (Ausiello et al., 2006). The role(s) of SlpA in virulence has not yet been definitively addressed by knock-out studies, and this may be particularly challenging if SlpA is essential for *C. difficile* growth. Therefore, although it appears likely that SlpA will play key roles as a virulence factor, these roles are yet to be definitively elucidated.

1.9.4.4 The *C. difficile* CWP family

In the genome of 630, 28 paralogs of the HMW SLP have been identified (Sebaihia et al., 2006). This family is known as the *C. difficile* cell wall protein (CWP) family. All the members contain two or three copies of the putative cell wall binding domain (PF04122) in addition to a second unique domain that is proposed to specify function. The *C. difficile* CWP family is illustrated in Figure 1.6. Given their homology to the HMW SLP the CWPs may contribute to the S-layer two-dimensional packing interactions, however the interactions between these proteins on the cell surface have not been investigated. Transcriptomic and proteomic studies have shown several of these proteins to be expressed *in vitro* (Wright *et al.*, 2005, Emerson *et al.*, 2008). Antibodies against some of these proteins are found in serum of infected patients (Wright *et al.*, 2008b) implying at least some of the CWPs are expressed *in vivo* during infection and exposed to the host immune system.

The cysteine protease Cwp84 responsible for the cleavage of SlpA (as discussed above) is a member of this CWP family. Biochemical studies using purified recombinant Cwp84 suggested a role for Cwp84 in the degradation of host extracellular matrix proteins, and proposed this as a virulence factor that may degrade host tissue integrity (Janoir et al., 2007). However, as mentioned above a recent study using a Cwp84 knock out strain did not show a key role for Cwp84 in virulence in the hamster model of infection (Kirby et al., 2009). More insight into the role of Cwp84 may come from studying these mutants in different models of infection as the acute nature of the standard hamster model of infection may not be very sensitive in detecting colonisation phenotypes.

Cwp66 is another member of the CWP family whose function has been investigated. It has been shown to be expressed on the cell surface and the C-terminus of the protein is variable between strains (Waligora *et al.*, 2001b). Antibodies raised against recombinant Cwp66 were shown to inhibit adhesion of *C. difficile* to a mammalian cell line, and therefore it was suggested that Cwp66 is an adhesive virulence factor (Waligora *et al.*, 2001a). Given the recent advances in genetic technology this work needs to be followed up with knock-out studies to determine whether or not Cwp66 does pay an important role in *C. difficile* adhesion to host cells and overall virulence.

There is no other published work on the CWP family, other than on CwpV (Emerson et al., 2009), which has been the focus of my PhD. CwpV will be introduced in the next section.

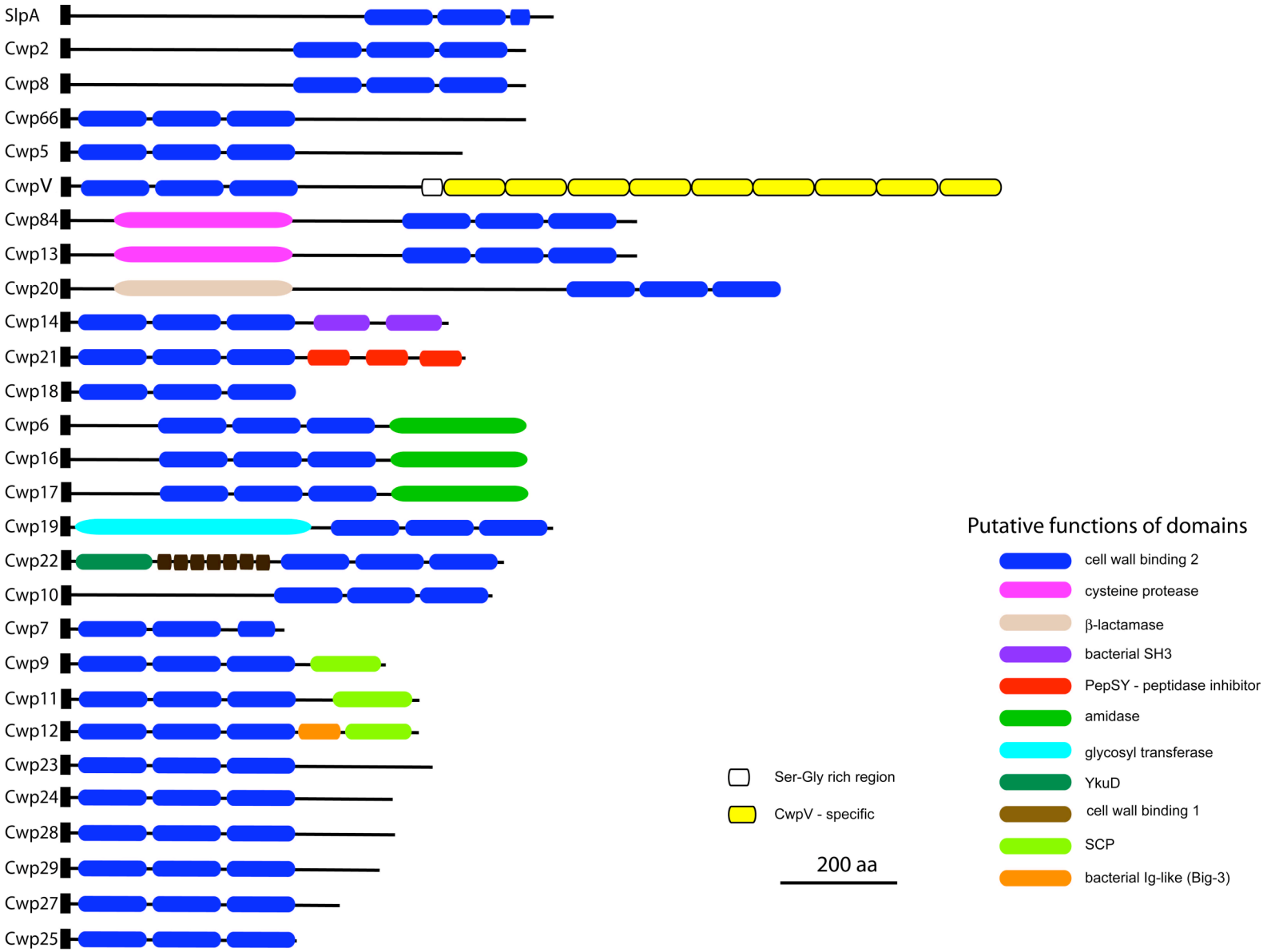


Figure 1.6 Diagrammatic representation of the *C. difficile* CWP family.

1.9.4.5 Other cell surface colonisation factors

A 68 kDa protein expressed on the surface of *C. difficile,* but not part of the CWP family has been named a fibronectin binding protein (Fbp68) based on biochemical detection of interactions between this protein and fibronectin (Hennequin et al., 2003). The relevance of this finding to *C. difficile* virulence has yet to be determined with targeted knock-out studies.

C. difficile has been shown to be motile in swimming assays due to the presence of flagella (Tasteyre *et al.*, 2000a). The flagellin gene *fliC* has been shown to be variable across *C. difficile* strains (Tasteyre *et al.*, 2000b). The flagellin proteins have been shown to be glycosylated, and across strains there also appears to be diversity in the flagellar biosynthetic locus with respect to glycan biosynthesis (Twine et al., 2009). The *fliD* gene encoding the flagellar cap has however been shown to be well conserved across *C. difficile* strains (Tasteyre *et al.*, 2001b). Flagella from a wide range of bacteria have been shown to be important as colonisation and virulence factors. Recombinant *C. difficile* flagellar subunits FliC and FliD, and crude flagellar preparations have been shown to bind to mouse mucus. Recombinant FliD was also shown to bind to a mammalian cell line. Flagellated strains have been shown to exhibit increased adherence to mouse tissue than non-flagellated strains, although there is no clear difference in the overall colonisation of germ-free mice between flagellated and non-flagellated strains (Tasteyre *et al.*, 2001a). These comparisons were carried out using strains of the same serogroup that did or did not express flagella. The role of flagella in *C. difficile* colonisation and virulence could now be further investigated using

isogenic mutants, as this may provide a clearer picture. Testing such mutants in animal models with a gut flora, rather than germ-free mice, may also provide more relevant information as to the role of flagella in *C. difficile* virulence.

A complete set of genes encoding putative type IV pili is present in the genomes of 630 and the ribotype 027 strain R20291. In many pathogens, type IV pili are involved in aspects of pathogenesis such as motility and adhesion (Mattick, 2002). Using antibodies against recombinant PilA proteins pili-like structures on the surface of *C. difficile* in close association with host tissues have been visualised in the hamster model of infection using immuno-gold electron microscopy (Goulding et al., 2009). This suggests that type IV pili are expressed *in vivo* and may be important colonisation factors, but further work is required to determine the role of type IV pili in *C. difficile* infection.

1.10 The *C. difficile* cell wall protein CwpV

For my PhD I have studied a member of the CWP family, CwpV. This is the largest member of the CWP family encoded in the 630 genome (see Figure 1.6). Initial characterisation of CwpV had been carried out prior to the start of my project, which identified this protein as an interesting candidate for further study. Here I summarise the preliminary findings that led into my project.

1.10.1 The *cwpV* gene in *C. difficile* 630

CwpV (CD0514, Cwp15) is a paralogue of the HMW-SLP. The *cwpV* gene is not situated in the *slpA* gene cluster, where many of the other CWPs are

situated (Calabi et al., 2001b). An N-terminal signal sequence is predicted by SignalP 3.0 (Bendtsen et al., 2004). In 630 the *cwpV* gene encodes a predicted protein of 167 kDa consisting of three distinct domains:

(1) An N-terminal HMW-SLP homology region incorporating three copies of the PF04122 motif (putative cell wall binding domain); (2) A region of unknown function terminating in a serine-glycine-rich flexible linker; (3) Nine repeats of 120 amino acids These repeats are almost perfect copies of one another with the exception of the first repeat, which is more divergent. A cartoon representation of the 630 CwpV protein is shown in Figure 1.7.



Figure 1.7 The CwpV protein in *C. difficile* strain 630.

Cartoon representation of the CwpV protein from 630. The predicted signal peptide is shaded black. The three Pfam04122 domains are shaded grey. The domain of unknown function is white. The serine-glycine-rich linker is pink. The nine repeats are shown in blue.

The CwpV protein has been previously detected in the cell wall of 630 (Wright et al., 2005), and is annotated as a cell surface putative haemagglutinin/adhesin protein (Sebaihia et al., 2006). This is due to some weak homology of the repetitive domain to proteins exhibiting these functions. No Pfam domains are detected in the region of unknown function or the repeat domains.

1.10.2 CwpV C-terminal variability across strains

All 75 strains so far tested by microarray contain one copy of the N-terminal cell wall binding domain of CwpV, but the C-terminal repeat region is not well-conserved (Stabler et al., 2006b). The *cwpV* genes from a number of different

strains have been sequenced. All *cwpV* genes characterised encode 4-9 repeating units of between 79-120 amino acids at the C-terminus of the protein. These repeating units can be grouped into 5 “types” based on differences in amino acid sequences of the repeating units. Some strains examined exhibit a mosaic-like repeat structure containing two different types of repeat. A diagrammatic representation of these five types is shown in Figure 1.8.

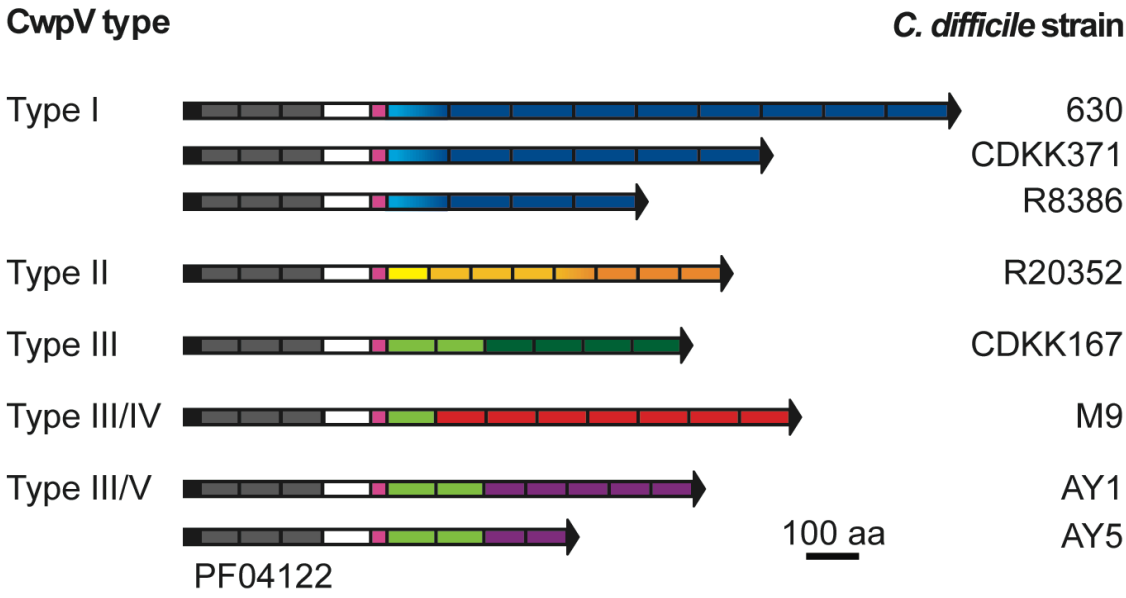


Figure 1.8 Sequence variation of CwpV across *C. difficile* strains.

The N-terminus up to the end of the serine-glycine-rich linker is conserved across diverse *C. difficile* strains. Different CwpV C-terminal repeat types are shown in different colours (type I – blue; type II - yellow; type III – green; type IV – red; type V – purple).

Therefore *cwpV* is a highly variable gene between *C. difficile* strains, encoding CwpV with a conserved N-terminus and a variable C-terminus. The significance of this C-terminal variation to CwpV antigenicity and function had not been investigated when I started this study.

1.10.3 CwpV is expressed in a phase variable manner in 630

CwpV is expressed on the bacterial surface in a phase variable manner. Antibodies raised against the Type I repeats (anti-CwpVrptI), which are those found in 630, were used to show that expression of CwpV within a clonal population of lab-cultured 630 is only detectable in a small proportion of cells. Evidence for the phase variable expression of CwpV in 630 is shown in Figure 1.9.

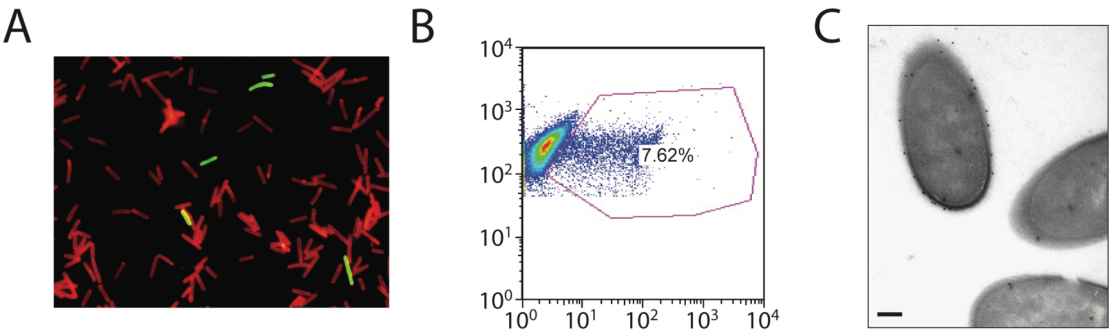


Figure 1.9 Detection of CwpV phase variable expression.

A. A clonal population of *C. difficile* 630 was labelled with anti-CwpV (green) and anti-HMW SLP (red) antibodies for fluorescent microscopy. Only a minority of cells labelled with the anti-CwpV antibody, indicative of phase variable expression of CwpV. **B.** Flow cytometry analysis of 630 stained with anti-CwpV antibody. 7.62% of cells are labelled, representing the CwpV-expression population. **C.** Analysis of gold-labelled anti-CwpV stained cells by electron microscopy. This confirms that only some cells are labelled and that in those cells CwpV is surface expressed.

1.10.4 Mechanism of CwpV phase variation

This phase variation has been shown to be mediated by inversion of 195 bp flanked by imperfect 21 bp inverted repeats (IRs). The inversion region, referred to as the *cwpV* switch, is located approximately 60 bp upstream of the translational start codon. The organization of this region in 630 and PCR

evidence of the inversion of the *cwpV* switch in a number of *C. difficile* strains is shown in Figure 1.10. The repeat sequence has not been found at any other location in the *C. difficile* genome sequences available. Therefore it appears that this switch only controls the expression of CwpV.

The mechanism by which one orientation results in expression of CwpV (ON) and one orientation precludes expression of CwpV (OFF) had not been determined at the start of my project. The IRs flanking the inversion region are characteristic of substrates for site-specific recombinases. The identity of the recombinase(s) responsible for catalysis of the inversion was not known at the start of my project. In previously studied systems where phase variation is mediated by site-specific recombination the inversion region either contains the promoter or is present in the coding region of the gene (van der Woude & Baumler, 2004). In *cwpV* the inversion region is upstream of the coding region and therefore inversion must mediate phase variation either by inverting the promoter or by a novel mechanism.

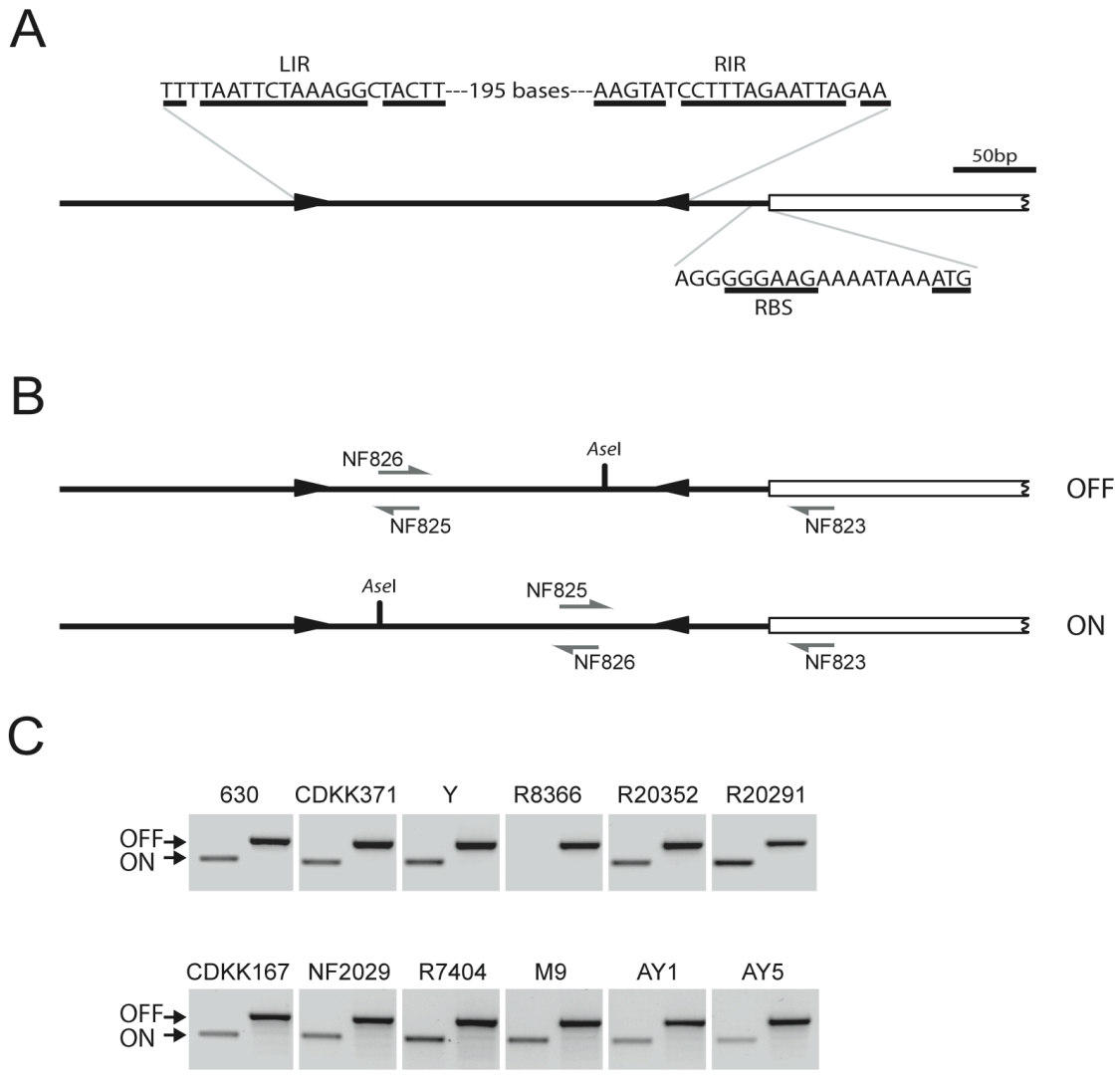


Figure 1.10 The *cwpV* switch.

A. Diagram of the location of the *cwpV* switch upstream of the *cwpV* gene in 630. The sequences of the IRs (LIR and RIR) are shown, with the perfectly repeated sequence underlined. The locations of the ribosome binding site (RBS) and ATG start codon are also shown. **B.** Orientation-specific PCR using primers NF823 and NF825 (ON) or NF826 (OFF) allows detection and distinction between the two orientations of the *cwpV* switch. C. Orientation PCR analysis of a panel of *C. difficile* strains detects both orientations of the *cwpV* switch in all strains tested, with the exception R8366, which carries a 2bp deletion in the LIR.

1.10.5 Function of CwpV

Despite discovery of the unusual regulation of CwpV expression, previous work had given little insight into CwpV function. The 630 *cwpV* gene is annotated as a cell surface putative haemagglutinin/adhesin protein (Sebaihia et al., 2006). In general bacterial pathogens are well known for expressing surface localised and secreted proteins containing repetitive sequences. *S. aureus* expresses fibronectin binding proteins (FnBPA) (Jonsson et al., 1991), clumping factor (ClfA) (McDevitt & Foster, 1995) and Protein A (Sjodahl, 1977), which all contain repetitive sequences that are involved in binding to host cellular components. *Streptococcus pyogenes* M protein contains repeats that mediate binding to fibronectin (Hollingshead et al., 1986). The family of large clostridial toxins (von Eichel-Streiber et al., 1996) also contain repetitive sequences in their C-terminal domains that mediate binding to host cell receptors. The presence of multiple repeats in these proteins increases the avidity of binding to the host receptor (Ingham et al., 2004). Given the repetitive nature of CwpV it was postulated that it may have a ligand-binding function, however no experimental evidence supporting this hypothesis had been obtained at the start of my project.

1.11 Project Aims

The current understanding of the molecular basis for *C. difficile* pathogenesis is limited. A better understanding of *C. difficile* surface proteins and the roles they play during host colonisation may aid development of interventions that aim to reduce or prevent colonisation of hosts by *C. difficile*.

Observation of phase variable expression of CwpV, the finding that this regulatory mechanism appears to be conserved across a diverse set of *C. difficile* strains, and the variation of CwpV repeat domains exhibited by different strains of *C. difficile* prompted a detailed molecular study into CwpV regulation, expression, strain-strain variation and function.

The major aims for this project were:

1. To characterise the molecular mechanism responsible for CwpV phase variation
2. To identify the site-specific recombinase(s) required for inversion of the *cwpV* DNA switch
3. To study CwpV protein processing and expression in the S-layer
4. To assess the level of conservation of the regulation and processing of CwpV across strains of different CwpV types
5. To determine whether different CwpV repeat types are antigenically distinct
6. To investigate the function of CwpV

In the subsequent chapters results are presented addressing these aims.

2 Materials and Methods

2.1 Bacterial strains and culture

2.1.1 *C. difficile* strains

C. difficile 630 was provided by Dr Peter Mullany, Eastman Dental Institute, London, and has been fully sequenced by the Wellcome Trust Sanger Institute. R20352 was obtained from Prof. Brendan Wren and was isolated from patients from an outbreak in Quebec (Canada). CDKK167 was provided by Professor C. P. Kelly, Harvard Medical School, Boston, USA and was isolated from a patient with a single episode of diarrhoea that responded well to metronidazole treatment. M9 was obtained from Dr. Denise Drudy, Dublin, Ireland and was isolated from a symptomatic patient in a $TcdA^{-}B^{+}$ outbreak. AY1 was provided by Dr Dale Gerding, it is a non-toxic strain isolated in Minnesota, USA. An erythromycin sensitive derivative of 630 (630Δ*erm*) (Hussain et al., 2005) was used for creation of mutants using ClosTron and therefore as a wild-type strain when carrying out phenotypic assays with mutant strains.

2.1.2 *C. difficile* culture conditions

C. difficile was routinely cultured either on blood agar base II (Oxoid) supplemented with 7% horse blood (TCS Biosciences); on brain-heart infusion (BHI) agar plates (Oxoid) or in BHI broth (Oxoid). Culture was undertaken in an anaerobic cabinet (Don Whitley Scientific) at 37ºC in an anaerobic atmosphere (10% CO_2, 10% H_2, 80% N_2). Long-term *C. difficile*

stocks were samples of overnight BHI cultures frozen at -80ºC in 15% glycerol.

2.1.3 *E. coli* strains

Commercial chemically competent TOP10 cells (Invitrogen) were used for initial transformation after cloning and were used for plasmid maintenance. Protein expression was carried out in either BL21(DE3) or Rosetta (Novagen) cells. Rosetta are BL21(DE3) cells carrying a plasmid encoding tRNAs specific to codons that are rare in *E. coli* but common in other organisms, dramatically enhancing protein expression from Clostridial genes.

2.1.4 *E. coli* culture conditions

E. coli strains were routinely grown at 37ºC on LB-agar plates or in LB liquid culture in the presence of selective antibiotics where appropriate to any transformed plasmid (50 µg/ml ampicillin, 50 µg/ml kanamycin, 30 µg/ml chloramphenicol). For long term storage, *E. coli* strains were grown overnight in LB, then made up to 15% glycerol and frozen at –80ºC.

2.1.5 Chemical transformation of *E. coli*

Purified plasmid DNA or an entire ligation reaction was mixed with a freshly thawed aliquot of chemically competent cells and incubated on ice for 30 min. The transformation tubes were then transferred into a 42ºC water bath for 30 s before returning to ice for 2 min. 250 µl SOC medium (Invitrogen) was then added to the cells and the tubes incubated with shaking at 37ºC for 1 h before spreading onto selective LB-agar plates.

2.1.6 Production of chemically competent *E. coli*

An overnight culture of *E. coli* in LB media was sub-cultured with shaking at 37°C and its optical density at 600nm (OD_{600}) monitored using an Ultrospec cell density meter (Amersham). When the cultures were in the exponential growth phase (OD_{600} 0.4-0.6) they were centrifuged at 3,400 x g for 10 min at 4°C and the pellet re-suspended in 5ml ice-cold 100mM $CaCl_2$. After 15 min on ice the sample was centrifuged at 3,400 x g for 10 min at 4°C and re-suspended in ice-cold 1 ml $CaCl_2$ containing 15% glycerol. The cells were incubated on ice for 2 h before freezing at -80°C. All steps except centrifugation were performed in a 4°C cold room.

2.1.7 *Clostridium perfringens* strain and culture conditions

Clostridium perfringens strain SM101 was kindly provided by Bruno Dupuy and was routinely cultured either on blood agar base II (Oxoid) supplemented with 7% defibrinated horse blood (TCS Biosciences), BHI agar (Oxoid) or in BHI broth (Oxoid). Cultures were grown in an anaerobic cabinet (Don Whitley Scientific) at 37°C in an anaerobic atmosphere (10% CO_2, 10% H_2, 80% N_2).

2.1.8 *E. coli-Clostridium* conjugation

1ml of an LB 30µg/ml chloramphenicol overnight culture of *E. coli* CA434 containing a pMTL960-derived plasmid was pelleted at 1700 x g for 10 min at 4 °C. The pellet was washed gently with 1 ml LB. Cells were pelleted again, the supernatant discarded and the pellet transferred to an anaerobic cabinet. 200 µl of an overnight *Clostridium* culture was used to resuspend the CA434 cell pellet. The resulting mixed bacterial suspension was spotted out onto a BHI agar plate and incubated at 37 °C in the anaerobic cabinet for 8 h to allow

conjugation to proceed. Then 1 ml of BHI was used to resuspend bacteria from the agar plate. Appropriate dilutions (typically neat, 1:5, 1:25) of this suspension were plated onto BHI plates supplemented with 15 µg/ml thiamphenicol to select for pMTL960-based plasmids and 250 µg/ml cycloserine to inhibit *E. coli* growth. Plates were incubated overnight at 37°C, in anaerobic conditions. Transconjugants were picked and streaked onto fresh selective plates (15 µg/ml thiamphenicol 250 µg/ml cycloserine) to single colony purify transconjugants and ensure no carry over of *E. coli* from the conjugation.

2.2 Bioinformatics and computer programs

2.2.1 Genome visualization

The genome sequence of 630 and its annotation were obtained from The Wellcome Sanger Institute website (Sebaihia et al., 2006) (Sebaihia et al., 2006, http://www.sanger.ac.uk/Projects/C_*difficile*/). The software Artemis (The Sanger Institute, http://www.sanger.ac.uk/Software/Artemis/) was used to visualise the genome sequences and annotation.

2.2.2 DNA sequencing and alignments

DNA sequencing chromatograms were visualized using Bioedit (Ibis biosciences). The Vector NTI suite of programs (Invitrogen) was used to create plasmid maps, construct contigs from multiple sequencing reactions and carry out alignment analysis after cloning.

2.2.3 Primer design

Annealing temperature and secondary structure of primers were predicted using OligoCalc (http://www.basic.northwestern.edu/biotools/oligocalc.html).

2.2.4 Nucleic Acid Folding and Hybridization prediction

Predictions of free energy of RNA folding were carried out using the Mfold web server (http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi).

2.2.5 Homology Searches

Homology searches were carried out using BLAST (basic local alignment search tool). Searches against all known protein or nucleotide sequences were carried out using BLASTP and BLASTN respectively (http://www.ncbi.nlm.nih.gov/BLAST/). Searches against *C. difficile* 630 sequences were carried out on the website of The Sanger Institute (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/c_*difficile*).

2.2.6 Gene and protein property predictions

Softberry BPROM (prediction of bacterial promoters) was used to predict the site of initiation of transcription (http://www.softberry.com). Basic properties of proteins, such as molecular weight and pI were calculated using tools from the ExPASy website (http://www.expasy.ch/tools). Searching for presence of repeats in proteins was carried out using RADAR (http://www.ebi.ac.uk/Tools/Radar/index.html).

2.3 DNA manipulation

Full details of recombinant plasmid production are given in Chapter 3. All recombinant plasmids used in this study are detailed in Table 2.3.

2.3.1 *C. difficile* genomic DNA preparation

Genomic DNA was isolated from 10 ml BHI *C. difficile* overnight liquid cultures. The Gentra Puregene Yeast/Bact. Kit (Qiagen) was used according to manufacturers instructions.

2.3.2 Chelex preparations

As an alternative to genomic DNA for PCR template, colonies of *C. difficile* grown on agar plates were subjected to chelex preparation. Chelex 100 resin (Bio-Rad) was prepared at 0.05g/ml in water. *C. difficile* colonies were resuspended in 100µl prepared chelex or 10µl of overnight liquid culture was added to 90µl prepared chelex. Samples were boiled for 10 min at 100°C then centrifuged at 20,000 x g for 10 min at 4°C. Supernatants were recovered and either used as PCR template immediately or stored at -20°C.

2.3.3 Polymerase chain reaction (PCR) primers

Primers used in this study were ordered for synthesis by Eurofins MWG Biotech and are described in Table 2.2.

2.3.4 Diagnostic PCR

For purposes where fidelity of the DNA polymerase was not critical, for example colony PCR to determine presence of vector inserts after ligation, Taq DNA polymerase was used. Redtaq polymerase kit (Sigma) was used containing a master mix of $MgCl_2$, dNTPs, Taq polymerase and loading buffer.

Primers were used at 10 pmol/µl. 1 µl of forward primer, 1 µl of reverse primer, 1 µl template and 7 µl RNase free water were added to 10 µl of Redtaq solution to make a final volume of 20 µl. Thermal cycler parameters were 95ºC, 5 minutes; thirty cycles of denaturation (95ºC, 30 s), annealing (25 s) and extension (72ºC, 1 minute per kb of the desired PCR product); and a final 10 min 72ºC incubation. Annealing temperature was generally 5ºC lower than the predicted melting temperature (T_m) of the primers used. If initial PCR was unsuccessful a set of PCRs with a gradient of annealing temperatures was set up to identify the optimum annealing temperature.

2.3.5 High fidelity PCR

For cloning reactions, where fidelity is of high importance a proofreading polymerase KOD Hot-start DNA polymerase (Novagen) was used. Thermal cycler parameters were 95ºC, 2 min; 25-35 cycles of: Denaturation (95ºC, 15 s); Annealing (temperature appropriate for primers, 30 s) and Extension (68-72ºC, 20-30s per kb of the desired PCR product). Annealing temperature was generally 5ºC lower than the predicted melting temperature (T_m) of the primers PCR, or optimized as described above.

2.3.6 PCR product purification

If a single band was visualized on an agarose gel the product was purified using a QIAquick PCR purification spin kit (Qiagen). If there were multiple bands present and a particular band was desired for purification it was purified from the gel. The band was excised using a scalpel and QIAquick Gel Extraction Kit (Qiagen) was used according to the manufacturer's instructions.

2.3.7 Agarose gels

PCR reactions were run on 0.8 - 2.0% TAE-agarose gels (TAE: 40mM Tris, 20mM acetic acid, 1mM EDTA) containing SYBR Safe DNA gel stain (Invitrogen). DNA was visualized using a Gel Doc system (Bio-Rad).

2.3.8 TOPO Cloning

A TOPO TA Cloning Kit (Invitrogen), which clones the insert using topoisomerase bound to the ends of the open vector, was used for primary cloning and also cloning for sequencing. Briefly, after PCR amplification of DNA fragment of interest with a proofreading polymerase, 0.2µl Taq (Sigma) was added and the reaction incubated at 72ºC for 10 min to add A-overhangs at the end of PCR products. The required PCR product was then subjected to either QIAquick PCR Purification or QIAquick Gel Extraction. 2µl cleaned amplification product was then mixed with 0.5µl TOPO kit salt solution and 0.5µl pCR4-TOPO vector (Invitrogen) and incubated at room temperature for 10 min before immediate transformation into TOP10 *E. coli* (Invitrogen).

2.3.9 Restriction enzyme digestion

For subcloning or cloning of PCR products into a specific vector DNA was digested with restriction endonucleases. All restriction enzymes were obtained from NEB and digestion conditions set up as recommended by the manufacturer.

2.3.10 Purification and Quantification of digested DNA

For digested PCR products the QIAquick PCR purification spin kit (Qiagen) was used to purify the DNA for ligation. For digested vectors where there

were two or more products produced from digestion the desired fragment was purified from an agarose gel. The band was excised using a scalpel and QIAquick Gel Extraction Kit (Qiagen) was used according to the manufacturer's instructions. DNA concentration was quantified using a NanoDrop Spectrophotometer (Thermo Scientific).

2.3.11 DNA Ligation

Following digestion, purification and quantification of the two fragments of DNA required for ligation 50ng of digested DNA vector was mixed with the insert at a molar ratio of 1:1 - 1:10. Freshly thawed DNA ligase buffer containing ATP was mixed with the DNA and Quick Stick ligase (Bioline). The reaction was incubated at room temperature for 15 min and then transformed into chemically competent TOP10 *E. coli* (Invitrogen).

2.3.12 Colony PCR

Colonies containing inserts in the vector were selected via colony PCR, using standard Taq PCR conditions, with an additional 10 min boil step at the start to promote lysis of the bacterial cells. Colonies were subjected to PCR with primers annealing either side of the insertion site, to allow distinction by size between empty vector and vector containing inserts.

2.3.13 Plasmid purification

Plasmids were purified from an overnight culture of *E. coli* grown in LB with the appropriate antibiotics to maintain the plasmid. Plasmids were purified using a QIAprep Spin Miniprep Kit (Qiagen) according to manufacturers instructions.

2.3.14 DNA Sequencing

Sequencing of purified plasmid DNA at a concentration of 50-100ng/μl was carried out by GATC Biotech (Konstanz, Germany) using dideoxy chain terminator DNA sequencing. Sequencing primers are provided at 10pmol/μl or ordered for synthesis by GATC.

2.4 Protein manipulation

2.4.1 Recombinant protein expression in *E. coli*

Vectors (pET28a or pACYCY-Duet-1) containing the genes of interest under the control of a lactose inducible promoter were used for recombinant protein expression. LB cultures of BL21(DE3) or Rosetta *E. coli* containing these plasmids were grown for 6 h with the appropriate antibiotic selection to maintain the plasmids. These cultures were then used to inoculate Overnight Express Autoinduction medium (Novagen). The cultures were grown overnight at 37°C with shaking to saturation of the medium. The cells were harvested by centrifugation and the pellet resuspended in lysis solution (50mM NaH_2PO_4 , 300mM NaCl, 10mM imidazole containing 10% BugBuster Protein Extraction Reagent (Novagen), 1% DNase and 1% lysozyme). Lysis was allowed to proceed for 30 min at room temperature with constant gentle mixing. Following lysis the suspension was centrifuged at 14,000rpm in a tabletop centrifuge for 5 min. The supernatant was harvested as it contains the soluble fraction of protein for purification.

2.4.2 Purification of His-tagged proteins

The pH of the supernatant from cell lysis was measured and if necessary adjusted to pH8. This supernatant containing the His-tagged protein was then applied to a gravity flow column containing Nickel-NTA agarose (Qiagen). The supernatant was incubated with the column for 1 h at 4°C. The solution was then allowed to flow through the column and the flow through was collected. The column was then washed with 5 column volumes of wash buffer (50mM NaH_2PO_4, 300mM NaCl, 10mM imidazole). Wash fractions were collected and A_{280nm} monitored to determine when washes have removed most contaminating protein. Finally 5 column volumes of elution buffer (50mM NaH_2PO_4, 300mM NaCl, 250mM imidazole) were applied to the column and fractions collected. A_{280nm} was used to monitor elution and if necessary more elution buffer was added to elute all protein from the column. Fractions of interest were then run on SDS-PAGE gel to determine which fractions contained purified protein and allow estimation of purity.

2.4.3 Dialysis of Purified Protein

In order to transfer purified protein from elution buffer following purification to a buffer more suitable for long-term storage dialysis was carried out. Fractions containing purified protein were pooled and protein transferred to a Slide-A-Lyzer Dialysis Cassette (Thermo Scientific). In general the dialysis buffer was 10mM HEPES, 150mM NaCl. The dialysis cassette was floated in 2 litres of pre-cooled dialysis buffer at 4°C for 2 h with stirring. The buffer was then changed and the dialysis left overnight with stirring at 4°C. The buffer was changed once more and left for a further 6 h before protein was retrieved from

the cassette. Recovery of proteins was confirmed by SDS-PAGE prior to determination of protein concentration.

2.4.4 Determination of protein concentration

Protein concentration was accurately determined using a BCA assay (Pierce) carried out according to the manufacturer's instructions. Briefly, standards of BSA of known concentration and samples were serially diluted in PBS. Reagents A and B were mixed in a 50:1 ratio, then 180µl reagent was added to 20µl sample. The reaction was incubated for 30 min at 37ºC, then the absorbance at 595nm determined using a Benchmark microplate reader (Bio-Rad).

2.4.5 *C. difficile* surface layer preparation

C. difficile cells from 50ml overnight cultures were harvested after OD_{600} measurement by centrifugation at 6000 x g for 10 min. Pellets were washed with 5ml sterile PBS and cells harvested again. Pellets were resuspended in 500µl 0.2M glycine–HCl pH 2.2 and incubated at room temperature with rotation for 30 minutes. Preparations were then centrifuged at 20,000 x g in a microcentrifuge for 15 minutes at 4°C. Supernatants were recovered and pH adjusted to pH 6–8 with 2M Tris, adding 10 µl at a time and monitoring pH with pH strips (Whatman).

2.4.6 Expression of Strep-tagged proteins in *C. difficile* and StrepTactin purification

C. difficile strains containing pMTL960-based expression plasmids were grown in static BHI culture overnight with 15 µg/ml thiamphenicol to ensure

maintainance of the plasmids. Surface layer preparations from the overnight culture were carried out as described above. Surface layer preparations were mixed 1:1 with StrepTactin wash buffer (150mM NaCl, 100mM Tris-HCl, 1 mM EDTA, pH 8.0). 100 µl of StrepTactin resin (50 µl bed volume) (Novagen) was washed twice with 0.5 ml dH_2O, then twice with 0.5 ml of wash buffer. The S-layer preparation/wash buffer mixture was then incubated with the resin overnight at 4°C with rotation. The next day, the mixture was centrifuged for 10 s at 20,000 x g to pellet the resin. The resin was washed with 1 ml wash buffer to remove non-specifically bound proteins. In order to determine the total number of washes, washes were monitored by A_{280nm} to determine when protein concentration of washes had returned to a baseline, and therefore most contaminating protein had been washed off. Elutions were carried out by addition of 50µl of elution buffer (150mM NaCl, 100mM Tris-HCl, 1 mM EDTA, 2.5mM desthiobiotin, pH 8.0) followed by centrifugation to pellet the resin and removal of the supernatant. In order to determine the total number of elutions, elutions were monitored by A_{280nm} to determine when protein concentration of elutions had returned to a baseline, and therefore most protein had been eluted. All fractions were collected for analysis via SDS-PAGE.

2.4.7 SDS-PAGE

SDS-PAGE gels a fixed concentration of acrylamide (8 – 15%) were prepared using mini-protean kits (Bio-Rad). The resolving gel (8 – 15% 39:1 acrylamide:bisacrylamide, 375mM Tris pH 8.8, 0.1% SDS, 0.5% APS, 0.07% TEMED) was poured and a layer of isopropanol was gently poured on top. Once the gel was set, the isopropanol was discarded. The stacking gel (5% 39:1 acrylamide:bisacrylamide, 62.5mM Tris pH 6.8, 0.1% SDS, 0.1% APS,

0.24% TEMED) was poured on top of the resolving gel and the comb put in place to form the wells. Once set, the comb was removed and the gel was immersed in electrophoresis buffer (25mM Tris, 192mM glycine, 0.1% SDS). Protein samples were diluted in sample buffer (final concentration 45mM Tris pH 6.8, 0.7% SDS, 0.7mg/ml sucrose, 3.5% β-mercaptoethanol, trace bromophenol blue). These, along with a 10µl aliquot of Broad Range Protein Marker (NEB), were heated to 100ºC for 5 minutes and then briefly centrifuged. Samples were loaded into the wells of the gel and a constant voltage of 200V applied for 40-50 min. Gels were then either Coomassie stained or transferred to PVDF membranes for immunoblotting.

2.4.8 Coomassie staining

For visualisation of proteins following SDS-PAGE, gels were stained in Coomassie blue (45% methanol, 10% acetic acid, 1mg/ml Coomassie blue stain) for 4 - 16 h, followed by destaining using 45% methanol, 10% acetic acid. A final destain in 10% acetic acid was carried out prior to scanning of gels for documentation.

2.4.9 N-terminal Sequencing

S-layer extracts were separated on 10% SDS-PAGE gel and transferred onto a PVDF membrane (see Western immunoblotting for transfer protocol). After protein transfer the membrane was washed in water for 10 min, and stained in 0.1% Coomassie Blue R in 50% MeOH for 5 min. After destaining in 50% MeOH, 10% acetic acid (3x 5 min with rocking) and washing in water (3x 5 min with rocking) the membrane was allowed to dry thoroughly. Bands requested for sequencing were indicated and the membrane sent to the

Protein and Nucleic Acid Chemistry Facility, Department of Biochemistry, University of Cambridge. Edman degradation was carried out to determine the first 5 amino acids of protein bands excised from the PVDF membrane.

2.4.10 Production of antibodies

Anti-CwpVNter and anti-CwpVrptI were produced prior to this study as reported (Emerson et al., 2009). For raising anti-CwpVrptII-V antibodies recombinant CwpV repeat domains were purified from *E. coli* and dialysed into 10mM HEPES, 150mM NaCl for use as antigens. Samples were sent to Covalab for immunization of rabbits according to their standard protocols. For each antigen two rabbits were immunized. Pre-immunisation serum and post-immunisation serum from all rabbits were analysed to confirm successful production of specific antibodies. The characteristics of the antigens used for production of these anti-CwpV antibodies is described in Table 2.1.

Table 2.1 Description of recombinant antigens used to raise antibodies

Antibody	***C. difficile* strain**	**No. of repeats**	**Size (bp)**	**MW (kDa)**	**Amplification primers**
CwpVNter*	CD630	n/a	1365	50	664/5
rptI*	CD630	1	207	7.6	345/6
rptII	R20352	2	506	18.5	1021/22
rptIII	CD167	2	560	20.5	1023/4
rptIV	M9	1	318	11.5	1025/6
rptV	AY1	1	280	10	1027/8

*These antibodies had been raised previously in our laboratory – information provided for completeness

2.5 Immunodetection

2.5.1 ELISA

96-well MaxiSorp PS Immuno F96 plates (Nunc, Roskilde, Denmark) were coated overnight at 4ºC with 50µl protein of interest at 10ug/ml. After three washes with PBS + 0.1% Tween-20 (PBS-T) the plates were blocked with 100µl PBS-T + 3% dried skimmed milk (milk) per well at 37ºC for 2 h. Plates were washed three times with PBS-T. Subsequent incubation steps with protein of interest or antibody were carried out at the appropriate concentration made up in PBS-T + 3% milk at room temperature for 1 h. Three PBS-T wash steps were carried out between each incubation. After final incubation with HRP-conjugated antibody and three washes with PBS-T, 50µl OPD substrate (5ml 0.1M citric acid, 5ml 0.2M Na_2HPO_4, 10ml H_20, 1 OPD tablet (Sigma), 8µl H_20_2) was added to each well and incubated on the plate at room temperature for up to 1 h. The reaction was stopped using 20µl 3M H_2SO_4 per well and the A_{492nm} was determined using a microplate reader.

2.5.2 Western immunoblotting

SDS-polyacrylamide gels were transferred to PVDF Immobilon-P Transfer membrane (Millipore) using a semi-dry protocol. SDS-PAGE gel and filter paper were pre-soaked in transfer buffer (25mM Tris, 192mM glycine, 20% methanol). Immobilon-P membrane was prepared for transfer by incubation sequentially in methanol, H_20 and transfer buffer according to manufacturer's protocol. The filter paper, membrane and gel were stacked in Trans-blot SD semi-dry transfer cell (Bio-Rad) and run at 25V for 15-30 min. Transferred proteins were visualised by staining with Ponceau S solution (Sigma) for 1

min and destained using H_2O. Locations of protein marker bands were marked in pencil. The membrane was then transferred to methanol for 1 min then allowed to dry completely. Antibody used for detection was then incubated with the membrane with shaking for 1 h at an appropriate dilution in 3% milk PBS-T. For His-tag detection anti-His antibody HRP Conjugate (Abcam) was used for detection. For strep-tag detection mouse anti-*Strep*-tag II (Novagen) was used. For non-conjugated antibodies a secondary HRP-conjugated antibody was required (goat anti-rabbit HRP (Dako) or goat anti-mouse HRP (Dako)). Between antibody incubation steps three washes in PBS-T were carried out. After final antibody incubation the blot was washed in PBS-T and then freshly prepared SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) was added to the membrane and allowed to develop for up to 5 min at room temperature. Chemiluminescence was detected using a LAS-3000 Image Reader (Fujifilm) and a white picture also taken to allow protein markers to be visualised.

2.5.3 Immunofluorescence microscopy

C. difficile cells from 800µl of a culture were harvested by centrifugation at 8000 rpm for 1 min. The supernatant was discarded and the pellet washed twice with 1 ml PBS. For fixation, pellets were then resuspended in 8% formaldehyde (made fresh from 36% stock solution (Sigma) diluted in PBS) and incubated for 15 min at room temperature. Fixed cells were pelleted and washed twice in 1 ml PBS by centrifuging at 8000 rpm for 2 min. After the final wash pellets were resuspended in 500µl 20 mM NH_4Cl_2 and incubated for 15 min at room temperature. Cells were pelleted at 8000 rpm for 2 min then washed twice with 1 ml PBS. The final wash supernatant was discarded

and pellets were resuspended in 40 µl 1% BSA in PBS and left for 10 min. Primary antibody was added to make the appropriate dilution (for anti-CwpV antibodies 1:20) and incubated with the cells for 45 min. Cells were then washed twice with PBS and once with 1% BSA in PBS. Again the pellet was resuspended in 40µl 1% BSA in PBS and fluorescent secondary antibodies added to the appropriate dilution (anti-rat and anti-rabbit Cy2/rhodamine redX conjugated antibodies from Jackson Immunoresearch, used at 1:200 dilution as recommended). Tubes were incubated in the dark for 45 min. Cells were washed twice with PBS, then twice with dH_20 before resuspending in 40µl dH_20. Stained cell suspension was then spotted onto a SuperFrost (Thermo Scientific) microscope slide and allowed to air dry in the dark. Once dry coverslips were mounted on top of the spot of cells using prolong gold anti-fade reagent (Invitrogen). Bacteria were visualized using a Nikon Eclipse E600 microscope fitted with a Nikon DMX1200 camera.

2.6 Reporter Assays

2.6.1 FbFP reporter assays

Cells from *C. difficile* cultures were harvested and washed with PBS, then resuspended in PBS to an OD_{600} of 1.0. Samples were transferred to black FluoroNunc 96-well plates (Nunc) then fluorescence of each well measured with a spectrophotometer (Cary Eclipse), using an excitation wavelength of 450nm and an emission wavelength of 495nm.

2.6.2 GusA reporter assays - *C. difficile*

The OD_{600} values of growing *C. difficile* cultures were monitored and from mid-logarithmic growth phase cultures 10ml samples were taken, bacterial cells were harvested by centrifugation and the pellet frozen at -80ºC. The pellet was then thawed and resuspended in 0.8ml Z buffer (60 mM NaH_2PO_4, 60 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50mM 2-mercaptoethanol, pH7). The cell suspensions were transferred to tubes containing 150-212μm diameter unwashed glass beads (Sigma) and the tubes treated in a BIO101 Fast Prep FP120 bead beater (BIO101 Savant) at setting 6 for 45s. Lysed suspensions were cooled on ice then spun at 14,000rpm in a bench top centrifuge for 10 minutes at 4ºC to pellet beads. 0.4ml lysis supernatant was mixed with 0.4 ml fresh Z buffer and incubated at 37ºC for 5 minutes. 0.16ml of 6mM p-nitrophenyl-β-D-glucuronide (Sigma) was then added to each tube and incubated at 37ºC for a defined length of time. The reaction was stopped by addition of 0.4ml of 1M Na_2CO_3. Tubes were centrifuged at 14,000rpm in a bench top centrifuge for 10 min to remove cell debris. The A_{405nm} of supernatant was measured and GusA activity calculated using the formula: (A_{405nm} x 1000)/(OD_{600} x reaction time (min) x 1.25 x vol of culture (ml)).

2.6.3 GusA reporter assays - *Clostridium perfringens*

The OD_{600} values of growing C. *perfringens* cultures were monitored and from mid-logarithmic growth phase cultures 1ml samples were taken, pelleted and frozen at -80ºC. The pellet was then thawed and resuspended in 0.8ml Z buffer (as above). To lyze the cells 8μl of toluene (Sigma) was added to the cell suspension. Tubes were capped and mixed vigorously for 1 min. Tubes

were then incubated on ice for 10 min then at 37ºC for 30 minutes. To start the enzymatic reaction 0.16ml of 6mM p-nitrophenyl-β-D-glucuronide (Sigma) was added to each tube. Tubes were incubated at 37ºC for a defined length of time. The reaction was stopped by addition of 0.4ml of 1M Na_2CO_3. Tubes were centrifuged at 14,000rpm in a bench top centrifuge for 10 min to remove cell debris. The A_{405nm} of supernatent was measured and GusA activity calculated using the formula: (A_{405nm} x 1000)/(OD_{600} x reaction time (min) x 1.25 x vol of culture (ml)).

2.7 Phenotypic assays

2.7.1 Colony morphology

For macroscopic visualization of *C. difficile* colony morphology strains were grown on BHI agar (Oxoid) for 5 days from an appropriate *C. difficile* inoculum such that isolated colonies grew on the plate. Plates were then scanned using a transmissive flat-bed scanner (Epson) to record images of the colonies or photographed. For microscopic visualization of *C. difficile* colony morphology glass-bottomed dishes (MatTek corporation) were inoculated directly on the glass with *C. difficile* taken with a sterile pipette tip from a 24 h-old *C. difficile* colony grown under standard conditions. Anaerobic BHI 0.7% agar at 40ºC was then carefully poured over the glass to cover the dish and allowed to set. *C. difficile* was grown between the glass and the agar for 3 days. Dishes were removed and images taken using a Zeiss Axiovert 200 widefield microscope at FILM, Imperial College London. Microscope images were analysed using Volocity 5.3.2 software (Improvision).

2.7.2 Flagella Expression Assay

C. difficile cultures were analysed for flagella expression based on a published protocol (Twine et al., 2009). BHI plates were inoculated with *C. difficile* from an overnight culture and grown for 24 hours to produce a lawn of bacterial growth. 500μl of distilled water was then added to the plate and bacteria harvested by scraping with an L-shaped spreader. Samples were normalized to an OD_{600} of 50. Samples in eppendorf tubes were subjected to 3 min of vortexing, following by centrifugation at 14,000rpm for 5 min. The supernatant was taken and concentrated 5-fold (200μl to 40μl). This concentrated supernatant was then analysed for flagella expression by SDS-PAGE and Coomassie staining, with 630 FliC running at ~33kDa.

2.7.3 Swimming Motility Assay

BHI 0.175% agar was poured into sterile FACS tubes and allowed to cool to room temperature. Tubes were then transferred to the anaerobic cabinet and left for 4 hours to equilibrate under anaerobic conditions. Tubes were inoculated to a defined depth with the straight end of sterile inoculation loops that had been dipped into *C. difficile* overnight liquid cultures. Tubes were incubated overnight then photographed to document motility exhibited.

2.7.4 *C. difficile* fibronectin ELISA

96-well MaxiSorp PS Immuno F96 plates (Nunc) were coated overnight at 4ºC with 50µl cellular fibronectin (Sigma) at 10ug/ml. After three washes with PBS + 0.1% Tween-20 (PBS-T) the plates were blocked with 100µl PBS-T + 3% dried skimmed milk (milk) per well at 37ºC for 2 hours. Plates were washed three times with PBS-T. *C. difficile* overnight cultures were harvested by

centrifugation then resuspended to an OD_{600} of 10. A two-fold dilution series of the bacterial suspension was carried out in PBS. 50µl samples of bacterial suspensions of the appropriate OD_{600} were then added to fibronectin-coated wells and incubated for 1 h at 37ºC. Following incubation wells were washed three times with PBS-T. Then anti-HMW SLP (1:15,000 dilution) and anti-LMW SLP (1:30,000 dilution) antibodies in milk were applied to each well and incubated for 1 h at 37ºC. Wells were washed three times with PBS-T then anti-rabbit HRP in milk (1:2000 dilution) added to each well, and incubated for 1 h at 37ºC. Wells were washed three times in PBS-T then 50µl OPD substrate (5ml 0.1M citric acid, 5ml 0.2M Na2HPO4, 10ml H_20, 1 OPD tablet (Sigma), 8µl H_20_2) was added to each well and incubated on the plate at room temperature for up to 1 h. The reaction was stopped using 20µl 3M H_2SO_4 per well and the optical density at 450nm was determined using a microplate reader.

2.7.5 *C. difficile* Caco-2 cell binding Assay

Caco-2 BBE1 cells (ATCC) were cultured under the conditions recommended by ATCC. Briefly, cells were propagated in high glucose (4.5g/L) DMEM (Gibco, Invitrogen) supplemented with 20% heat-inactivated fetal calf serum, 100µM MEM non-essential amino acids (Gibco, Invitrogen) and GlutaMAX (Invitrogen). In order to subculture cells trypsin-EDTA solution (Sigma) was used to remove cells from the flask. Cells were subcultured prior to confluency (maximum ~80% confluent) for routine propagation. For binding assays cells were grown in 12-well plates and 5 days after confluency differentiated cells were used for *C. difficile* binding assays. The 12-well plate containing cells was transferred to the anaerobic cabinet and 100µl of overnight *C. difficile*

culture was added directly into the cell culture medium. To allow comparison between strains overnight cultures were normalized to the same OD_{600} prior to addition to Caco-2 cells. Cells were incubated with the bacteria for 4 h in the anaerobic cabinet. The medium was carefully removed from cells and 1ml BHI added to each well, and gently shaken for 1 min to wash the cells. This gentle wash step was repeated twice more. Finally 100μl of BHI was added to the well, and the cells scraped and resuspended in this volume. This neat sample was diluted in BHI (10^{-1}, 10^{-2}, 10^{-3} dilutions), and plated out on BHI plates to determine the number of adherent bacteria per well.

2.7.6 *C. difficile* Caco-2 cell infection immunofluorescence

After infection of Caco-2 BBE1 cells grown on coverslips with *C. difficile* (same infection conditions as described for binding assay) cells were washed gently twice with PBS. Cells were then fixed with 3% formaldehyde in PBS for 20 minutes at room temperature. Cells were washed twice with PBS, then twice with 50mM NH_4Cl in PBS, then twice more with PBS. One wash was carried out with 1% BSA in PBS then all liquid removed from the cells. A 40μl drop of 1% BSA in PBS, containing anti-HMW SLP (1:100) was then added to the cells, and incubated for 45 min. Two PBS washes and one 1% BSA in PBS wash were carried out prior to addition of a 40μl drop of 1% BSA in PBS, containing Hoechst (1:500) and anti-rabbit rhodamine red-X (1:200) to the cells. Two PBS washes and one dH_2O wash of coverslips were carried out then excess liquid drained prior to mounting on a microscope slide with prolong gold antifade mount (Invitrogen). Cells were visualized using a Nikon Eclipse E600 microscope fitted with a Nikon DMX1200 camera.

2.7.7 Induction of *C. difficile* sporulation

C. difficile strains were streaked onto BHIS agar (For 1 litre: 37g BHI (Oxoid), 5g Yeast Extract (BD Bacto), 1g L-Cysteine (Sigma), 15g Agar (BD Bacto) and grown overnight anaerobically at 37ºC. A single colony was used to inoculate 15 ml of TGY broth (For 1 litre: 30g Tryptic Soy Broth (BD Bacto), 2% Glucose (VWR), 10g Yeast Extract (BD Bacto) 1g L-Cysteine (Sigma). Following overnight growth to stationary phase in TGY broth under standard culture conditions, bacteria were subcultured at a dilution of 1 in 10 in SMC medium (For 1 litre: 90g Bacto Peptone (BD Bacto), 5g Proteose Peptone (Sigma), 1g Ammonium Sulphate (Sigma), 1.5g TRIS (VWR). When the OD_{600} of the SMC culture reached 0.4, 100 µL of the culture was spread on SMC agar (SMC broth with 1.5% agar (BD) and incubated anaerobically at 37°C for 5 days. Addition of 2ml of dH_2O followed by mixing with an L-shaped spreader allowed harvesting of the **crude spore preparation**. Samples were taken for analysis by phase microscopy.

2.7.8 Quantification of sporulation efficiency

Spores were purified using HistoDenz (Sigma) based on a published protocol (Sorg & Dineen, 2009). Crude spore preparations harvested from SMC plates (see above) were washed five times in 1ml ice-cold dH_20 by centrifugation and resuspension/vortex steps. The OD_{600} of this washed spore preparation was measured to reflect the amount of material present prior to purification, a sample for microscopy was also taken. The washed spore preparation was then pelleted by centrifugation and resuspended in 500 µL 20% (w/v) HistoDenz. This was carefully layered onto 1 ml 50% (w/v) HistoDenz in an

eppendorf tube to form a density gradient. This gradient was then centrifuged at 14,000 x g for 15 minutes. Cell debris and any remaining vegetative cells collected at the boundary between 20% and 50% HistoDenz, and was discarded. The pellet containing spores was washed five times in ice-cold dH_2O to remove HistoDenz and yield the purified spore preparation of which OD_{600} measurements and samples for microscopy were taken. Purified spores were stored in water at 4°C, or frozen at -80ºC for long-term storage. Sporulation yield was calculated as OD_{600} purified spores (post-HistoDenz)/OD_{600} washed spore preparation (pre-HistoDenz).

2.7.9 Germination Efficiency Assay

Purified spores were normalized to an OD_{600} of 1.0 in dH_2O. A ten-fold dilution series of this purified spore preparation was carried out in PBS 0.05% Tween-80 (Sigma). Samples (typically of 10^{-4}, 10^{-5}, 10^{-6} and 10^{-7} dilutions) were spread onto BHI plates containing 0.1% sodium taurocholate (Sigma) in triplicate. Plates were then incubated overnight anaerobically at 37ºC. Colonies were counted to determine the number of spores germinating from the samples. Germination efficiency was calculated as the number of colony forming units (cfu) per ml of an OD_{600}=1.0 purified spore preparation.

Table 2.2 Primers used in this study

Name	Sequence (5' to 3')	Characteristic or use
NF812	GGTAAATAAGGAAAAAATAATAAGAATTCGAATTTTTTAGGGGGAAAATACCATG	Site-directed mutagenesis to introduce *Eco*RI site 10 bp upstream of RBS in pCBR001, creating pCBR003
NF813	CATGGTATTTTCCCCCTAAAAAATTCGAATTCTTATTATTTTTTCCTTATTTACC	Site-directed mutagenesis to introduce *Eco*RI site 10 bp upstream of RBS in pCBR001, creating pCBR003
NF883	CTTTAACTATGCCGGGATACATCGCAGCGTAATGC	Site-directed mutagenesis to remove *Bam*HI site from *gusA*
NF884	GCATTACGCTGCGATGTATCCCGGCATAGTTAAAG	Site-directed mutagenesis to remove *Bam*HI site from *gusA*
NF885	CCGGAATTCCTGCAGTAAAGGAGAAAATTTTATGTTACG	Amplification of *gusA*
NF886	CGCGGATCCGGTGCGCCAGGAGAGTTGTTGATTC	Amplification of *gusA*
NF814	GGTACCAATATCTAATTTTAAAGAGGTTTCTAACTATAATG	Amplification of short ON from CD630 gDNA
NF815	GAATTCACTTAATTTTTTAACTTTGTTCCCGTTCTAATTCT	Amplification of short ON/OFF from CD630 gDNA
NF816	GGTACCTTTTTATACAAAACCCTATTTTTTATAACAATTC	Amplification of short OFF from CD630 gDNA
NF855	CCGCCGGGTACCGTAATTTGAATTAAAATGTAAATATAAAGAATAATTAAAAAAATATCGC	Amplification of ON/OFF from CD630 gDNA
NF856	CCGCCGGAATTCACTTAATTTTTTAACTTTGTTCCCGTTCTAATTCT	Amplification of ON/OFF from CD630 gDNA
NF864	AAAGTGCTACTCTTTTTACAAATAAATTTTATTATTTATACTTG	Inverse PCR to delete 2 bp from IR
NF865	ATTCTAAAGGATACTTAATATCTAATTTTAAAGAGGTTTC	Inverse PCR to delete 2 bp from IR
NF866	ATTCTAAAGGCTACTTTTTTTATACAAAACCCTATTTTTTATAAC	Inverse PCR to delete 2 bp from IR

Name	Sequence (5' to 3')	Characteristic or use
NF867	GAACGGGAACAAAGTTAAAAAATTAAGTGAATTCGAATTTTTTAG	Inverse PCR to delete 2 bp from IR
NF894	GTGCTACTCTTTTTACAAATAAATTTTATTATTTATACTTG	Inverse PCR deletion of whole region of inversion, or parts of region of inversion
NF895	CGGGAACAAAGTTAAAAAATTAAGTGAATTCGAATTTTTTAG	Inverse PCR deletion of whole region of inversion
NF952	CT TAA AAT ATA AAT TTA CAA ATT TTT CAA TCA ATA AGA AAA TTT C	Inverse PCR to create Δ1 deletion in inversion region
NF953	GTA AGT TTG ATT TTT ATG TTA ATG AAT TGT TAT AAA AAA TAG G	Inverse PCR to create Δ2 and Δ4 deletions in inversion region
NF954	CAA TGT AAC AAA ATA ATT ACA AAT TAA TTA TGA AAT CAT TAT AG	Inverse PCR to create Δ2 and Δ5 deletions in inversion region
NF956	CGG GAA CAA AGT TAA AAA ATT AAG TGA ATT CCT G	Inverse PCR to create Δ4 deletion in inversion region
NF909	ATT CTA AAG GAT ACT TAA TAT CTA CAA AAC CCT ATT TTT TAT AAC AAT TC	For mutation of 5bp in the OFF orientation of the *cwpV* switch (OFF>ON)
NF910	GAA TTG TTA TAA AAA ATA GGG TTT TGT AGA TAT TAA GTA TCC TTT AGA AT	For mutation of 5bp in the OFF orientation of the *cwpV* switch (OFF>ON)
NF911	TT CTA AAG GCT ACT TTT TTT ATA ATT TTA AAG AGG TTT CTA ACT ATA ATG	For mutation of 5bp in the OFF orientation of the *cwpV* switch (ON>OFF)
NF912	CAT TAT AGT TAG AAA CCT CTT TAA AAT TAT AAA AAA AGT AGC CTT TAG AA	For mutation of 5bp in the OFF orientation of the *cwpV* switch (ON>OFF)
NF913	CGCGGATCCAAATAATGATAAGAAAATTATAAAAGTACATAA	Amplification of recombinase CD1932 for cloning into pACYCDuet-1
NF914	ATAAGAATGCGGCCGCTTAATCGTCCTCAATCCATTCAA	Amplification of recombinase CD1932 for cloning into pACYCDuet-1
NF915	CGCGGATCCAAAAGCTGCAATTTATTCAAGA	Amplification of recombinase CD1905 for cloning into pACYCDuet-1
NF916	ATAAGAATGCGGCCGCCTAATTGCTTAAAGAAACAGAATTATC	Amplification of recombinase CD1905 for cloning into pACYCDuet-1

Name	Sequence (5’ to 3’)	Characteristic or use
NF917	CGCGGATCCAAATGCATTTATTAGAAAAAGAAATAA	Amplification of recombinase CD2066 for cloning into pACYCDuet-1
NF918	ATAAGAATGCGGCCGCTTAATTAGCATTAAATAAACTTTCAGA	Amplification of recombinase CD2066 for cloning into pACYCDuet-1
NF919	CGCGGATCCAAATGTGGCAATTTATTTACGCAAAA	Amplification of recombinase CD1882 for cloning into pACYCDuet-1
NF920	ATAAGAATGCGGCCGCTTAATCTCTTGGGATATTAGGGTAA	Amplification of recombinase CD1882 for cloning into pACYCDuet-1
NF921	CGCGGATCCAAAAATGGCAACAAGACCTATAGAAA	Amplification of recombinase CD1167 for cloning into pACYCDuet-1
NF922	ATAAGAATGCGGCCGCTTAACCAATAAAGAAATTTTCACTAGC	Amplification of recombinase CD1167 for cloning into pACYCDuet-1
NF1043	CGCGGATCCAATGATTATTATATATGGATATTGCCGTG	Amplification of recombinase CD3578 for cloning into pACYCDuet-1
NF1044	ATAAGAATGCGGCCGCCTAAATTGATTTTTTAGCACGAATAAGTGTGC	Amplification of recombinase CD3578 for cloning into pACYCDuet-1
NF1045	CGCGGATCCAATGGATATTATAGAGGGATATATAG	Amplification of recombinase CD1222 for cloning into pACYCDuet-1
NF1046	ATAAGAATGCGGCCGCCTATTTAGCTTCTGCTTTTTGTAATTCTAC	Amplification of recombinase CD1222 for cloning into pACYCDuet-1
NF823	GCATTTTTCCATCCATCTTG	For cwpV switch orientation PCR, anneals outside region of inversion
NF825	TTTAAGGTAAGTTTGATTTTTATGTTAATGAAT	For cwpV switch orientation PCR, anneals in region of inversion
NF826	ATTCATTAACATAAAAATCAAACTTACCTTAAA	For cwpV switch orientation PCR, anneals in region of inversion
NF793	CACCTCCTTTTTGACTTTAAGCCTACGAATACC	For cwpV switch orientation PCR, anneals in pMTL960 backbone
NF1215	CTTAATCACATATATCACACTTGTATAAGTTTACTGG	Analysis of *recV* ClosTron insertion site
NF1356	CATTAATAGCTTGCTGAGTTACCCTTATATATCGC	Analysis of *recV* ClosTron insertion site

Name	Sequence (5' to 3')	Characteristic or use
NF1357	AGA GCT CAA TCA AAC TAA GAG GAG TGG TTG AAA ATG GC	For amplification of *recV* for cloning into pRPF146 with a C-terminal strep tag
NF1359	GTG GAT GAC TCC ACT CGA GAC CAA TAA AGA AAT TTT CAC TAG CTT C	For amplification of *recV* for cloning into pRPF146 with a C-terminal strep tag
NF1360	TGG ATC CTT ATT TTT CAA ATT GTG GAT GAC TCC ACT CGA GAC C	For amplification of *recV* for cloning into pRPF146 with a C-terminal strep tag
NF1411	GAT TTA TAA GGG TAA CTC AGC AAG C	To create Y176F mutation in *recV*
NF1412	GCT GTG TAG TGG CTA CAG AGC	To create Y176F mutation in *recV*
NF1019	GAT ATA CCA TGG GCT CTT CAT CAA AGG TAG GTC AAA TAA AAT TTA G	Amplify two and nine 630 type I repeats for pET28a cloning
NF1020	GTG GTG CTC GAG ACC TAC TGC TGG CGC TGC CAC ATA TTT TTT AAC	Amplify two 630 type I repeats for pET28a cloning
NF1021	GATATACCATGGGCTCATCAAAAGTAGATAAAGTAGTTTTG	Amplify two R20352 type II repeats for pET28a cloning
NF1022	GTGGTGCTCGAGTACTACTTTTCCTGCTTCAAATGTCTC	Amplify two R20352 type II repeats for pET28a cloning
NF1023	GATATACCATGGGCGTGGCAAAAATAACAAGCGTTCCTGC	Amplify two CDKK167 type III repeats for pET28a cloning
NF1024	GTGGTGCTCGAGTATAGCTGTTCTGGCTGG	Amplify two CDKK167 type III repeats for pET28a cloning
NF1025	GATATACCATGGGCACACCAACAGTAGACCCAGTTGAAGG	Amplify one M9 type IV repeat for pET28a cloning
NF1026	GTGGTGCTCGAGTGTTGGTGTTTTATCTAATACTATTG	Amplify one M9 type IV repeat for pET28a cloning
NF1027	GATATACCATGGGCCCAGCCAAATCAACAAAAGTAATAATAGG	Amplify one AY1 type IV repeat for pET28a cloning
NF1028	GTGGTGCTCGAGTCCTGCTGTTGCTCCTGATAC	Amplify one AY1 type IV repeat for pET28a cloning
NF1095	CTA ATT TTT ATT CAT ATT TCA TCT CGA GTT TTA ATA TAT TAA ATA AAC	Amplify nine 630 type I repeats for pET28a cloning

Name	Sequence (5' to 3')	Characteristic or use
NF1096	CCGGAATTCAAATGGACAAGGGGGAAGAAAATAAAATG	Amplification of *cwpV* gene from CD630 gDNA for cloning into pCBR023 to produce pCBR044, and for amplification of all other types of *cwpV* gene
NF1097	CGCGGATCCCCAATACTAGTATTATCAAGAAAATTTACATC	Amplification of cwpV gene from CD630 gDNA for cloning into pCBR023 to produce pCBR044
NF1123	CCA CGG CGC GCT GAT GAA GAA CTT CCT CCA CTA C	For deletion of the *cwpV* repeat region to produce truncated *cwpV*
NF1124	GAG TGG GCT CGA GTA AAT AAT GAA ATA TGA ATA AAA ATT AG	For deletion of the *cwpV* repeat region to produce truncated *cwpV*
NF1296	CACAATTTGAAAAATAAATAATGAAATATGAATAAAAATTAG	To add strep tag to *cwpV* in pCBR044
NF1297	GATGACTCCACTCGAGTTTTAATATATTAAATAAACTTGATAGC	To add strep tag to *cwpV* in pCBR044
NF1298	CGCCTCGAGTTTCATAAATCCTAGTATTCTTGATAACATTAATGC	Amplification of R20352 *cwpV* gene for cloning into *C. difficile* expression vector, pairs with NF1096
NF1299	CGCCTCGAGTTTATTAAAATTAAAAATATTTACTGCTATTCTACCC	Amplification of CDKK167 *cwpV* gene for cloning into *C. difficile* expression vector, pairs with NF1096
NF1300	CGCCTCGAGATTAATTACAGATTTTGCTACATTTGTAAGATTTG	Amplification of M9 *cwpV* gene for cloning into *C. difficile* expression vector, pairs with NF1096
NF1301	CGCCTCGAGTTTAACCAGATTAGTAAACTTAAGCATTAGTC	Amplification of AY1 *cwpV* gene for cloning into *C. difficile* expression vector, pairs with NF1096
NF1308	GCC AAG ATA GAA AAT AAT TCT AAA GG	CwpV cleavage domain mutagenesis, primer F1
NF1309	ACT AAC TTA AAG GTT AAT GAT AAA AAT GG	CwpV cleavage domain mutagenesis, primer F2
NF1310	GAT GGA ACT TTT GTT AAC TAT GG	CwpV cleavage domain mutagenesis, primer F3
NF1311	GTG GAT GAT GTA GTT ATA GAT GAT GTA AAA GAT GG	CwpV cleavage domain mutagenesis, primer F4
NF1312	CCT AAT GGT GAT GTA AAT AAC TAT GC	CwpV cleavage domain mutagenesis, primer F5

Name	Sequence (5’ to 3’)	Characteristic or use
NF1313	TAT TTT TCC ATA GTT AAC AAA AGT TCC	CwpV cleavage domain mutagenesis, primer R1
NF1314	TTT TAC ATC ATC TAT AAC TAC ATC ATC C	CwpV cleavage domain mutagenesis, primer R2
NF1315	TTT TGC ATA GTT ATT TAC ATC ACC ATT AGG	CwpV cleavage domain mutagenesis, primer R3
NF1316	CAT GTT GAT AGT CAC AGT CTT TGT ATA TG	CwpV cleavage domain mutagenesis, primer R4
NF1317	ACC ATT TAA GTT TAC TGT TAC TGC	CwpV cleavage domain mutagenesis, primer R5

Table 2.3 Plasmids used in this study

Name	Relevant characteristics	Source/ reference
pTUM177	Contains promoterless *gusA* gene with *C. difficile tcdB* RBS.	Mani and Dupuy (2001)
pMTL960	*E. coli – C. difficile* shuttle vector.	Purdy *et al.,* (2002)
pCBR001	Contains synthetic gene for FbFP; pUC19 backbone.	This study and Drepper *et al.,* (2007b)
pCBR003	pCBR001 modified by introduction of *Eco*RI site upstream of RBS.	This study.
pCBR012	Derived from pCBR001; contains *fbfp* gene under control of *cwp2* promoter; pMTL960 backbone.	This study.
pCBR023	Contains modified *gusA* gene under control of *cwp2* promoter; pUC19 backbone.	This study.
pCBR024	pUC19 containing *gusA* and short upstream sequence. *cwpV* switch in ON orientation. (short ON)	This study.
pCBR025	pUC19 containing *gusA* and short upstream sequence. *cwpV* switch in OFF orientation. (short OFF)	This study.
pCBR026	pUC19 containing *gusA* and long upstream sequence of *cwpV* with switch in ON orientation. (ON)	This study.
pCBR027	pUC19 containing *gusA* and long upstream sequence of *cwpV* with switch in OFF orientation. (OFF)	This study.
pCBR028	Derivative of pCBR026 containing 2 bp deletion in LIR of *cwpV* switch. (ON LIR*)	This study.
pCBR029	Derivative of pCBR027 containing 2 bp deletion in LIR of *cwpV* switch. (OFF LIR*)	This study.
pCBR030	Derivative of pCBR026 containing 2 bp deletion in RIR of *cwpV* switch. (ON RIR*)	This study.

pCBR031	Derivative of pCBR027 containing 2 bp deletion in RIR of *cwpV* switch. (OFF RIR*)	This study.
pCBR032	Derivative of pCBR027 with entire switch region deleted (ΔIR)	This study.
pCBR033	pUC19 containing the full-length *cwpV* gene from 630 under the control of the *cwp2* promoter	This study.
pCBR034	Insert from pCBR023 cloned into pMTL960. (P_{cwp2}-*gusA*)	This study.
pCBR035	Insert from pCBR024 cloned into pMTL960. (short ON-*gusA*)	This study.
pCBR036	Insert from pCBR025 cloned into pMTL960. (short OFF- *gusA*)	This study.
pCBR037	Insert from pCBR026 cloned into pMTL960. (ON- *gusA*)	This study.
pCBR038	Insert from pCBR027 cloned into pMTL960. (OFF- *gusA*)	This study.
pCBR039	Insert from pCBR028 cloned into pMTL960. (ON LIR*- *gusA*)	This study.
pCBR040	Insert from pCBR029 cloned into pMTL960. (OFF LIR*- *gusA*)	This study.
pCBR041	Insert from pCBR030 cloned into pMTL960. (ON RIR*- *gusA*)	This study.
pCBR042	Insert from pCBR031 cloned into pMTL960. (OFF RIR*- *gusA*)	This study.
pCBR043	Insert from pCBR032 cloned into pMTL960. (ΔIR- *gusA*)	This study.
pCBR044	pMTL960 containing the full-length *cwpV* gene from CD630 under the control of *cwp2* promoter (pOECwpV)	This study.
pCBR045	pACYC-Duet1 containing CD1932	This study.
pCBR046	pACYC-Duet1 containing CD1905	This study.
pCBR047	pACYC-Duet1 containing CD2066	This study.

pCBR048	pACYC-Duet1 containing CD1882	This study.
pCBR049	pACYC-Duet1 containing CD1167 (*recV*)	This study.
pCBR050	pACYC-Duet1 containing CD3578	This study.
pCBR051	pACYC-Duet1 containing CD1222	This study.
pCBR059	Derivative of pCBR038 with section of *cwpV* switch deleted - Δ1	This study.
pCBR060	Derivative of pCBR038 with section of *cwpV* switch deleted - Δ2	This study.
pCBR061	Derivative of pCBR038 with section of *cwpV* switch deleted - Δ4	This study.
pCBR062	Derivative of pCBR038 with section of *cwpV* switch deleted - Δ5	This study.
pCBR063	Derivative of pCBR037 with 5bp in predicted stem loop mutated (ON>OFF)	This study.
pCBR064	Derivative of pCBR038 with 5bp in predicted stem loop mutated (OFF>ON)	This study.
pCBR065	Derivative of pCBR033 with truncated cwpV gene	This study.
pCBR066	Insert from pCBR065 cloned into pMTL960 (pOE Nter)	This study.
pCBR068	pET28a for expression of two CwpV Type I repeats (630)	This study.
pCBR069	pET28a for expression of two CwpV Type II repeats (20352)	This study.
pCBR070	pET28a for expression of two CwpV Type III repeats (CDKK167)	This study.

pCBR071	pET28a for expression of one CwpV Type IV repeat (M9)	This study.
pCBR072	pET28a for expression of one CwpV Type V repeat (AY1)	This study.
pCBR073	pET28a for expression of nine type I repeats (630)	This study.
pCBR079	Derivative of pCBR034 encoding CwpV with a C-terminal strep-tag	This study.
pCBR080	Insert from pCBR079 cloned into pMTL960 (pOE I)	This study.
pCBR081	Derivative of pCBR080 with CwpV cleavage domain deletion – CDM1	This study.
pCBR082	Derivative of pCBR080 with CwpV cleavage domain deletion – CDM2	This study.
pCBR083	Derivative of pCBR080 with CwpV cleavage domain deletion – CDM3	This study.
pCBR084	Derivative of pCBR080 with CwpV cleavage domain deletion – CDM4	This study.
pCBR085	Derivative of pCBR080 with CwpV cleavage domain deletion – CDM5	This study.
pCBR086	Derivative of pCBR080 with CwpV cleavage domain deletion – CDM6	This study.
pCBR100	Derivative of pCBR079 encoding 20352 CwpV	This study.
pCBR101	Derivative of pCBR079 encoding CDKK167 CwpV	This study.
pCBR102	Derivative of pCBR079 encoding M9 CwpV	This study.
pCBR104	Derivative of pCBR079 encoding AY1 CwpV	This study.

pCBR105	Insert from pCBR079 cloned into pMTL960 (pOE II)	This study.
pCBR106	Insert from pCBR079 cloned into pMTL960 (pOE III)	This study.
pCBR107	Insert from pCBR079 cloned into pMTL960 (pOE IV)	This study.
pCBR109	Insert from pCBR079 cloned into pMTL960 (pOE V)	This study.
pCBR113	Derivative of pRPF146 encoding RecV with a C-terminal strep-tag under the control of a tetracylcine-inducible promoter	This study.
pCBR115	Derivative of pCBR113 with a mutation to encode RecV Y176F	This study.
pRPF146	pMTL960 with *gusA* under the control of a tetracylcine-inducible promoter	Dr. Robert Fagan

3 Production of recombinant plasmids

3.1 Reporter Plasmids

Plasmid pCBR001 was constructed (Codon Devices, USA) carrying a synthetic gene for the blue light photoreceptor FbFP (Drepper et al., 2007b) under the control of the constitutive cell wall protein Cwp2 promoter (P_{cwp2}) and its ribosome binding site (RBS) with a pUC19 backbone. pCBR001 was digested with *Bam*HI and *Acc*65I and the *Bam*HI-*Acc*65I fragment containing the P_{cwp2} sequence and the *FbFP* gene were subcloned into the *E.coli-C. difficile* shuttle plasmid pMTL960 (Purdy et al., 2002) to create pCBR012. In order to allow substitution of the gene and/or promoter the pCBR001 plasmid was modified by the introduction of an *Eco*RI site 10 bp upstream of the RBS using primers NF812 and NF813 to create pCBR003.

The *gusA* reporter plasmid pTUM177 (Mani and Dupuy, 2001), containing a promoterless *gusA* gene and the *toxB* ribosome-binding site, was mutated by deletion of the *Bam*HI site within *gusA* by SDM using primers NF883 and NF884. The *toxB* ribosome-binding site and the *gusA* ORF were amplified with primers NF885 and NF886, to produce an 1860 bp *Eco*RI-*Bam*HI fragment. This fragment was ligated to pCBR003 to produce pCBR023. An illustration of pCBR023 is shown in Figure 3.1.

The plasmid pCBR023 was then cleaved with *Acc*65I-*Eco*RI and the P_{cwp2} fragment was replaced with one of the following fragments derived by PCR of *C. difficile* CD630 DNA from upstream of *cwpV:* a short fragment not including the putative *cwpV* promoter (P_{cwpV}) and having the *cwpV* switch in the ON orientation

(primers NF814 and NF815) was used to form pCBR024; a short fragment not including the P_{cwpV} and having the *cwpV* switch on the OFF orientation (primers NF814 and NF815) forming pCBR025; a long fragment including the putative *cwpV* promoter was used with primers NF855 and NF856 and the transformants screened by PCR with the primers followed by *Ase*I digestion to identify a plasmid containing the *cwpV* switch in the ON orientation (pCBR026) and one in the OFF orientation (pCBR027). To produce pUC19-ON LIR*-GusA and pUC19-ON RIR*-GusA plasmids (pCBR028 and pCBR030) 2 bp deletions in the IRs of pCBR026 were carried out by inverse PCR. For ON LIR* primers NF864 and NF865 were used, for ON RIR* primers NF866 and NF867 were used. To produce pUC19-OFF LIR*-GusA and pUC19-OFF RIR*-GusA plasmids (pCBR029 and pCBR031) 2 bp deletions in the IRs of pCBR027 were carried out by inverse PCR. For OFF LIR* deletion primers NF864 and NF866 were used, for OFF RIR* production primers NF865 and NF867 were used. Inverse PCR was carried out with primers NF894 and NF895 to delete the entire *cwpV* switch from pCBR027 to form pCBR032.

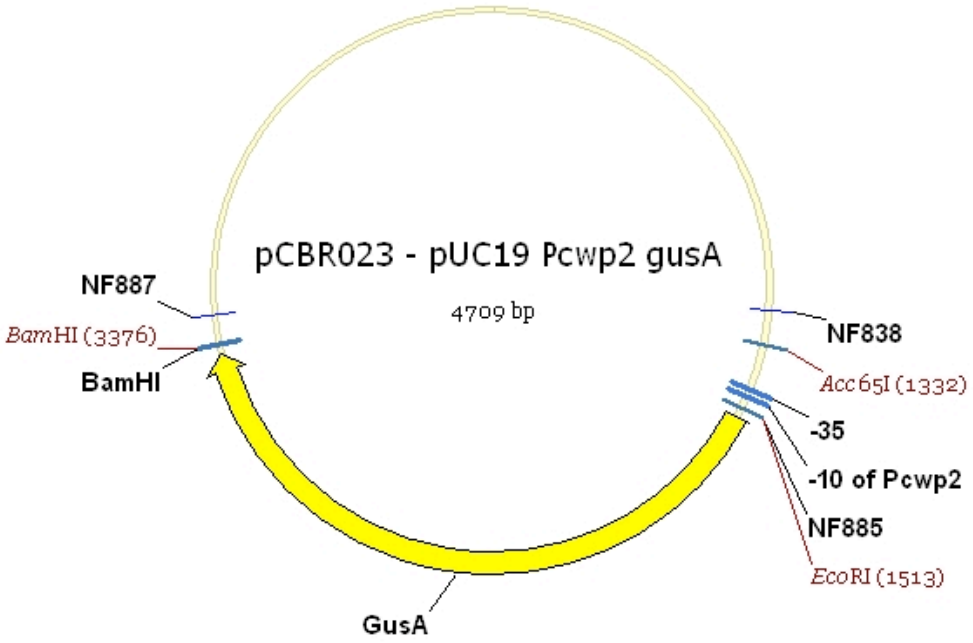


Figure 3.1 pCBR023

The plasmid pCBR023 has a pUC19 backbone. The *Acc65*I-*Bam*HI fragment contains the Cwp2 promoter and the *gusA* reporter gene. The *Eco*RI site is situated 10bp upstream of the ribosome binding site to enable promoters and genes to be switched by restriction digest using *Eco*RI and either *Bam*HI (to switch the gene) or Acc65I (to switch the promoter). The *Acc*65I-*Bam*HI fragment can be cloned into pMTL960 for subsequent transfer to *C. difficile*. pUC19-based plasmids were often used to plasmid mutagenesis, due to the smaller size of pUC19 compared to pMTL960 and also to eliminate the risk of introducing secondary plasmid backbone mutations in pMTL960.

For plasmids pCBR023-032 the *Bam*HI-*Acc*65I fragment containing the putative promoter sequences and the *gusA* gene were subcloned into pMTL960 to produce plasmids pCBR034-043. An illustration of pCBR034 is shown in Figure 3.2.

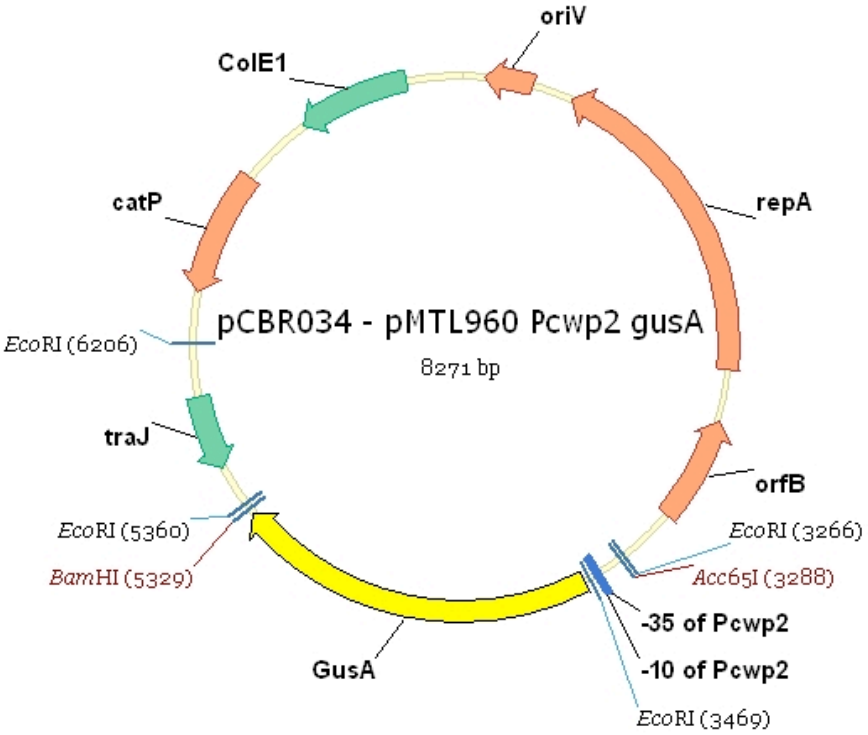


Figure 3.2 pCBR034

The plasmid pCBR034 has the *Acc*65I-*Bam*HI fragment from pCBR024 cloned into pMTL960. The pMTL960 backbone encodes the functions necessary for conjugative transfer of the plasmid from *E. coli* CA434 to *C. difficile*. Due to the presence of *Eco*RI sites in the pMTL960 backbone promoters and genes cannot be simply exchanged by restriction digest, this must be done in the pUC19 backbone plasmids and constructs subsequently subcloned into pMTL960.

To test our model by deletion of sections of the OFF orientation of the *cwpV* switch inverse PCR was carried out with pCBR038 as template. For Δ1 primers NF894 and NF952 were used to create pCBR059. For Δ2 primers NF953 and NF954 were used to create pCBR060. For Δ4 primers NF953 and NF956 were used to create pCBR061. For Δ5 primers NF894 and NF954 were used to create pCBR062. For mutation of 5bp in the stem loop region referred to as OFF>ON, pCBR038 was used as template with primers NF909 and NF910 to create pCBR064. For ON>OFF pCBR037 was used as template with primers NF911 and NF912 to create pCBR063.

3.2 Recombinase expression in *E. coli*

Genes for seven recombinase genes (CD1932, CD1905, CD2066, CD1822, CD1167, CD3578 and CD1222) were amplified from CD630 gDNA using specific primers, as detailed in Table 2.2. These PCR products were digested with BamHI and NotI and cloned into pACYCDuet-1 (Novagen) digested with the same enzymes to generate plasmids pCBR045-51, encoding N-terminally his-tagged recombinases. An illustration of pCBR045 is shown in Figure 3.3.

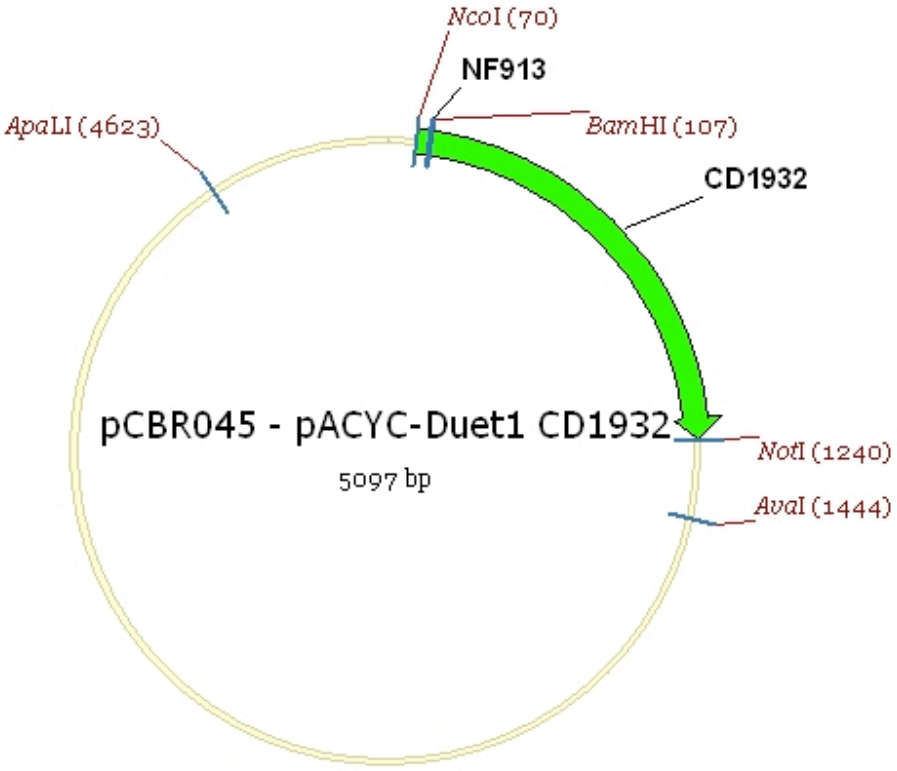


Figure 3.3 pCBR045

The plasmid pCBR045 contains the CD1932 recombinase gene in the *Bam*HI-*Not*I cloning site. This generates a construct encoding the recombinase with an N-terminal His-tag.

3.3 RecV (CD1167) expression in *C. difficile*

For complementation of *ΔrecV* the *recV* gene from *C. difficile* 630 was amplified with two rounds of PCR using primers NF1357 and NF1359, then NF1357 and NF1360. These primers were designed to add an XhoI site and a C-terminal StrepII-tag-encoding sequence to the 5' end of the PCR product. This amplified fragment was directionally cloned using *Sac*I and *Bam*HI into the *C. difficile* tetracycline-inducible expression vector pRPF146 to generate pCBR113 (pRecV), encoding RecV with a C-terminal StrepII-tag. pCBR113 is illustrated in Figure 3.4 Inverse PCR with primers NF1411 and NF1412, using pCBR113 as the template, was carried out to mutate the *recV* tyrosine176 codon to a phenylalanine codon. This produced the plasmid pCBR115 ($pRecV_{Y176F}$).

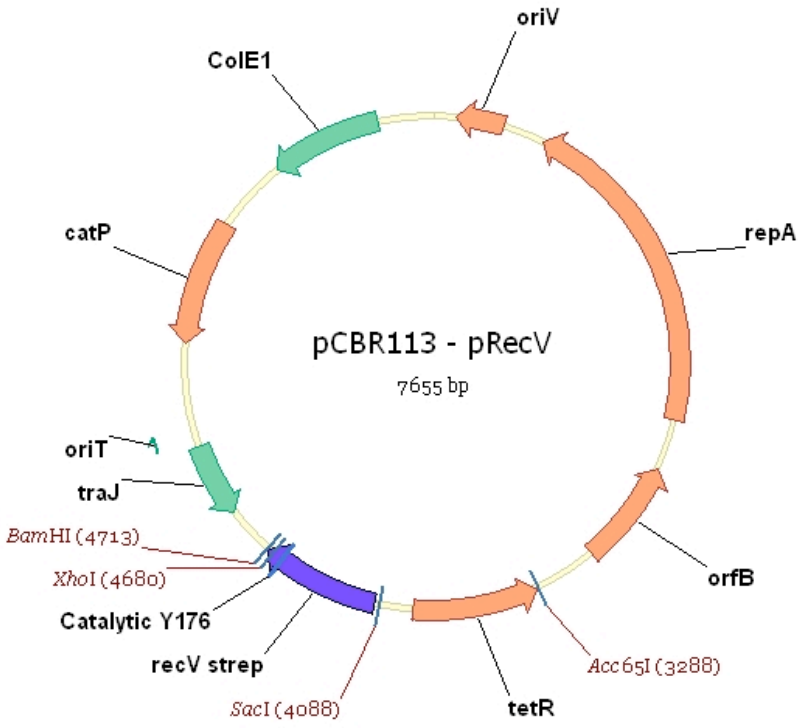


Figure 3.4 pCBR113

The plasmid pCBR113 encodes RecV with a C-terminal StrepII-tag under the control of a tetracycline inducible promoter. From this construct genes can be exchanged using *Sac*I and *Xho*I, without affecting the C-terminal StrepII-tag.

3.4 Expression of CwpV repeat fragments in *E.coli*

To clone fragments of CwpV for expression of recombinant proteins CwpV repeats from strains 630 (type I), R20352 (type II), CDKK167 (type III), M9 (type IV) and AY1 (type V) were amplified by PCR and primers described in Table 2.2. These fragments were directionally cloned using *Nco*I and *Xho*I into pET28a *E. coli* expression vector to generate plasmids pCBR068-072. pCBR068 is illustrated in Figure 3.5. For expression of nine type I CwpV repeats primers NF1019 and NF1095 were used to amplify nine repeats from 630 gDNA. This fragment was then cloned into pET28a to produce pCBR073.

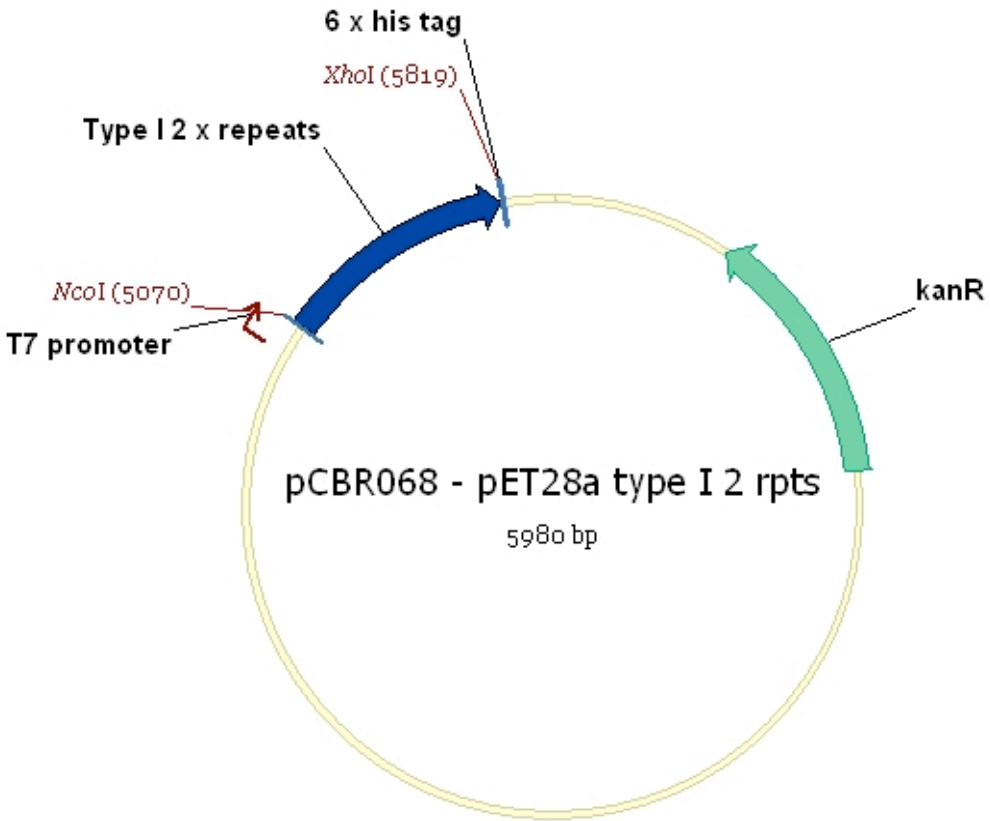


Figure 3.5 pCBR068

The plasmid pCBR068 was constructed by insertion of sequence encoding for two typeI CwpV repeats into the *Nco*I-*Xho*I cloning site, yielding a plasmid encoding for recombinant protein with a C-terminal His-tag.

3.5 Expression of CwpV in *C. difficile*

The full-length *cwpV* gene was amplified from CD630 gDNA using primers NF1096 and NF1097. This PCR product was digested with EcoRI and BamHI and cloned into pCBR023 in place of the *gusA* gene to create pCBR033. The Acc65I-BamHI fragment containing PCwp2 and the *cwpV* gene was then subcloned into pMTL960 to produce pCBR044, illustrated in Figure 3.6.

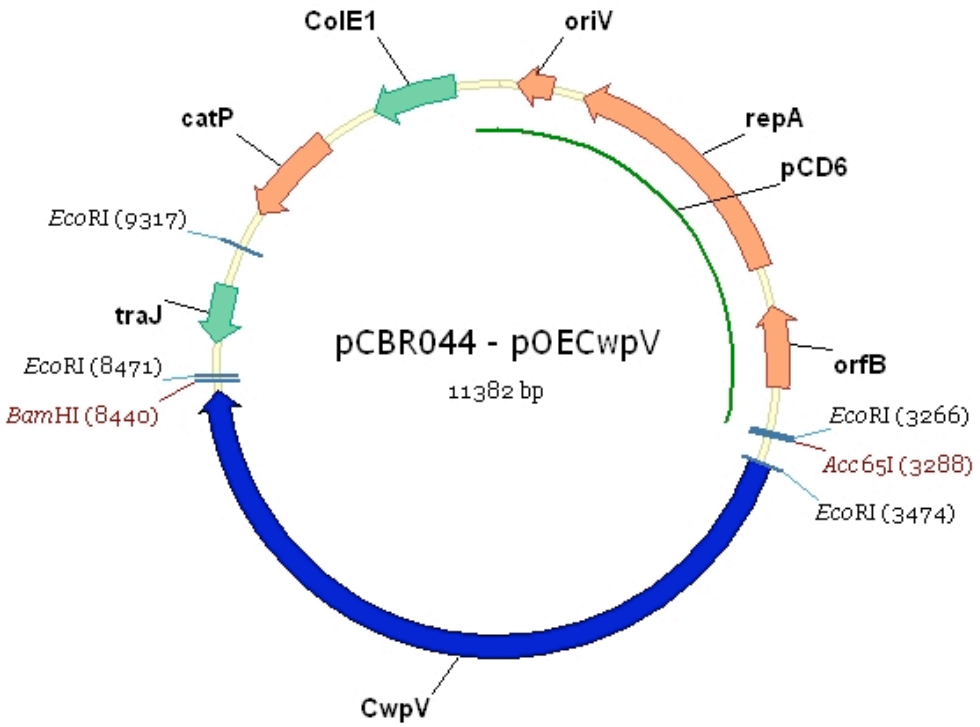


Figure 3.6 pCBR044

The plasmid pCBR044 (pOECwpV) allows expression of CwpV in *C. difficile*. It utilises the Cwp2 promoter that had been assessed using GusA reporter assays. It was constructed from the GusA construct, by exchange of *gusA* for *cwpV*.

To truncate the *cwpV* gene, such that it only encodes the N-terminal part of CwpV up to the end of the serine/gycine rich linker, inverse PCR was carried out using primers NF1123 and NF1124 on pCBR033 to generate pCBR065. The *Acc*65I-

*Bam*HI fragment containing PCwp2 and the truncated *cwpV* gene was then subcloned into pMTL960 to produce pCBR066.

In order to add an 5' XhoI site and a C-terminal StrepII-tag (WSHPQFEK)-encoding sequence to *cwpV* the plasmid pCBR033 was used as template for inverse PCR with primers NF1296 and NF1297 to generate pCBR079. The Acc65I-BamHI fragment containing PCwp2 and the StrepII-tagged *cwpV* gene was then subcloned into pMTL960 to produce pCBR080, illustrated in Figure 3.7.

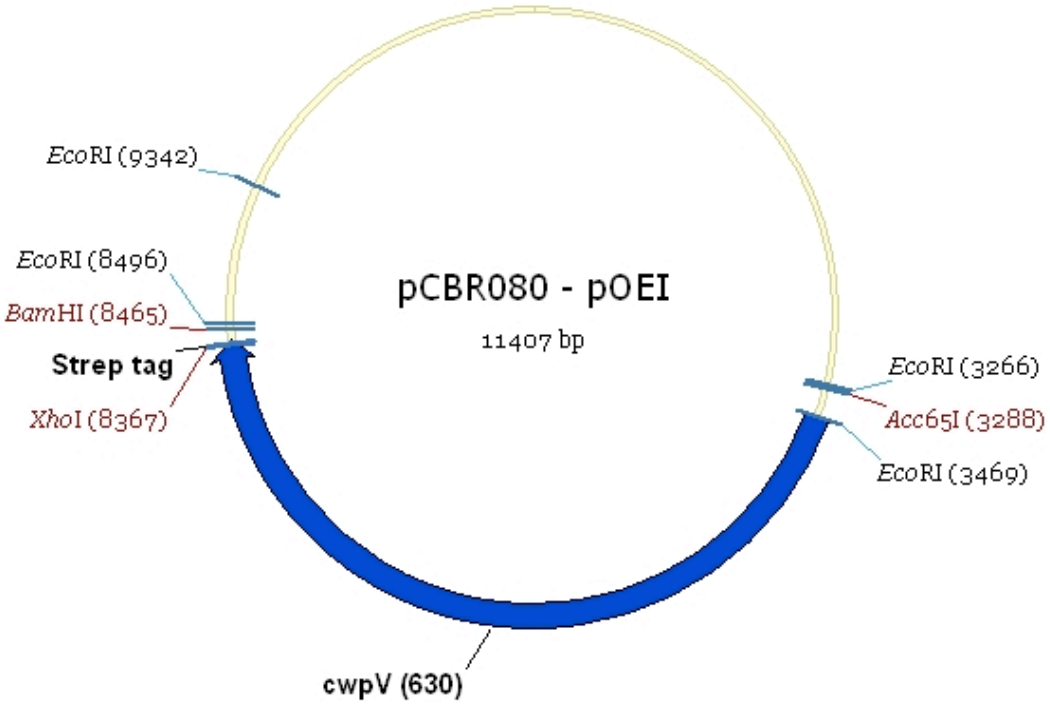


Figure 3.7 pCBR080

The plasmid pCBR080 encodes CwpV with a C-terminal StrepII-tag. The StrepII-tag was added by inverse PCR and an *Xho*I site prior to the tag was included to allow exchange of genes without disrupting the tag or exchange of the tag without disrupting the gene.

Cleavage domain mutants were created by inverse PCR with primers NF1308-1317 using pCBR080 as template to create plasmids pCBR081-086. The *cwpV* genes from R20352, CDKK167, M9 and AY1 were amplified with the primers

NF1096 and NF1298-1301 and cloned in place of the 630 gene in pCBR080 to create plasmids pCBR105-107 and 109 (pOE II-V).

4 Regulation of CwpV expression

4.1 Introduction

As described in the general introduction CwpV had been observed to be expressed in a phase variable manner, based on assays that relied upon the use of polyclonal antibodies raised to the CwpV N-terminus (anti-CwpVNter) and first C-terminal repeat (anti-CwpVrptI). In order to confirm the specificity of these antibodies, and therefore the reliability of the observations of CwpV phase variability, a CwpV knock-out strain was desired. Using the ClosTron system (Heap et al., 2007) Dr. Robert Fagan constructed a CwpV knock-out in *C. difficile* 630Δ*erm* and this strain, designated Δ*cwpV,* was therefore available for characterisation with anti-CwpV antibodies. For further confirmation of CwpV expression observations a *C. difficile* plasmid was desired, conferring constitutive expression of CwpV in *C. difficile*. Prior to the start of this project complementation of knock-outs in *C. difficile* using expression from plasmids had not been described, although stable *C. difficile* plasmids were available (Purdy et al., 2002). In this chapter the specificity of anti-CwpV antibodies is verified using Δ*cwpV* and constitutive expression of CwpV from a plasmid. The observations of CwpV phase variable expression are also confirmed.

In order to analyse the mechanism underlying phase variation of CwpV a reporter gene system for use in *C. difficile* was desired. Reporter gene use in *C. difficile* had been previously described; the reporter enzyme beta-glucuronidase (GusA) was used to investigate the *C. difficile* cytotoxin promoters (Mani et al., 2002). However, in this study the plasmids used were unstable in *C. difficile*. More stable *C. difficile* plasmids (Purdy et al., 2002) facilitated the development of a robust

reporter system for *C. difficile*, which is desribed in this chapter. Fluorescent reporters are useful for studying phase variation as they give information about expression at the individual cell level. Novel fluorescent proteins, which, unlike commonly used GFP-family proteins, are active under anaerobic conditions had been reported (Drepper *et al.*, 2007a) therefore the use of these fluorescent proteins as reporters in *C. difficile* was investigated.

Other inversion-based phase variation systems, where the mechanisms underpinning phase variation have been determined, rely upon inversion of the DNA sequence encoding the promoter of the gene or part of the open reading frame (ORF) (Wisniewski-Dye & Vial, 2008, van der Woude & Baumler, 2004). From the *C. difficile* genomic DNA sequence it can be seen that the region of inversion is upstream of the *cwpV* ORF, and therefore the ORF is not directly affected by DNA inversion. The location of the CwpV promoter and the effect of DNA inversion upon its activity were therefore investigated by use of reporter genes. The results presented in this chapter support the proposal of a novel mechanism responsible for phase variable expression of CwpV, where neither the orientation of the promoter nor the ORF are directly affected by DNA inversion.

Inversion of a DNA sequence flanked by inverted repeats is characteristic of the action of site-specific recombinases. In some other systems of bacterial phase variation involving DNA inversion, the site-specific recombinases are encoded in the same genetic locus as the phase variable gene, for example the *E. coli* Fim recombinases (Klemm, 1986) and the *Salmonella enterica* Hin recombinase (Heichman & Johnson, 1990). However, in the direct vicinity of *cwpV* in 630 no obvious candidate recombinase can be found. A large number of site-specific

recombinases are encoded in the 630 genome in part due to the high prevalence of mobile genetic elements (Sebaihia et al., 2006), which carry recombinases for chromosome integration and excision events. Inversion of the *cwpV* DNA switch has been detected in a diverse set of *C. difficile* strains, therefore the recombinase responsible for inversion is likely to be well conserved between strains. Based on a microarray study (Stabler et al., 2006b) it was known which of the 630 site-specific recombinases were well conserved across diverse strains. This allowed selection of recombinase candidates for study, and the results presented in this chapter identify a single site-specific recombinase that is required for *cwpV* switch inversion in *C. difficile*. Creation of a knock out of this recombinase in 630 is reported in this chapter and the effects of knocking out this recombinase on CwpV expression are determined.

4.2 Results

4.2.1 CwpV is expressed in a phase variable manner in *C. difficile* 630

The *cwpV* gene was targeted for disruption with a group II intron using the ClosTron system. Dr. Robert Fagan carried out the mutagenesis procedure and isolated a *C. difficile* clone in which the *cwpV* gene had been successfully disrupted, designated Δ*cwpV* (NF2015). The 630 *cwpV* gene was cloned into a stable *E. coli-C. difficile* shuttle plasmid pMTL960 under the control of the *cwp2* promoter, known to confer expression of Cwp2 under normal growth conditions. This plasmid, designed for constitutive expression of CwpV in *C. difficile* was named pOECwpV (pCBR045). This plasmid was transferred by conjugation to Δ*cwpV* to yield Δ*cwpV*(pOEcwpV). *C. difficile* surface layer extracts were prepared from 630Δ*erm* wild-type (WT), Δ*cwpV* and Δ*cwpV*(pOEcwpV). These were

analysed for CwpV expression by SDS-PAGE followed by Coomassie staining and Western blotting using anti-CwpVNter and anti-CwpVrpt1 antibodies. The results, shown in Figure 4.1, verify the specificity of both anti-CwpV antibodies, as no signal is detected in Δ*cwpV.* Successful expression of CwpV from pOECwpV can be seen, with CwpV constituting a significant proportion (estimated at ~25% by densitometry) of the total S-layer protein in Δ*cwpV*(pOEcwpV). CwpV is expressed on the surface of *C. difficile* as two fragments, one `~45 kDa (open triangle) and one ~120 kDa (black triangle) in size. The predicted mass of CwpV after signal peptide cleavage is 163 kDa. Given the reactivity of the ~45 kDa fragment with anti-CwpVNter and the ~120 kDa fragment with anti-CwpVrpt1 it was therefore postulated that the mature surface-expressed form of CwpV consists of two fragments, the ~45 kDa N-terminal fragment and the ~120 kDa C-terminal fragment.

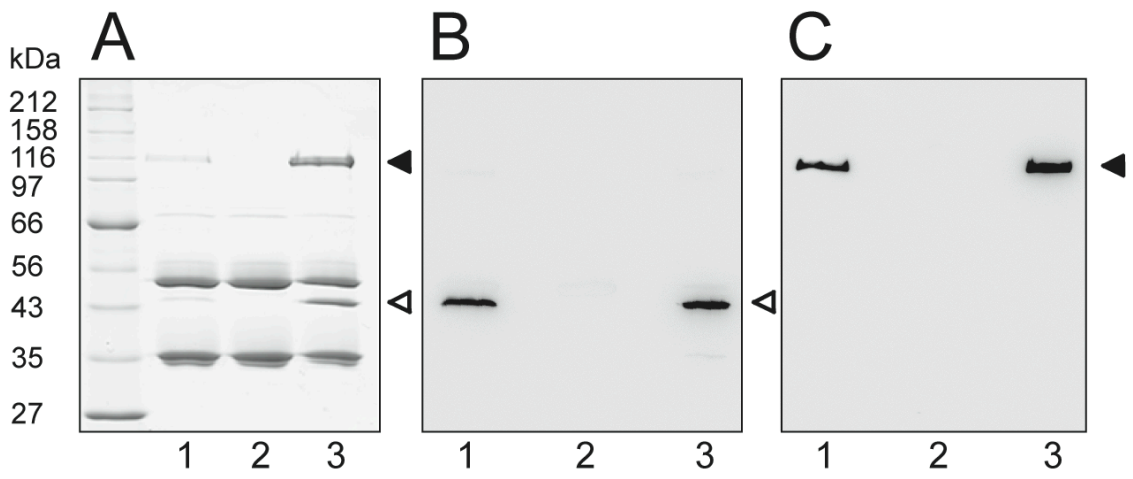


Figure 4.1 Surface expression of CwpV.

C. difficile strains were grown overnight in BHI broth. S-layer extracts were prepared and analysed by SDS-PAGE and Western blotting. **A.** Coomassie blue-stained gel. **B.** Western blot using anti-CwpVNter (1:5000). **C.** Western blot using anti-CwpVrpt1 (1:5000). Lane 1: WT; lane 2: Δ*cwpV*; lane 3: Δ*cwpV*(pOECwpV). The ~43 kDa (open triangle) and ~116 kDa (closed triangle) fragments of CwpV are indicated. For Western blots the total protein loaded in lane 3 was fivefold less than in lanes 1 and 2 to allow all bands to be visualized using one exposure time.

To characterize CwpV expression in these strains at the cellular level, rather than the average of a batch of millions of cells as seen in a surface extract, bacterial cells were stained with anti-CwpVrptI and anti-rabbit rhodamine red-X. Overnight *C. difficile* cultures were grown from single *C. difficile* colonies inoculated into standard BHI growth medium. As can be seen in Figure 4.2 WT cells express CwpV in a phase variable manner, as had been previously observed. Approximately 5% of cells express CwpV. This proportion of CwpV-positive cells was seen reproducibly in 630 cultures, with measured CwpV-positive proportions ranging from ~1-10%. The validity of this conclusion, based on staining with the anti-CwpVrptI antibody was verified by the lack of staining of all Δ*cwpV* cells. Expression of CwpV from the pOECwpV plasmids resulted in all cells expressing CwpV. This was as expected based on the usage of the *cwp2* promoter in this plasmid, as Cwp2 is known to be expressed in all cells of a clonal population (data not shown).

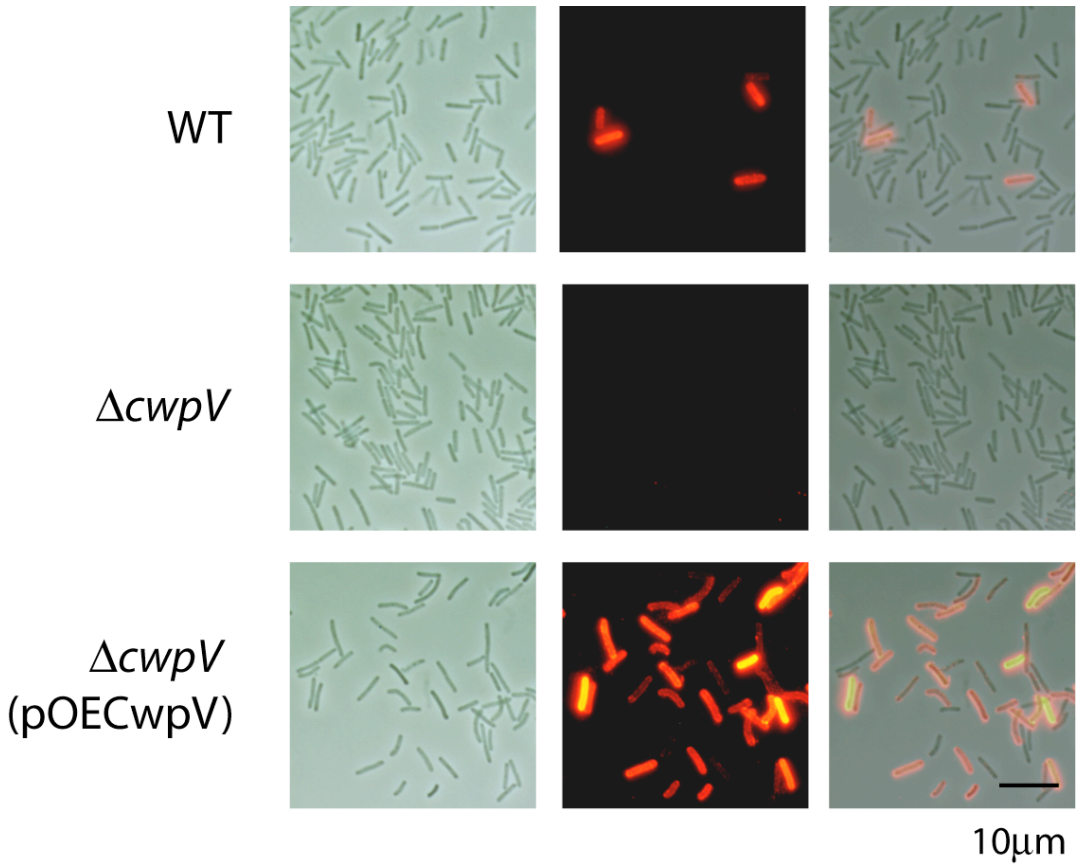


Figure 4.2 Expression of CwpV at the cellular level.

C. difficile strains; top to bottom: WT, Δ*cwpV*, Δ*cwpV*(pOECwpV), were grown overnight in BHI broth. Cells were harvested and stained for CwpV expression using anti-CwpVrpt1 and anti-rabbit rhodamine red-X antibodies. CwpV expression was detected by fluorescence microscopy. Images shown are (left to right) phase contrast, fluorescence, merged.

4.2.2 Development of a *C. difficile* reporter assay

Having confirmed and validated the observations of CwpV phase variability, the next aim was to investigate the molecular mechanism underpinning this phase variation. This required the development of a robust reporter system for use in *C. difficile*. FbFP is a flavin-based fluorescent protein developed from photoreceptors (Drepper et al., 2007a). Unlike the GFP family of fluorescent proteins it is active under both aerobic and anaerobic conditions, and therefore seemed a good candidate for use as a reporter gene in *C. difficile*. The *fbfp* gene was synthesized commercially according to the published sequence (Drepper et al., 2007a) in a *C.*

difficile codon-optimised form. This *fbfp* gene was then cloned into the pMTL960 *C. difficile* expression vector under the control of the *cwp2* promoter, in the same position as the *cwpV* gene in the previously described pOECwpV to produce pCBR012, which was transferred from *E. coli* to *C. difficile* by conjugation. For robust, reliable reporter assays it is important that the reporter plasmid is well tolerated by the bacterium, which can be assessed by determining the stability of the plasmid. The stability of the pCBR012 plasmid in *C. difficile* was assessed by serially subculturing cells containing the plasmid in the absence of thiamphenicol selection, then plating the resultant cells onto plates with and without thiamphenicol, allowing determination of the proportion of cells still containing the plasmid. In Figure 4.3A it can be seen that the plasmid is gradually lost over the 7 serial subcultures carried out, as expected due to plasmid segregation during replication. However, there is no sudden loss of the plasmid once selection is removed, indicative of a well-tolerated stable reporter plasmid.

In order to assess the usefulness of FbFP as a reporter the fluorescence of WT cells (no plasmid) was compared to cells containing pCBR012. FbFP expression was measured using a fluorescent spectrophotometer, exciting at 450nm and recording the emission at 495nm. In Figure 4.3B it can be seen that pCBR012 does increase the fluorescence of *C. difficile* compared to WT cells. However with the *cwp2* promoter, known to be a relatively strong promoter from the pOECwpV work, the fluorescence conferred by pCBR012 is only ~2.5-fold that of WT cells. This is in part due to relatively high intrinsic fluorescence of *C. difficile* cells, which has been documented and sometimes used to aid in *C. difficile* identification in a medical microbiology setting. A greater degree of sensitivity was desired for a

reporter to use for investigation into the mechanisms underpinning CwpV phase variation. Therefore the use of FbFP was not pursued further.

As mentioned in the introduction the enzyme beta-glucuronidase (GusA) has previously been used successfully in *C. difficile* (Mani et al., 2002), therefore the *fbfp* gene in pCBR012 was substituted for the *gusA* gene, producing pCBR035. GusA activity of WT 630 and 630(pCBR034) cell lysates was measured using a colorimetric GusA substrate as previously described (Mani & Dupuy, 2001). As can be seen in Figure 4.3C GusA expression driven from the *cwp2* promoter in pCBR034 can be detected and the signal:background ratio is much higher than for FbFP. This reporter therefore exhibits the sensitivity desired for investigating the molecular mechanism of CwpV phase variation. However it cannot easily give information at the cellular level, as the activity is measured on a batch culture. This is a drawback for studying phase variation, but due to the paucity of options for fluorescent reporters in anaerobes it was decided to use GusA.

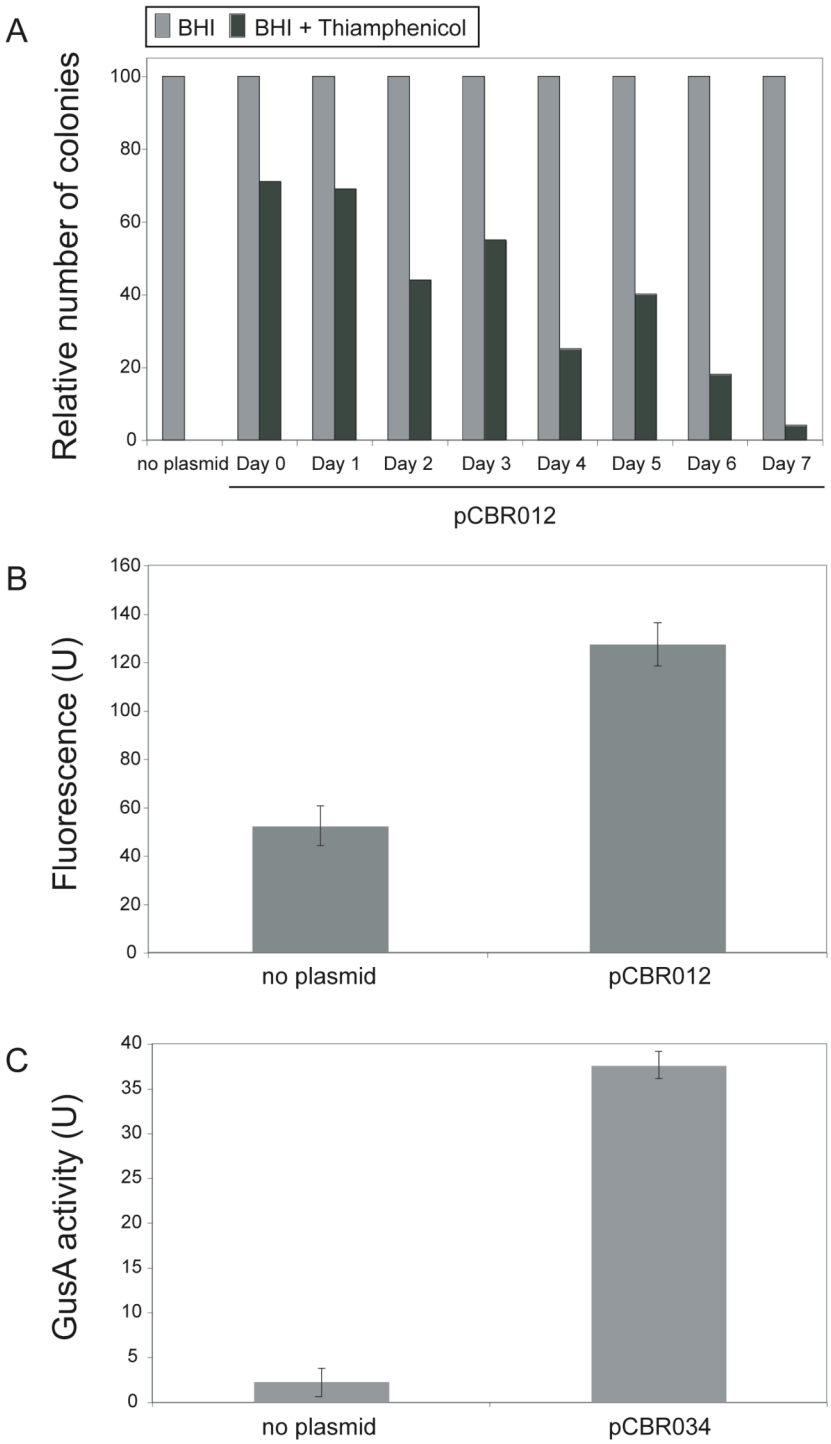


Figure 4.3 Development of a robust *C. difficile* reporter assay.

A. Stability of pCBR012 in 630. Seven successive overnight cultures of 630(pCBR012) were carried out without thiamphenicol selection. Viable counts were then carried out on BHI and BHI + thiamphenicol plates to determine the proportion of bacteria containing the plasmid. **B.** FbFP expression in 630 driven by the Cwp2 promoter. Cultures of 630 and 630(pCBR012) were grown in liquid culture to an OD of 0.4-0.8. The fluorescence at 495nm of the bacteria was measured. Three individual cultures were measured and the mean values are shown with standard deviation error bars. **C.** GusA expression in 630 driven by the Cwp2 promoter. Cultures of 630 and 630(pCBR034) were grown in liquid culture to an OD of 0.4-0.8. Bacteria were lysed and the GusA activity determined using a colorimetric assay as previously described (Mani & Dupuy, 2001).

4.2.3 Using the GusA reporter system to investigate the mechanism underlying phase variation of CwpV

Concomitant with development of the GusA reporter assay, some properties of the *cwpV* transcript were investigated in the laboratory. I will briefly summarise these findings as they were not my own work but contributed to the subsequent design of reporter constructs enabling elucidation of the phase variation mechanism. In order to locate the start site of transcription of *cwpV* 5′ RACE (Rapid Amplification of cDNA Ends) analysis on mRNA isolated from strain 630 was carried out. The 5′ end of the transcript was found to be localized 36 or 37 bases upstream of the left inverted repeat (LIR). Putative -10 and -35 sequences, closely matching those recognized by the major bacterial sigma factor σ70 in other Gram-positive promoters were identified appropriately positioned upstream of the site of transcript initiation. Interestingly, all RACE clones sequenced were observed to be in one orientation, termed the ON orientation. PCR-RFLP analysis was carried out with primers specific for the inversion region, and also clearly showed that the *cwpV* mRNA is present in the ON orientation. In contrast, the majority of the genomic DNA is present in the other orientation, termed the OFF orientation. It therefore appears that stable full-length transcripts can only be produced from genomic DNA in the ON orientation. This is consistent with the detection of CwpV expression in only around 5% of cells, possibly corresponding to those few cells with genomic DNA in the ON orientation.

To study the ability of the *cwpV* DNA switch to control transcription, a series of *gusA* transcriptional reporter constructs was created in *E. coli.* GusA expression driven by the *cwp2* promoter could act as a positive control for the reporter assay. DNA fragments located upstream of *cwpV* were amplified from 630 genomic DNA

and cloned upstream of *gusA* in order to assay their ability to drive expression. Two different lengths of fragment were used: (i) long fragments containing DNA extending 432 bp upstream of the start codon of *cwpV*, which contain the experimentally determined transcriptional start site and predicted promoter; and (ii) short fragments which extended 273 bp upstream of the start codon of *cwpV* and did not include the LIR. Plasmids with the long upstream sequence were also constructed containing a 2 bp deletion in the LIR (LIR*) or right IR (RIR*). This mutation is naturally found in the LIR of *C. difficile* strain R8366, which had been observed to not undergo DNA inversion (Emerson et al., 2009). All plasmids were generated such that the *cwpV* switch was in each of the two orientations, termed ON and OFF. Construction of these single orientation plasmids was possible because the *cwpV* switch does not invert in *E. coli* (see below). A diagrammatic representation of the different sequences tested for promoter activity is shown in Figure 4.4A.

These plasmids were transferred to *C. difficile* 630 by conjugation. Exponential cultures were analysed for GusA activity (Figure 4.4B) and for the plasmids containing long sequences the *cwpV* switch orientation was assayed by PCR (Figure 4.4C). The plasmid containing the long upstream sequence in the ON orientation (pCBR037) was able to drive expression of GusA (Figure 4.4B). Expression was also observed from the long OFF plasmid (pCBR038). PCR analysis of the plasmids with long upstream sequences after growth in *C. difficile* confirmed that both originally ON and originally OFF plasmids contained DNA in both the ON and OFF orientations (Figure 4.4C). This suggests that in *C. difficile* the *cwpV* switch on this plasmid can undergo inversion, resulting in sufficient levels of the switch in the ON orientation to allow expression of *gusA*. Strains

containing LIR* or RIR* mutant derivatives of the ON plasmids (pCBR039 and pCBR041) also expressed GusA, whereas those in the OFF orientation (pCBR040 and pCBR042) did not. PCR analysis showed that OFF plasmids with the LIR* and RIR* mutations contained DNA only in the OFF orientation, confirming that the 2 bp deletion did indeed lock the DNA in one orientation and prevent switching. This is consistent with a complete lack of GusA activity detected in these strains and a lack of expression of CwpV in strain R8366, which contains the same upstream sequence as OFF LIR*. DNA in the ON orientation and containing the 2 bp deletion in either IR drove expression of GusA, indicating that the deletion itself does not prevent expression. Therefore it is the inability to switch that causes the OFF phenotype of strain R8366. PCR analysis of ON LIR* detected some of the *cwpV* switch in the OFF orientation showing that the orientation of ON LIR* does not remain fully locked ON (Figure 4.4C), suggesting that this 2 bp deletion inhibits the OFF to ON inversion more effectively than ON to OFF. Plasmids containing the short fragment of DNA upstream of *cwpV* (pCBR035 and pCBR036) did not drive the expression of *gusA*, regardless of the orientation of the *cwpV* switch, consistent with the promoter being located upstream of the LIR. Interestingly, deletion of the entire *cwpV* switch region (pCBR043) resulted in expression of *gusA*, demonstrating that the *cwpV* promoter functions in the absence of the *cwpV* switch. This indicates that the *cwpV* switch is a negative regulator of expression when present in the OFF orientation.

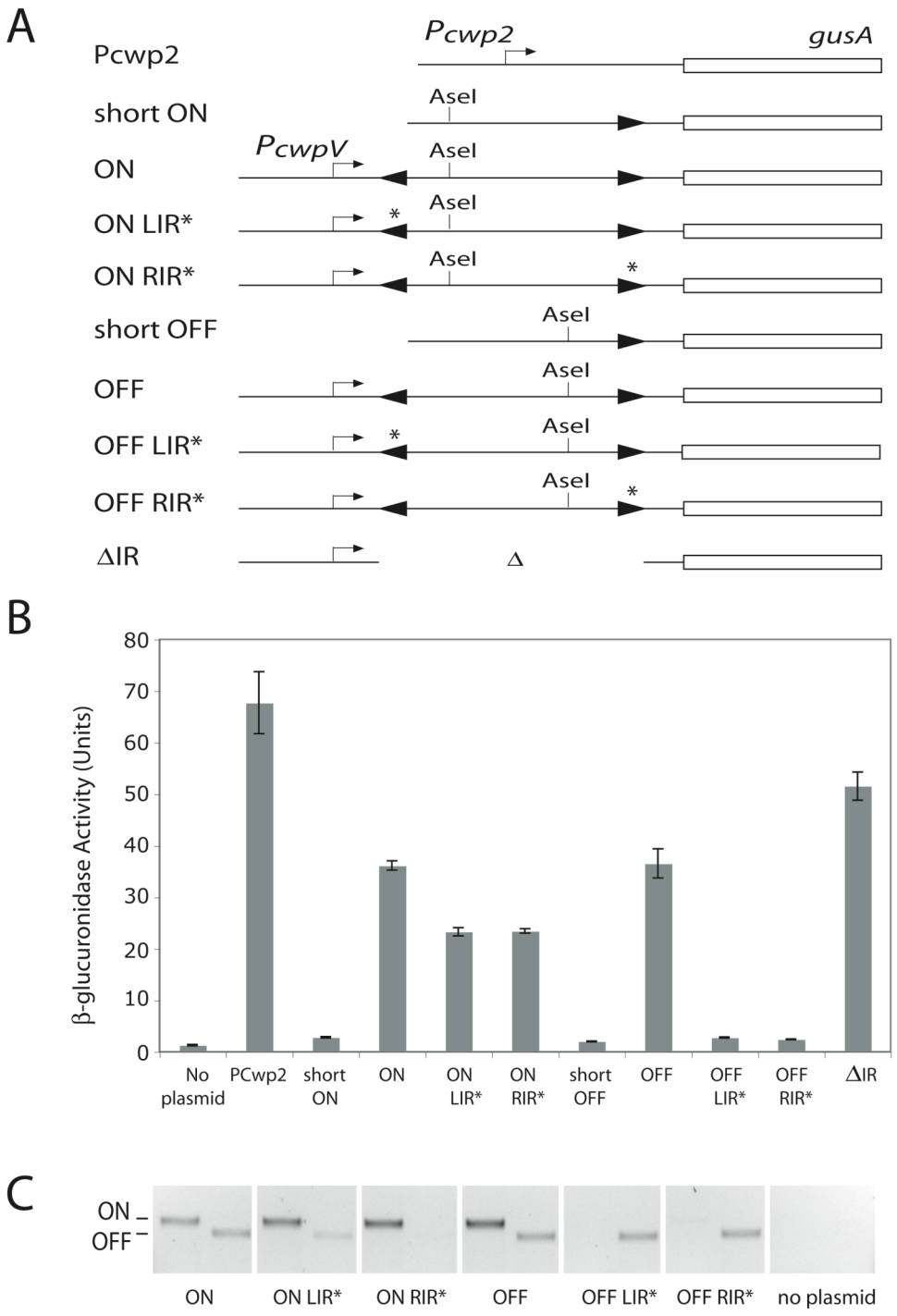


Figure 4.4 Analysis of transcription from the CwpV promoter in *C. difficile* using a GusA reporter.

C. difficile 630 containing plasmids with different components of the *cwpV* switch upstream of *gusA* were grown in liquid medium, harvested and the β-glucoronidase activity measured. **A.** Diagrammatic representation of the promoter regions present in the plasmids used in this experiment. Plasmids used were: P*cwp2,* pCBR034; short ON, pCBR035; ON, pCBR037; ON LIR*, pCBR039; ON RIR*, pCBR041; short OFF, pCBR036; OFF, pCBR038; OFF LIR*, pCBR040; OFF RIR*, pCBR042; ΔIR, pCBR043. **B.** Enzyme activities of cell lysates. Error bars represent the standard deviations of independent cell lysates (*n* = 3). **C.** PCR analysis of DNA isolated from cultures using primers to detect the ON (left) (NF793 and 826) and OFF (NF793 and 825) orientations of the *cwpV* switch.

4.2.4 A model for the *cwpV* switch mechanism

Considering the mechanisms whereby DNA inversion could act as a negative regulator of *cwpV* expression, two possibilities seem likely. First, one or more proteins could bind to the DNA or RNA in the OFF orientation and prevent transcription or translation. Second, intrinsic properties of the DNA or RNA in the OFF orientation could prevent transcription or translation. These two mechanisms are not mutually exclusive as there may be a specific structure in the OFF DNA/RNA to which protein binding could mediate the negative regulation. To explore the second hypothesis and search for the presence of intrinsic terminators, the RNA structure of transcripts from the two orientations were predicted using *mfold*, a program that predicts the structure and free energies of RNA sequences (Zuker, 2003). A stable stem loop structure consisting of a 9 bp stem containing four G-C bp was predicted to form in the OFF orientation at the junction between the LIR and the region of inversion. This stem loop structure is followed by a poly-U tract, as illustrated diagrammatically in Figure 4.5A. In the ON orientation the sequence of this region is altered by a number of base pairs (coloured red in Figure 4.5A) due to inversion of the *cwpV* switch and the predicted RNA structure is therefore different. A model for predicting intrinsic terminators in *E. coli* has been developed (Lesnik et al., 2001) and has been used to predict intrinsic terminators in a number of bacterial species including *Clostridium acetobutylicum* (Paredes et al., 2004). This model is built from sequence constraints based on known terminator sequences from *E. coli*. Briefly there must be a stem of 4–18 bp, containing at least four G-C bp, a loop of 3–10 nt and following the stem a poly-U tract with sufficient U-content. Quantitative models explaining how such structures destabilize the transcription complex have been reported (Yager & von Hippel, 1991). The structure predicted for the OFF

orientation fulfils all the criteria for an intrinsic terminator, whereas the ON stem only contains two G-C bp and has a large internal loop and therefore does not fulfil the criteria.

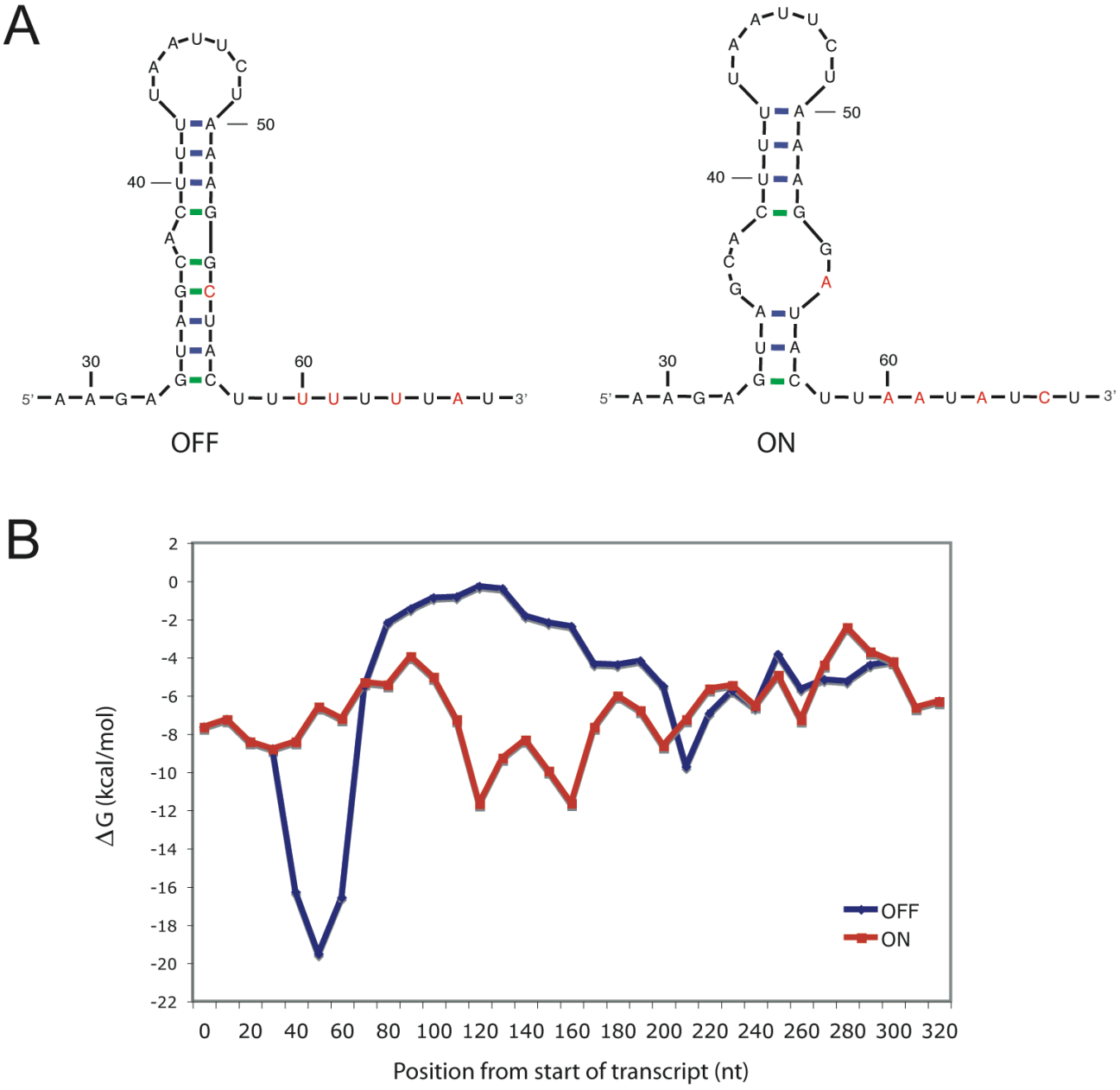


Figure 4.5 RNA structure and free energy predictions of the *cwpV* switch.

A. Predicted RNA structures found at the junction between the LIR and the region of inversion. Nucleotides that differ between the two orientations are shown in red. GC base pairing is shown in green and AU base pairing in blue. Numbering is from the predicted transcriptional start site. **B.** Free energy predictions for the two orientations, measuring the ΔG of 60 nt regions and sequentially moving along the sequence in 10 nt segments. Data were obtained using the on-line program *mfold* (Zuker, 2003). The free energy peak seen at 50 nt in the OFF orientation corresponds to the predicted RNA structure shown in A.

To further compare the stability of predicted RNA structures of the ON and OFF transcripts, quantitative predictions of free energy of RNA folding for the transcript in the two orientations were calculated for 60 nt long stretches of sequence, systematically moving along the transcript in 10 nt jumps. A peak in free energy, characteristic of an intrinsic terminator (Washio et al., 1998), is seen in the OFF orientation at the junction between the LIR and the inversion region; this peak is not present in the ON orientation (Figure 4.5B).

Based on these predictions our model for the mechanism underlying phase variation of *cwpV* is that transcription initiates upstream of the LIR regardless of the orientation of the *cwpV* switch, but in the OFF orientation transcription is terminated upon reaching the region of inversion due to an intrinsic terminator. In the ON orientation, transcription can continue through the inversion region to the ORF and produce a full-length transcript. Based on these predictions a set of six new mutants was made in the *cwpV* switch, these are depicted in Figure 4.6A. Four deletions were made to test whether the repression of expression could be attributed to a specific region of the switch. The other two new mutants were designed to change only the five base pairs that differ in the predicted stem loop region between the OFF and ON orientation (coloured red in Figure 4.5A). For ON these 5 base pairs were changed to those present in the OFF orientation and this mutant designated ON>OFF. For OFF these 5 base pairs were changed to those present in ON and this mutant designated OFF>ON. These mutants were then tested for GusA activity.

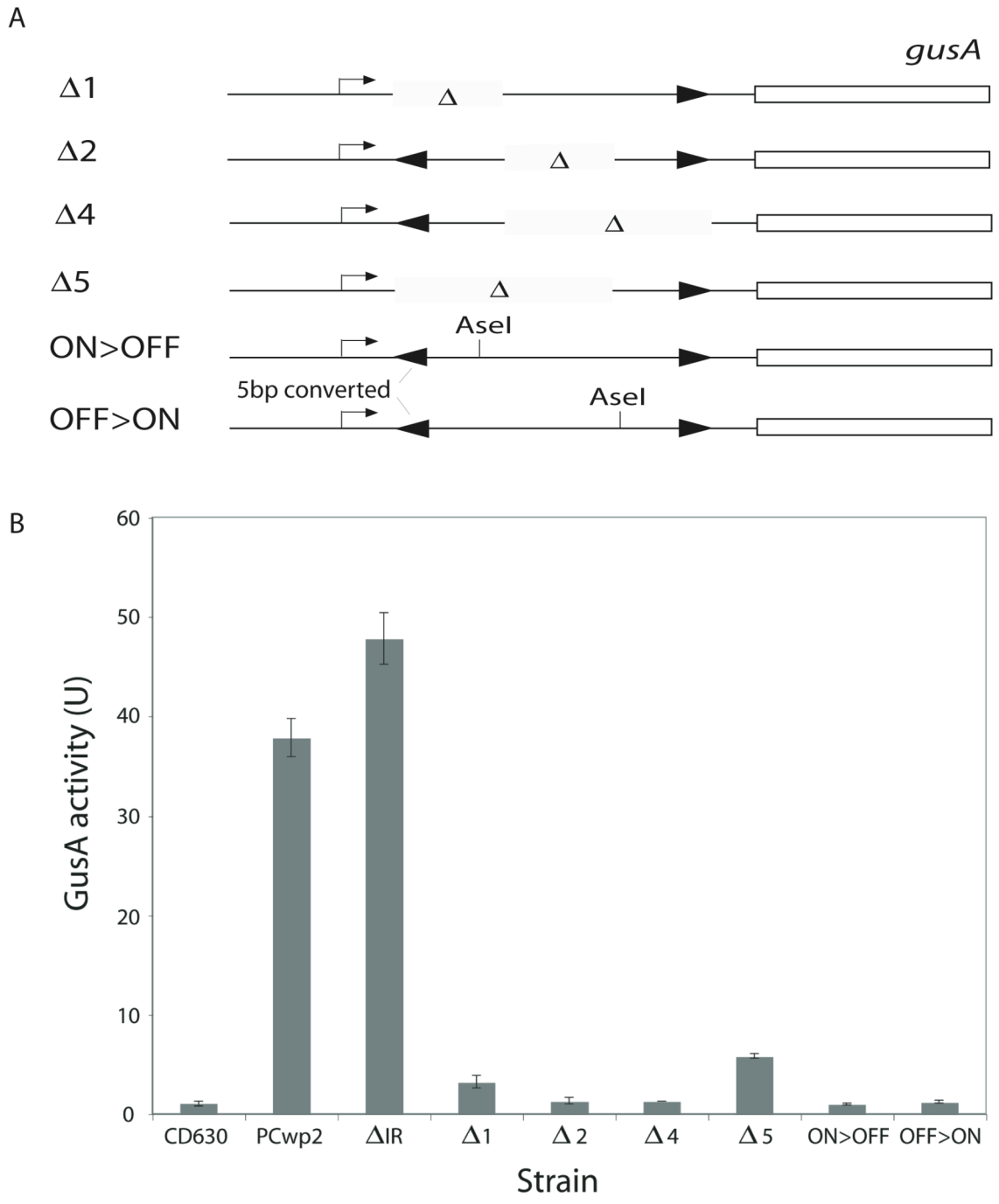


Figure 4.6 Testing the model of the *cwpV* switch mechanism.

A. Six *cwpV* switch mutants are depicted. Δ1, Δ2, Δ4 and Δ5 are deletions of specific lengths of the region of inversion. ON>OFF and OFF>ON differ from ON and OFF respectively by 5bp changes in the predicted stem loop region. **B.** The mean GusA activity of *C. difficile* strains containing these reporter plasmids was determined by colormetric assay. Three independent cultures of each strain were tested. The mean GusA activity is plotted with standard deviation error bars.

From the GusA activities shown in Figure 4.6B it can be seen that, as predicted from the stem loop terminator model, Δ2 and Δ4 exhibit the OFF phenotype. This indicates that the 5' end of the inversion region is sufficient for repression of protein expression. However, Δ1 and Δ5, where the 5' region is deleted only show a low level of GusA expression, much less than the full ΔIR phenotype. This suggests that the 3' end of the inversion region can also somehow inhibit protein expression, although it is not necessary for the OFF phenotype. Thus, there seems to be some redundancy in the system. The OFF>ON phenotype does not show GusA activity, again suggesting that there is redundancy in the mechanism, such that mutation of 5bp alone is not sufficient to lead to an ON phenotype. However, the ON>OFF does display the expected OFF phenotype, again suggesting that the predicted stem loop sequence alone is sufficient for inhibition of transcription. Therefore, overall the activity of these mutants is consistent with our model, but there appears to be redundancy in the mechanism by which the OFF orientation of the *cwpV* switch prevents expression from the *cwpV* promoter.

To further test the validity of the proposed model it was decided to test the activity of the described GusA reporter plasmids in *Clostridium perfringens.* The rationale behind this was that a mechanism relying on specialist protein factors would be unlikely to function in the same way in a different species, whereas the mechanism we propose relies upon intrinsic properties of RNA transcripts, and therefore would exhibit the same activity in different species. *Clostridium perfringens* was chosen, as being another member of the Clostridia has a very similar G:C content to *C. difficile* and it therefore seems reasonable that sensitivity of the transcriptional apparatus to RNA secondary structure would be similar between the two species. The reporter plasmids illustrated in Figure 3.4A were

transferred by conjugation to *C. perfringens* and the resulting strains tested for GusA activity. The results shown in Figure 4.7 indicate that the activity of the strains mirrors that seen for *C. difficile* in Figure 4.4B, with the exception of OFF, which does not exhibit any activity in *C. perfringens*. This is due to a lack of *cwpV* switch inversion of OFF in *C. perfringens*, meaning that the OFF strain does not contain any *cwpV* switch in the ON orientation, in contrast to what is seen in *C. difficile*. Switch inversion is investigated further below. The fact that these GusA constructs in *C. perfringens* exhibit the same pattern of activity as seen in *C. difficile* supports the model we propose for the phase variation mechanism, as it does not rely on any specialist factors, and stem loop terminators are expected to function across species boundaries.

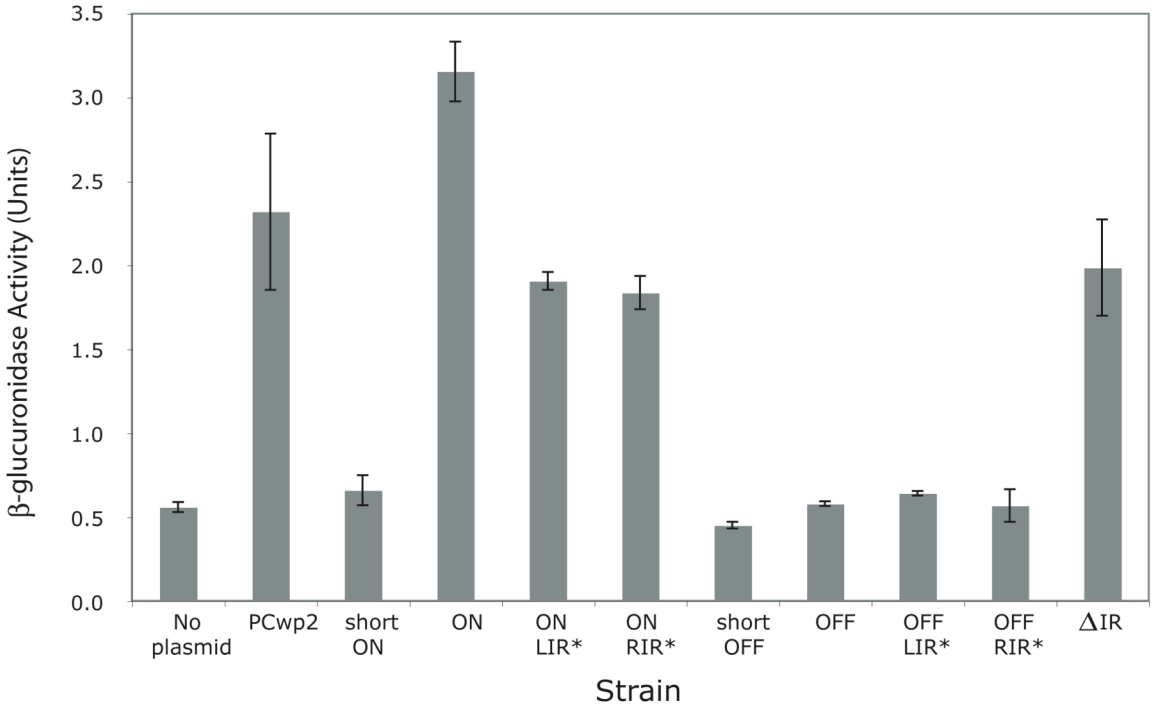


Figure 4.7 Analysis of transcription from the CwpV promoter in *Clostridium perfringens*.

C. perfringens containing plasmids with different components of the *cwpV* switch upstream of *gusA* (as illustrated in Figure 3.4A) were grown in liquid medium, harvested and the β-glucoronidase activity of cell lysates measured. Error bars represent the standard deviations of independent cell lysates (*n* = 3).

4.2.5 Inversion of the *cwpV* switch does not occur in *E. coli* or *C. perfringens*

During cloning of the *cwpV* switch region in *E .coli* and transfer of *cwpV* switch-containing plasmids to *C. perfringens* it was apparent that inversion was not occurring in these species. These observations suggest that one or more site-specific recombinases in *C. difficile* mediate DNA inversion to control expression of CwpV. The ability of different species to mediate inversion of the *cwpV* switch was further assessed by orientation PCR analysis of plasmids pCBR037 and pCBR038. These plasmids contain the *cwpV* switch in the ON and OFF orientations respectively. The orientation of the *cwpV* switch can be determined by PCR using primers which anneal to opposite strands within the region of inversion, paired with a primer outside the region of inversion, as illustrated in Figure 4.8A. In *C. difficile*, inversion of the *cwpV* switch occurred in both the ON and OFF plasmids, as shown by the appearance of PCR products using either primer NF825 or NF826 (Figure 4.8B). In contrast, in *E. coli* (Figure 4.8C) or *C. perfringens* (Figure 4.8D), no inversion was seen in four independent isolates containing either plasmid. Therefore it is likely that *C. difficile* encodes one or more specific recombinases which catalyse the inversion of the *cwpV* switch and that these enzymes are not present in *C. perfringens* or *E. coli*.

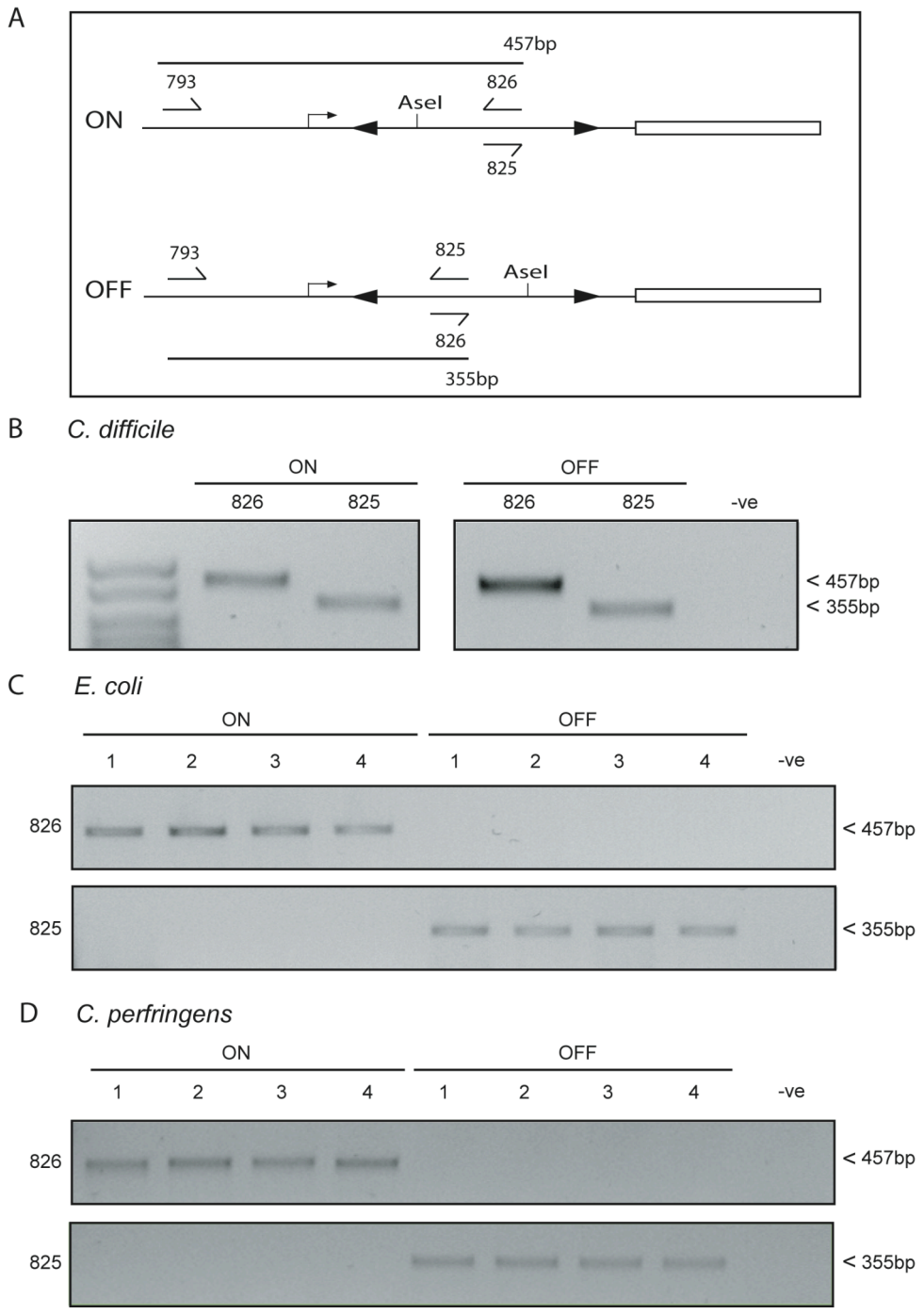


Figure 4.8 DNA inversion is specific to *C. difficile* and does not occur in C. *perfringens* or *E. coli*.

Orientation-specific PCRs were carried out to determine whether DNA inversion of the cwpV switch occurs in *E. coli* or *C.* perfringens. **A.** Diagrammatic representation of the cwpV switch region in the ON and OFF orientations, in pCRB037 and pCRB038, respectively, showing the position of the primers used to assay switch orientation. **B.** PCR products of the ON and OFF plasmids after introduction into *C. difficile*. DNA inversion occurs in each case as seen by amplification of both PCR products. **C and D**. The plasmids were introduced into *E. coli* (**C**) and *C.* perfringens (**D**). Four independent transformants or transconjugants were picked and in each case no DNA inversion is seen.

4.2.6 Identification of a *C. difficile* site-specific recombinase mediating inversion of the *cwpV* switch

A search of the *C. difficile* 630 genome (Sebaihia et al., 2006) identified 22 putative recombinases. However, 16 of these are associated with the many mobile genetic elements (MGEs) in this strain, and are largely absent from other strains as determined by microarray analysis (Stabler et al., 2006b). All *C. difficile* strains analysed to date exhibit *cwpV* switch inversion (except R8366), making it likely that the site-specific recombinase(s) responsible for inversion are well conserved across strains. A total of seven putative site-specific recombinase genes were chosen for analysis: three of the best conserved MGE-associated recombinases, one putative housekeeping recombinase and all three non-MGE associated recombinase genes. The genes were amplified from *C. difficile* 630 genomic DNA (Figure 3.9A) and each cloned into pACYCDuet-1 to allow their expression as His-tagged proteins. These recombinase-encoding plasmids were co-transformed into *E. coli* BL21(DE3) with pUC19-based plasmids containing the *cwpV* switch in either the ON or OFF orientation (pCBR026 or pCBR027). In Figure 3.9B/C it can be seen that orientation-specific PCR showed that only one recombinase, CD1167, mediated DNA inversion. This gene, which we designate *recV*, is a tyrosine recombinase of the XerC/XerD family of DNA recombinases. It can catalyse inversion of the *cwpV* switch in both the ON to OFF and OFF to ON directions. None of the other recombinases tested catalysed DNA inversion in *E. coli*. Western blotting using anti-His tag antibodies verified that all recombinases were expressed in *E. coli*, producing bands of the expected size (data not shown).

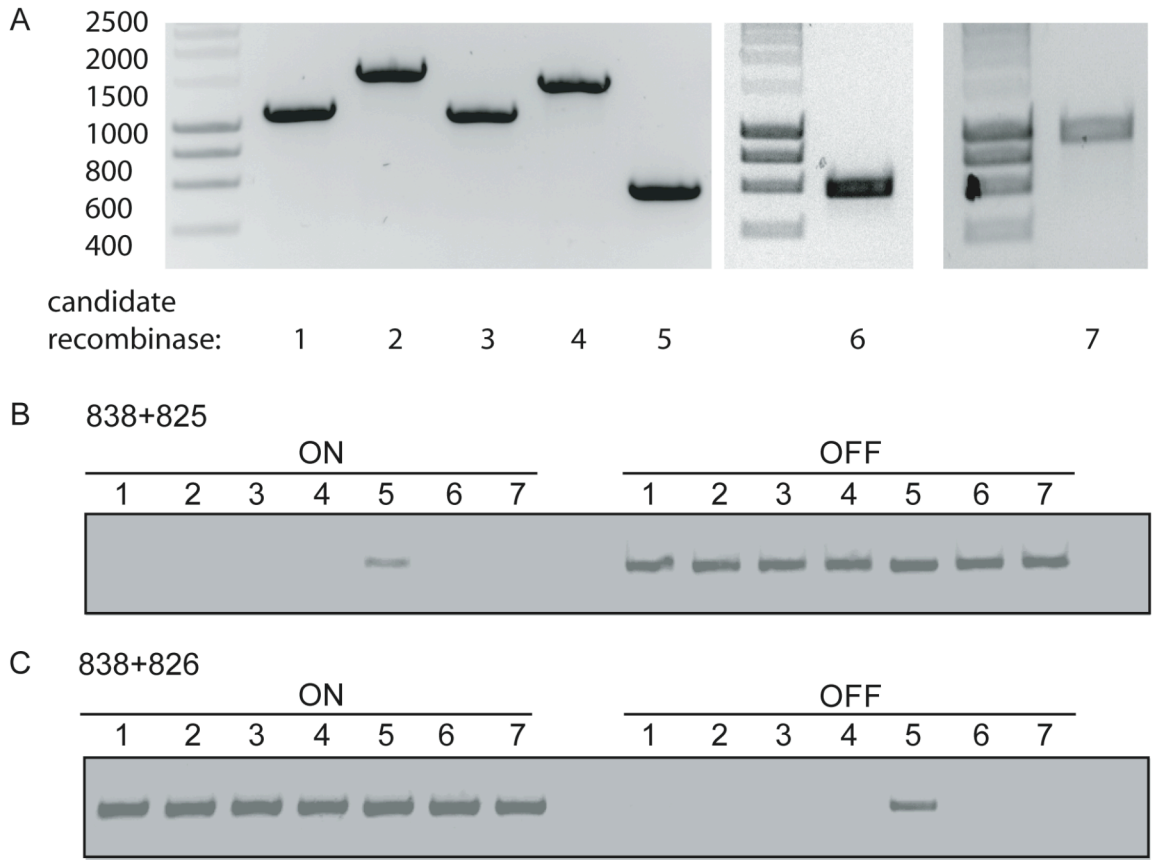


Figure 4.9 Identification of a *C. difficile* site-specific recombinase mediating DNA inversion of the *cwpV* switch.

A. Genes for seven recombinases (1; *CD1932,* 2; *CD1905,* 3; *CD2066,* 4; *CD1822,* 5; *CD1167,* 6; *CD3578* and 7; *CD1222*) were amplified from *C. difficile* 630 genomic DNA. These genes were cloned into pACYCDuet-1 and co-transformed into *E. coli* BL21(DE3) with pUC19 plasmids carrying the *cwpV* switch in either the ON or OFF orientations (pCBR026 or pCBR027). PCR reactions with primer pairs **B.** NF838 and NF825 and **C.** NF828 and NF826 were then carried out to assay *cwpV* switch orientation. Recombinase 5 (*CD1167*) mediated inversion in both the ON to OFF and OFF to ON directions.

To definitively determine the role of *recV* in inversion of the *cwpV* switch, this gene was targeted for disruption in *C. difficile* with a group II intron, using ClosTron technology (Heap et al., 2010) by Dr. Lucia de la Riva-Perez. Primers flanking the desired insertion site were designed that would amplify a 665bp product from the WT copy of *recV*. A product of 2839bp was expected following successful disruption of *recV.* WT and four erythromycin resistant clones were

tested by PCR. Figure 4.10A shows that a 2839bp fragment was amplified from all four erythromycin resistant clones, indicative of successful *recV* disruption. These clones were therefore designated Δ*recV* 1-4. In order to assay inversion of the *cwpV* switch, orientation-specific PCR was carried out. As expected orientation PCR with four clones of WT indicated inversion of the *cwpV* switch was occurring as expected, shown by products from both primer pairs in Figure 4.10Bi. All four Δ*recV* clones only amplified an OFF product (Figure 4.10Bii) and were therefore designated Δ*recVOFF*. The lack of inversion in these strains confirmed the necessary role of RecV in *cwpV* DNA switch inversion, and suggests that this is the only *C. difficile* recombinase responsible for inversion.

The *recV* gene was cloned into the *C. difficile* expression plasmid pRPF146 to yield pRecV (pCBR113). This plasmid was introduced into Δ*recVOFF* by conjugation. Four Δ*recV*(pRecV*)* transconjugants were tested by orientation PCR for inversion of the *cwpV* switch. In Figure 4.10Biii it can be seen that plasmid-encoded *recV* successfully complemented the Δ*recVOFF* mutation resulting in detection of *cwpV* switch inversion in all four transconjugants.

Bioinformatic analysis of RecV indicates that it is a member of the phage integrase family of tyrosine recombinases, and allows prediction of active site residues by comparison to characterized members of this family (for example P1 bacteriophage Cre recombinase). CwpV tyrosine-176 was identified as a putative catalytic residue by sequence alignment. To confirm the key role of this residue, site-directed mutagenesis by inverse PCR was carried out to produce pRecV$_{Y176F}$ (pCBR115). This plasmid was introduced to Δ*recVOFF* by conjugation and transconjugants selected by growth on thiamphenicol. In Figure 4.10Biv it can be

seen that all four Δ*recV*(pRecV$_{Y176F}$) transconjugants tested only had the OFF orientation of the *cwpV* switch. Therefore RecV$_{Y176F}$ could not complement the Δ*recVOFF* phenotype confirming the key role of tyrosine-176 in *recV*-mediated inversion of the *cwpV* switch.

In order to isolate Δ*recVON* mutants, Δ*recV*(pRecV) was serially subcultured without thiamphenicol selection to allow loss of the pRecV plasmid. Thiamphenicol-sensitive colonies were isolated and tested for *cwpV* switch orientation. All were found to contain a single orientation of the *cwpV* switch, as the plasmid-encoded copy of *recV* had been lost (data not shown). Four clones were isolated from which only the ON orientation of the *cwpV* switch could be amplified, as shown in Figure 4.10 Cv, and are referred to as Δ*recVON.*

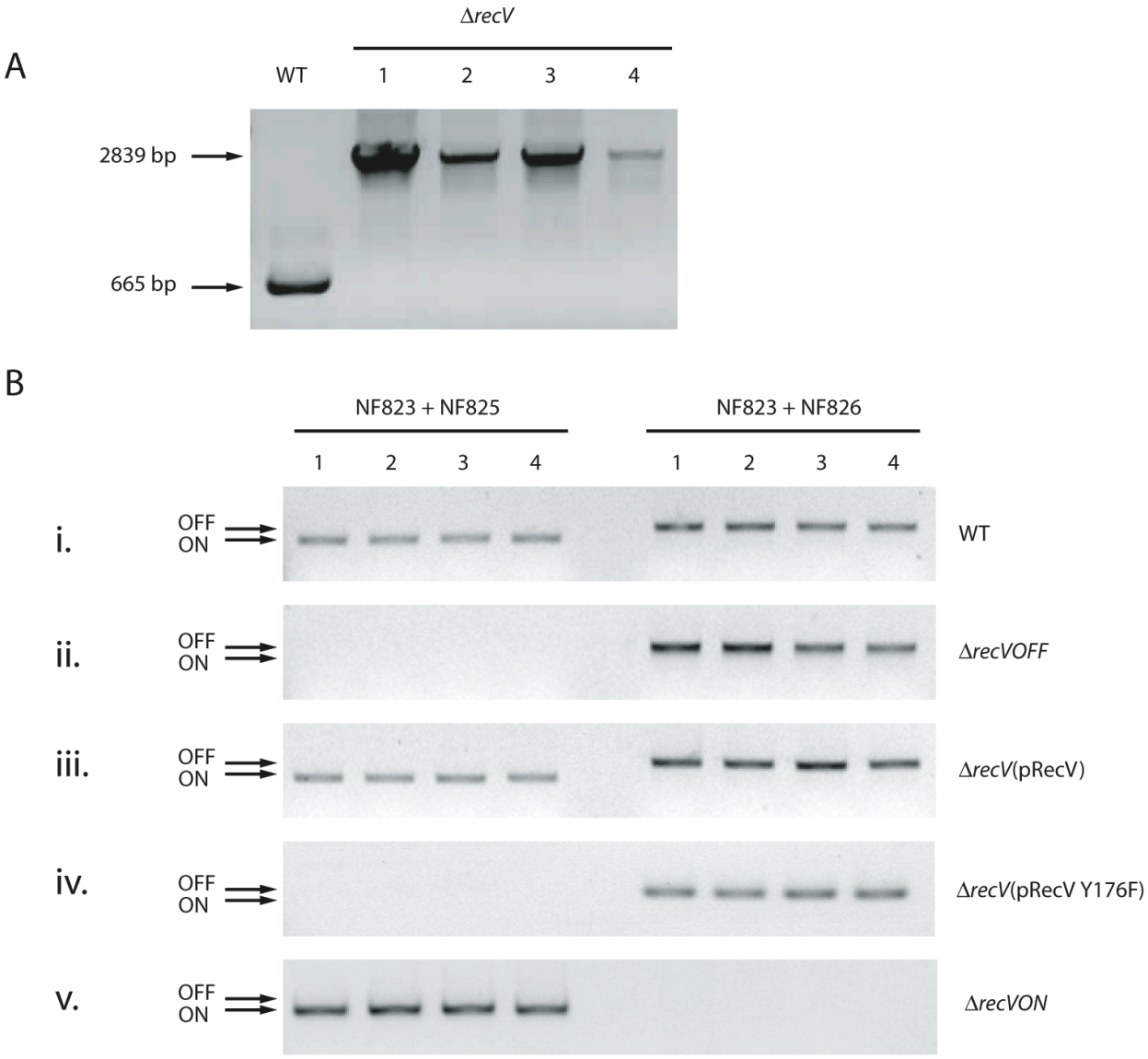


Figure 4.10 Isolation of *ΔrecV* mutants.

A. PCR confirmation of *recV* (*CD1167*) gene disruption using ClosTron. Primers NF1215+1356 flanking the ClosTron target site in *recV* give a 665bp product in 630Δ*erm* (WT). After the ClosTron procedure four erythromycin resistant colonies were tested with these primers and a 2839bp product was amplified, indicative of insertion of the group II intron into *recV.* These clones were designated Δ*recV*. **B.** Analysis of the orientation of the *cwpV* DNA switch in *C. difficile* clones by orientation PCR. i. Products for the ON and OFF orientations are amplified from WT. ii. All four isolated Δ*recV* mutants contain only the OFF orientation of the *cwpV* DNA switch, therefore these strains are referred to as Δ*recVOFF*. iii. Complementation of Δ*recVOFF* using a plasmid encoding RecV reconstituted the switching phenotype (Δ*recV*(pRecV)). iv. Alignment of RecV with other tyrosine recombinases identified Y176 as the catalytic tyrosine. A $RecV_{Y176F}$ mutant was unable to complement Δ*recVOFF* confirming the key role of this tyrosine residue in RecV activity. v. Δ*recV*(pRecV) was serially subcultured without thiamphenicol selection. Four thiamphenicol sensitive colonies were isolated from which only the ON orientation of the *cwpV* DNA switch could be amplified. These strains were therefore referred to as Δ*recVON*.

4.2.7 CwpV expression in Δ*recV* mutants

Overnight cultures of WT, Δ*recV*(pRecV), Δ*recV*(pRecV$_{Y176F}$), Δ*recVOFF* (x4) and *ΔrecVON* (x4) were grown. Samples were taken for *cwpV* switch orientation-specific PCR to confirm orientation of the *cwpV* switch in each culture. All cultures gave the orientation results expected; WT - ON/OFF, *ΔrecV*(pRecV) - ON/OFF, *ΔrecV* OFF(pRecV$_{Y176F}$) - OFF, *ΔrecVOFF* (x4) - OFF and *ΔrecVON* (x4) – ON, as shown in Figure 4.11A and B.

In Figure 4.11C *C. difficile* S-layer extracts from each culture are visualized by SDS-PAGE and Coomassie staining. The two major bands seen in each lane at ~40 kDa and ~35 kDa are the HMW SLP and LMW SLP, derived from post-translational cleavage of SlpA by the cysteine protease Cwp84. In 630 and Δ*recV*(pRecV) extracts a low level of CwpV expression can be seen with the N-terminal fragment running at ~ 42kDa (open triangle) and the C-terminal fragment running at ~ 116kDa (closed triangle), reflecting the average level of CwpV expression in the cells of the culture, where some cells are ON and others are OFF. No expression of CwpV is seen in Δ*recV*(pRecV$_{Y176F}$) or Δ*recVOFF* strains, as the *cwpV* switch is only present in the OFF orientation, and therefore full-length *cwpV* transcripts cannot be produced. Compared to WT cultures, Δ*recVON* strains express a high amount of CwpV, suggesting that in these cells CwpV constitutes a relatively high proportion of protein present in the S-layer. These findings are confirmed by Western blot using anti-CwpVrpt1 (detects the ~116 kDa CwpV protein) and anti-CwpVNter (detects the ~42 kDa CwpV protein) antibodies, as shown in Figure 4.11 D and E respectively.

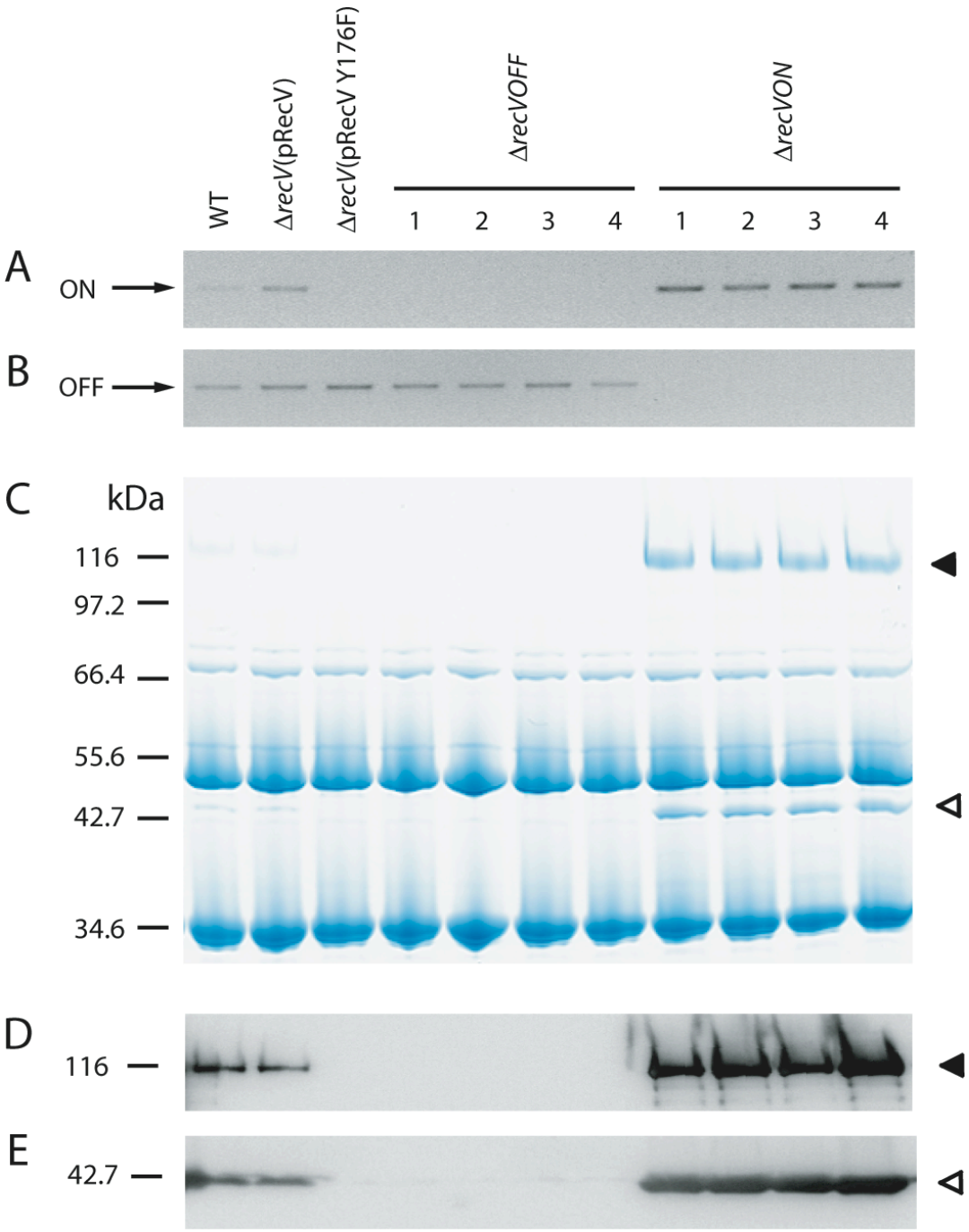


Figure 4.11 CwpV expression in Δ*recV* mutants.

Overnight *C. difficile* cultures of WT, Δ*recV*(pRecV), Δ*recV*(pRecV$_{Y176F}$), Δ*recVOFF* (x4) and Δ*recVON* (x4) were grown and analysed in the following ways: **A.** NF823+825 orientation-specific PCR amplifying the ON orientation of the *cwpV* DNA switch. **B.** NF823 + 826 orientation-specific PCR amplifying the OFF orientation of the *cwpV* DNA switch. **C.** *C. difficile* S-layer extracts analysed by SDS-PAGE and Coomassie staining. **D.** Western blot of S-layer extracts using anti-CwpVrpt1 antibody (1:5000). **E.** Western blot of S-layer extracts using anti-CwpVNter antibody (1:5000). The N-terminal (open triangle) and C-terminal (black triangle) fragments of CwpV are indicated.

In order to assess CwpV expression at the individual cell level, cultures were stained using CwpVrpt1 antibody and a fluorescent secondary antibody, then analysed by fluorescent microscopy. Representative fields of view for each strain are shown in Figure 4.12. In the WT culture ~5% of cells are ON, as previously observed. In ΔrecVOFF expression cannot be detected in any cells. In ΔrecV(pRecV) phase variable expression of CwpV can be seen with approximately 15% of cells expressing CwpV. The higher proportion of CwpV-positive cells compared to the WT is possibly due to a different level of RecV expression driven from the multi-copy expression plasmid compared to the single genomic copy of *recV* in the WT strain. In Δ*recV*(pRecVY176F) no CwpV expression can be detected, as expected due to a lack of switch inversion from OFF. In Δ*recVON*, CwpV expression can be detected in all cells. This suggests that *cwpV* switch orientation is the primary determinant of CwpV expression; provided the switch is in the ON orientation, CwpV is expressed.

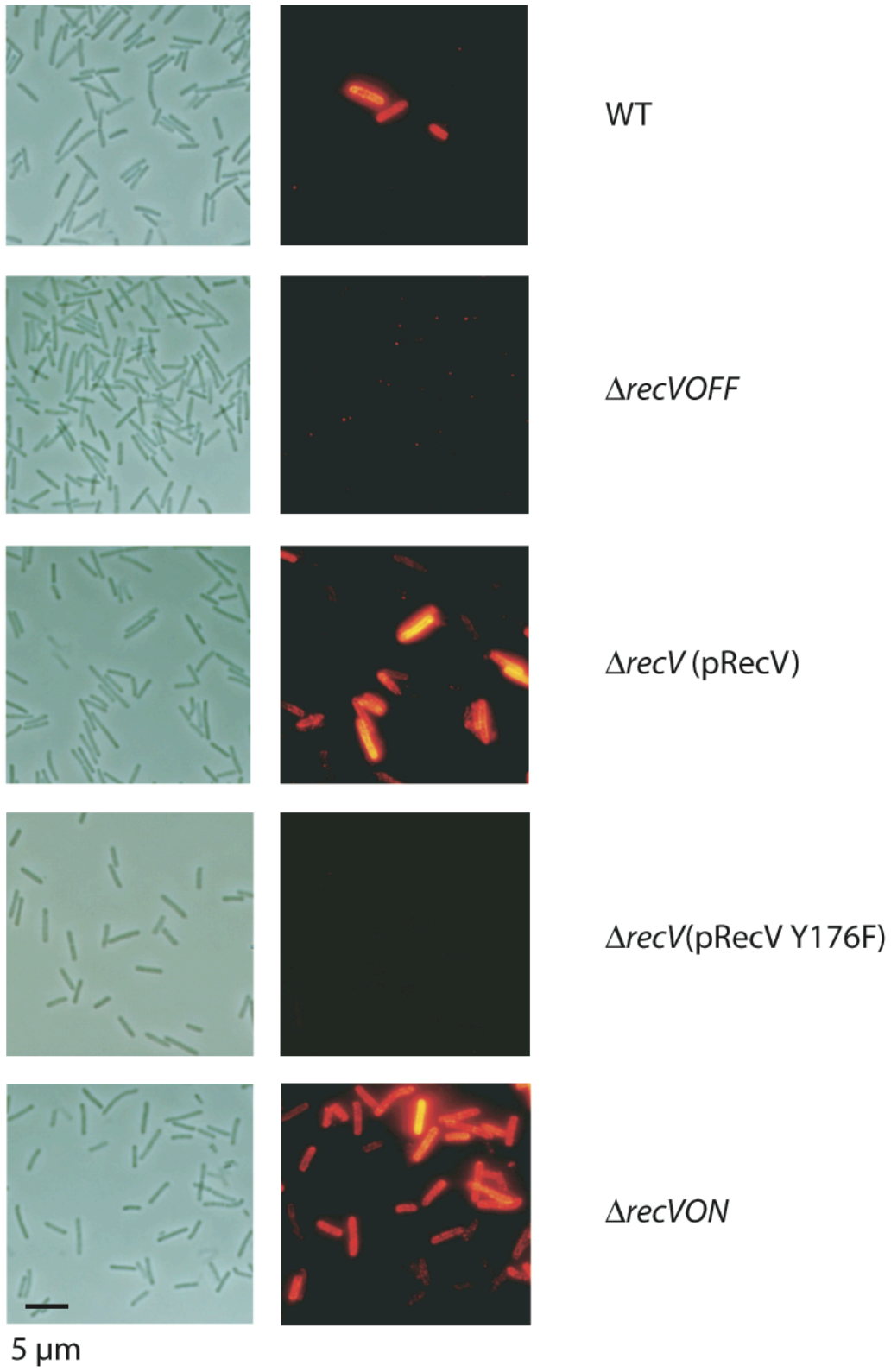


Figure 4.12 CwpV expression at the cellular level in Δ*recV* mutants.

Immunofluorescent analysis of expression of CwpV using anti-CwpVrpt1 antibody. Phase and fluorescence images of a representative field of view are shown.

4.3 Discussion

Phase variation mediated by DNA inversion has been described in a small number of bacterial systems (Wisniewski-Dye & Vial, 2008). The best-characterized systems involve inversion of a promoter or part of the ORF. In the former, a promoter in one orientation actively transcribes a gene whereas in the opposite orientation either no gene transcription occurs or an alternative gene is transcribed. For example, in the *E. coli fim* system, the *fimA* gene product is either expressed or not, depending on the promoter orientation (Abraham et al., 1985). In the *Salmonella hin* system, in one orientation the promoter directs expression of the H1 flagellin, whereas in the opposite orientation it directs expression of H2 flagellin (Zieg et al., 1977).

We have established that the CwpV promoter and transcriptional initiation site are situated upstream of the *cwpV* switch and that DNA inversion controls the expression of *cwpV.* The production of stable full-length *cwpV* transcripts is only possible from template genomic DNA with the *cwpV* switch in the ON orientation. This mechanism for controlling phase variable gene expression involving DNA inversion encompassing neither the promoter nor the ORF is to our knowledge completely novel. By using a *gusA* transcriptional reporter in *C. difficile*, we have shown that the *cwpV* switch region is not required for expression from the *cwpV* promoter, and that the *cwpV* switch must therefore act *in cis* to negatively regulate expression. We propose a model to account for these observations based on predicted secondary structures of the mRNA transcripts from the two orientations. This mechanism involving transcriptional termination modulated by DNA inversion has not been reported for any other system. Transcripts in the OFF orientation are

predicted to form a stable stem loop structure followed by a poly-U tract, which induces transcriptional termination at some 60–70 nucleotides downstream of the transcript initiation site. In the ON orientation the predicted structure does not fulfil the criteria for an intrinsic terminator and transcription can proceed unhindered. Further mutagenesis studies support the key role of the stem-loop terminator in conferring the OFF phenotype of the OFF orientation. This is particularly clear from the ON>OFF mutation that was made, where changing just 5 bp of ON to confer the properties of the stem-loop completely abrogated expression. However, these mutants also suggest that there is some redundancy in the system, because other regions of the OFF orientation of the *cwpV* switch also hinder expression. We acknowledge the possibility that, in the OFF orientation, transcription could also be affected by binding a repressor molecule or that control could be mediated by a combination of these mechanisms. However, the similar functioning of the *cwpV* switch in *C. perfringens* suggests that there are not specialist factors involved, as these would be unlikely to be present in C. *perfringens*. Our model predicts function of the *cwpV* switch across species boundaries, as stem-loop termintors are ubiquitous in bacterial species. The observations from C. *perfringens* therefore support our model.

Interestingly modulation of transcription termination is also seen in the *E. coli fim* system, where a Rho-dependent transcriptional terminator is present downstream of the *fimE* recombinase gene in phase OFF cells ((Joyce & Dorman, 2002, Hinde *et al.*, 2005). In phase ON bacteria, termination of the *fimE* transcript does not occur and the mRNA transcript is longer and more stable, leading to increased levels of FimE recombinase and biasing the switch to the phase OFF. However, in the *fim* switch, the primary level of control of *fimA* expression is inversion of the

promoter, whereas in the *cwpV* switch described here, transcription termination is the key event determining expression of *cwpV*.

The recombinase RecV was able to mediate DNA inversion in both the ON to OFF and OFF to ON orientations when expressed in *E. coli*, suggesting that if other protein factors are necessary for *cwpV* switch DNA inversion they are highly conserved across the bacterial kingdom. Using targeted gene knock-out we have shown that in *C. difficile* 630 the tyrosine recombinase RecV is essential for *cwpV* DNA switch inversion. This suggests that there are no other site-specific recombinases expressed in 630 that can mediate *cwpV* DNA switch inversion. In well-characterised phase variation systems employing DNA inversion the recombinase(s) involved in DNA inversion are located either within the inverted DNA (*S. typhimurium hin*) or within the immediate vicinity of the switch (*E. coli fim*) (Wisniewski-Dye & Vial, 2008). It is interesting that *recV* is not located in the vicinity of *cwpV*. This fact, along with the presence of the *cwpV* switch and *recV* in all *C. difficile* strains analysed within a diverse set of isolates, suggests that this switch and *recV* have not been recently acquired by *C. difficile* via horizontal gene transfer. Rather it appears that *recV* is a member of the relatively small core set of genes common to all *C. difficile* strains, and has been part of the evolution of *C. difficile* over a long timescale.

Analysis of CwpV expression in Δ*recV* mutants showed that the *cwpV* switch orientation is the primary determinant of CwpV expression. If the switch is in the ON orientation CwpV expression will proceed, and if the switch is in the OFF orientation CwpV is not expressed. Analysis of CwpV expression in the Δ*recVON* and Δ*recVOFF* mutants throughout the growth phase gave no indication of further

levels of CwpV expression regulation, as expression levels were constant throughout (data not shown). Complementation of *ΔrecV* with a plasmid-encoded copy of *recV* restored *cwpV* switch inversion, and in fact led to a higher proportion of cells expressing CwpV (~15%) than in the WT (~5%). The is likely due to differences in RecV expression levels between *ΔrecV*(pRecV+) and WT. The relationship between RecV expression level and *cwpV* switch orientation has not yet been investigated. In the well-characterised *E. coli fim* switch, inherent differences in affinity of site-specific recombinases for different orientations of the switch determine the relative kinetics of switching in either direction and therefore determine the overall proportion of switch orientations (Gally et al., 1996). Having definitively determined the role of RecV in *cwpV* switch inversion the mechanistic details of this switch system can be investigated.

A common theory to account for phase variation in bacterial pathogens is that alteration of surface structures allows evasion of the immune responses. Other explanations include modulation of adhesin expression that could facilitate detachment of bacteria from host substrates, resulting in dissemination of the bacteria from the host, or modulation of factors involved in biofilm formation. However, in many cases the biological significance of phase variation remains a mystery (van der Woude, 2006). Investigation into the function of CwpV is the focus of chapter 6.

5 CwpV post-translational processing and expression in the *C. difficile* S-layer: Conservation and Diversity across strains

5.1 Introduction

As shown in Chapter 4 CwpV is expressed on the *C. difficile* surface as two fragments; the N-terminal fragment (~42 kDa) and the C-terminal fragment (~116 kDa). The implications of this finding are unclear. The N-terminal fragment contains the putative cell wall binding domains, and is therefore presumed to bind to the cell wall. The function of these PF04122 domains has however never been addressed experimentally. The ability to express CwpV family from a plasmid-encoded copy of the gene in *C. difficile* allows investigation into the role of different CwpV domains by mutagenesis, and the results of such a mutagenesis are shown in this chapter.

The mechanism of CwpV cleavage is also unknown. SlpA is cleaved into the HMW and LMW SLP fragments by the cysteine protease Cwp84 (Kirby et al., 2009, Dang et al., 2010). Cwp84 and another protease of the CWP family Cwp13 are therefore candidates for carrying out CwpV cleavage. The location of the CwpV cleavage site may provide insight into the mechanism of cleavage. The experimentally determined sequence of the CwpV cleavage site is reported in this chapter.

Following cleavage of CwpV it is known that both fragments are expressed anchored to the cell surface (Emerson et al., 2009). The C-terminal fragment does

not contain any known cell wall binding domains, and therefore the mechanism by which it is anchored to the cell surface required investigation. The LMW SLP is thought to be anchored to the cell surface due to the complex that it forms with the HMW SLP, which contains the putative cell wall binding domains (Fagan et al., 2009). It therefore seemed possible that the C-terminal domain of CwpV may form a complex with proteins containing cell wall binding domains in order to remain anchored to the cell wall after CwpV cleavage. The results of an investigation into this hypothesis are presented in this chapter.

Given the known diversity of the CwpV C-terminus across different *C. difficile* strains, the significance of this variation was investigated. Bacterial cell surface proteins are under significant immune pressure, and this can lead to diversifying selection based on differing antigenicities of proteins with sequence variability. For example, the LMW SLP exhibits a high degree of sequence variation, which is thought to affect its antigenicity (Karjalainen *et al.*, 2002b). Therefore the antigenicity of the different CwpV repeat types was investigated, and the results are presented in this chapter.

Although the C-terminal repeat domains are highly variable across diverse *C. difficile* strains, the N-terminal sequences are well-conserved. Therefore the level of conservation of mechanisms underpinning CwpV post-translational processing and expression in the S-layer was investigated for CwpV proteins of different types.

As discussed in Chapter 4, CwpV is expressed in a phase variable manner in 630. The *cwpV* switch and *recV* appear well conserved across *C. difficile* strains,

therefore it was postulated that CwpV is likely to be expressed in a phase variable manner in a diverse set of *C. difficile* strains. Strains encoding all types of CwpV repeat were therefore investigated for regulation of CwpV expression at the cellular level.

An understanding of the differences and similarities between strains with respect to CwpV expression informed subsequent investigations into the function of CwpV, which is the focus of Chapter 6.

5.2 Results

5.2.1 CwpV protein domains in *C. difficile* 630

The sequence of *cwpV* in *C. difficile* 630 encodes the largest protein of the CWP family. Three copies of the putative cell wall binding domain PF04122 are encoded at the N-terminus of the protein. There is a domain of unknown function downstream of the cell wall binding domains, followed by a 30 amino acid serine/glycine rich linker (S/G linker) containing 18 serine residues and 9 glycine residues. This is presumed to function as a flexible linker. Then at the C-terminal end of the protein nine 120 amino acid repeats are encoded. A diagrammatic representation of the 630 CwpV protein is shown in Figure 5.1A with the amino acid sequence colour coded in the same manner shown in Figure 5.1B. An alignment of the nine repeat sequences against each other is shown in Figure 5.1C. It can be seen that the degree of homology between the repeats is very high, with the first repeat being slightly more divergent that the other repeats.

A



Black = signal peptide
Grey = Cell wall binding domains
White = unknown function
Pink = S/G linker
Blue = repeats

B

```
MKIPKKIAAM LTVTMIAGSS TAGIASAQTV ATNLTGQERY ETAVKISQDG WKNADEVVIV      60
NDSSIADALS ATPFAKAKNA PILLTSKDKL NDKTKAEIQR LKAKKVYLIG GTSVLSTNIE     120
KEIKDLKISF ERISGAERYQ TSLELAKKLD AISDVKKIAV VNGEKGLADA VSVGAPAAQN     180
NMPVILADSK NGTAVADKFI KDAGITQSYV VGGESSISEA VKNKLPNSTR LGGTDRNDTN     240
AKVIKEFYKK TDLKNAYVTK DGMNKQDQLI DALAVGVLGA KNQSPVVLVG KNLSASQKSL     300
VNSKSFDKIT KVGGNGNETA FNEMKSLQEV KTVEAKTISE LKSAIDKATA NDVINFKPTS     360
EVKEAFTIQT DKAVTVNLNG TYTKTVTINM PNGDVNNYAK VDDVVIDDVK DGTFVNYGKI     420
TNLKVNDKNG AKIENNSKGE IGSLTVASGA SQVKVTNGGK ITTVTNNSKG TTIDNKGTIS     480
SVKGDNSPTI SGNSPSSNSS GGSSSSGGSS GGSSGGSSSS KVGQIKFSAT DKQIKMASGD     540
KTVDPSDNTY ILKLTNGTVK DGDVKANVKV AGLPSGLDYT AVGNRSANTI TITVSGTASQ     600
SIQNDLNNVS VLVKAGAVSG TGATDSIANT TFEIKKYVAP AIGQLTFDAT DKQIKMASGD     660
KTVDPSDDTY VLTLTKGTAK DGDVKANVEV TGLPSGLDYT AEGDSSTNTI TITVSRTASQ     720
AVQTDLDNVS VLVKAGAVSE TTATDSAANA TFEVKKYVAA PAVGQLTFDA TDKQIKMASG     780
DKTVDPSDDT YVLTLTKGTA KDGDIKANVE VTGLPSGLDY TAEGDSSANT ITITVSGTAS     840
QAVQTDLDNV SVLVKAGAVS ETTATDSAAN ATFEVKKYVA APAVGQLTFD ATDKQIKMAS     900
GDKTVDPSDD TYVLTLTKGT AKDGDVKANV EVTGLPSGLD YTAEGDSSAN TITITVSGTA     960
SQAVQTDLDN VSVLVKAGAV SETTATDSAA NATFEVKKYV APAIGQLTFD ATDKQIKMAS    1020
GDKTVDPSDD TYVLTLTKGT AKDGDVKANV EVTGLPSGLD YTAEGDSSAN TITITVSGTA    1080
SQAVQTDLDN VSVLVKAGAV SETTATDSAA NATFEVKKYV APAIGQLTFD ATDKQIKMAS    1140
GDKTVDPSDD TYVLTLTKGT AKDGDVKANV EVTGLPSGLD YTAEGDSSAN TITITVSGTA    1200
SQAVQTDLDN VSVLVKAGAV SETTATDSAA NATFEVKKYV APAIGQLTFD ATDKQIKMAS    1260
GDKTVDPSDD TYVLTLTKGT AKDGDVKANV EVTGLPSGLD YTAEGDSSAN TITITVSGTA    1320
SQAVQTDLDN VSVLVKAGAV SETTATDSAA NATFEVKKYV APAIGQLTFD ATDKQIKMAS    1380
GDKTVDPSDD TYVLTLTKGT AKDGDVKANV EVIGLPSGLD YTAEGDSSAN TITITVSGTA    1440
SQAVQTDLDN VSVLVKAGAV SETTATDSAA NATFEVKKYV APAIGQLTFD ATDKQIKMAS    1500
GDKTVDPSDD TYVLTLTKGT AKDGDVKANV EVIGLPSGLD YTAEGDSSAN TITITVSGTA    1560
SQAVQTDLDN ISVLVKAGAV SETTATDSAA NATFEVKKHA SDPISTIGKI SS
```

C

```
rpt6 GQLTFDATDKQIKMASGDKTVDPSDDTYVLTLTKGTAKDGDVKANVEVTGLPSGLDYTAE 60
rpt7 GQLTFDATDKQIKMASGDKTVDPSDDTYVLTLTKGTAKDGDVKANVEVTGLPSGLDYTAE 60
rpt5 GQLTFDATDKQIKMASGDKTVDPSDDTYVLTLTKGTAKDGDVKANVEVTGLPSGLDYTAE 60
rpt4 GQLTFDATDKQIKMASGDKTVDPSDDTYVLTLTKGTAKDGDVKANVEVTGLPSGLDYTAE 60
rpt2 GQLTFDATDKQIKMASGDKTVDPSDDTYVLTLTKGTAKDGDVKANVEVTGLPSGLDYTAE 60
rpt3 GQLTFDATDKQIKMASGDKTVDPSDDTYVLTLTKGTAKDGDIKANVEVTGLPSGLDYTAE 60
rpt9 GQLTFDATDKQIKMASGDKTVDPSDDTYVLTLTKGTAKDGDVKANVEVIGLPSGLDYTAE 60
rpt8 GQLTFDATDKQIKMASGDKTVDPSDDTYVLTLTKGTAKDGDVKANVEVIGLPSGLDYTAE 60
rpt1 GQIKFSATDKQIKMASGDKTVDPSDNTYILKLTNGTVKDGDVKANVKVAGLPSGLDYTAV 60
     **:.*.*******************:**:*.**:**.****:****:* **********

rpt6 GDSSANTITITVSGTASQAVQTDLDNVSVLVKAGAVSETTATDSAANATFEVKKYVAPAI 120
rpt7 GDSSANTITITVSGTASQAVQTDLDNVSVLVKAGAVSETTATDSAANATFEVKKYVAPAI 120
rpt5 GDSSANTITITVSGTASQAVQTDLDNVSVLVKAGAVSETTATDSAANATFEVKKYVAPAI 120
rpt4 GDSSANTITITVSGTASQAVQTDLDNVSVLVKAGAVSETTATDSAANATFEVKKYVAPAI 120
rpt2 GDSSTNTITITVSRTASQAVQTDLDNVSVLVKAGAVSETTATDSAANATFEVKKYVAAPA 120
rpt3 GDSSANTITITVSGTASQAVQTDLDNVSVLVKAGAVSETTATDSAANATFEVKKYVAAPA 120
rpt9 GDSSANTITITVSGTASQAVQTDLDNISVLVKAGAVSETTATDSAANATFEVKKHASDPI 120
rpt8 GDSSANTITITVSGTASQAVQTDLDNVSVLVKAGAVSETTATDSAANATFEVKKYVAPAI 120
rpt1 GNRSANTITITVSGTASQSIQNDLNNVSVLVKAGAVSGTGATDSIANTTFEIKKYVAPAI 120
     *: *:******** ****::*.**:*:********** * **** **:***:**:.: .
```

Figure 5.1 The CwpV protein from 630.

A. Diagrammatic representation of the CwpV protein from 630. **B.** Amino acid sequence of the CwpV protein from 630 colour coded in the same way as the diagram in A. Predicted PF04122 domains are underlined. **C.** Multiple sequence alignment (ClustalW2) of the nine 120 amino acid repeats from 630 CwpV.

5.2.2 Determination of the 630 CwpV cleavage site

Analysis of cell surface expression of CwpV indicates that CwpV is cleaved. This was shown in the previous chapter and is clear from analysis of surface extracts analysed by SDS-PAGE followed by Coomassie staining and Western blotting, as shown in Figure 5.2A-C. In order to determine the CwpV cleavage site, a *C. difficile* surface extract from Δ*cwpV*(pOECwpV) was transferred to PVDF membrane and visualized via Coomassie staining. The C-terminal fragment (~116 kDa) of CwpV was identified and this membrane sent to the PNAC at the University of Cambridge for N-terminal sequencing by Edman degradation. This yielded the sequence TFVNY, unambiguously locating the CwpV cleavage site in the domain of unknown function, between the cell wall binding domains and the S/G linker, as depicted in Figure 5.2D diagrammatically.

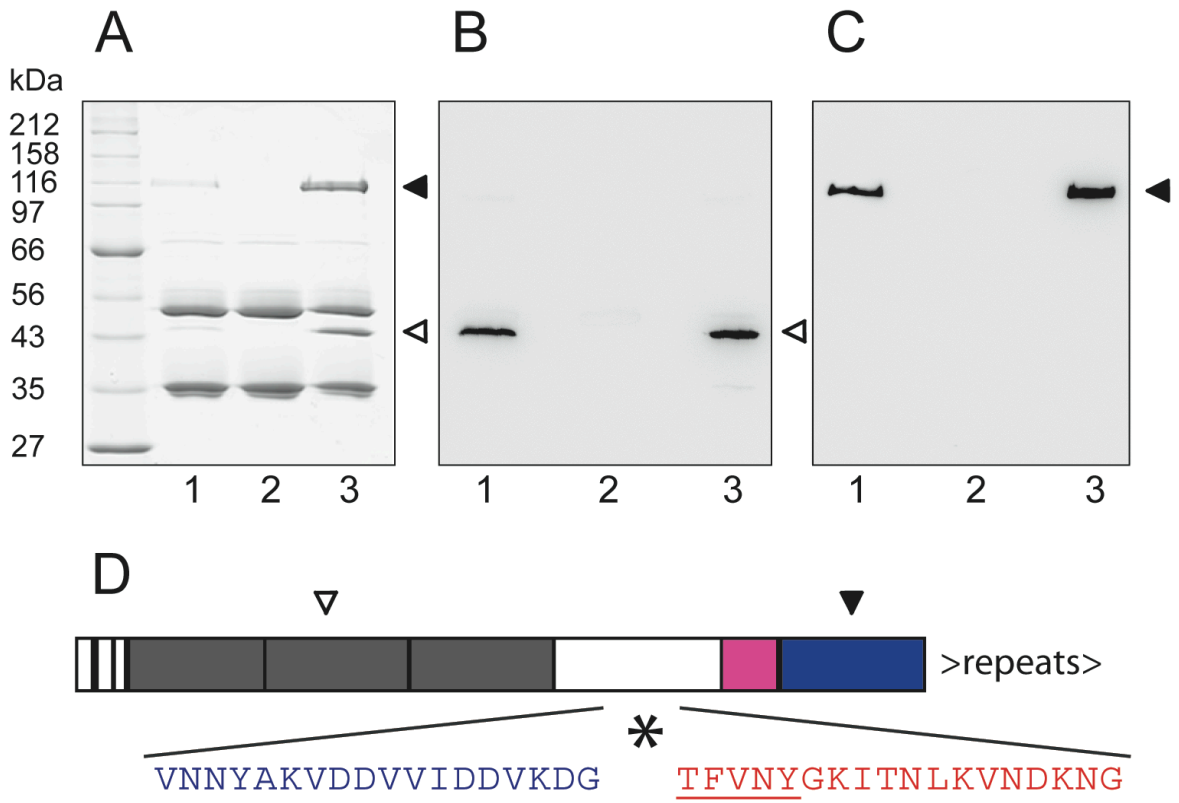


Figure 5.2 Determination of the 630 CwpV cleavage site.

C. difficile strains were grown overnight in BHI broth. S-layer extracts were prepared and analysed by SDS-PAGE and Western blotting. **A.** Coomassie blue-stained gel. **B.** Western blot using anti-CwpVNter (1:5000). **C.** Western blot using anti-CwpVrpt1 (1:5000). Lane 1: WT; lane 2: Δ*cwpV*; lane 3: Δ*cwpV*(pOECwpV). The ~42 kDa (open triangle) and ~116 kDa (black triangle) fragments of CwpV are indicated. For Western blots the total protein loaded in lane 3 was fivefold less than in lanes 1 and 2 to allow all bands to be visualized using one exposure time. **D.** N-terminal sequencing of the CwpV C-terminal fragment (black triangle) yielded the amino acid sequence TFVNY. This unambiguously locates the cleavage site in the middle of the domain of unknown function, between the cell wall binding domains and the S/G linker, as depicted here diagrammatically.

5.2.3 The N-terminal fragment of CwpV mediates cell wall binding

In order to experimentally address the function of the N-terminal fragment of CwpV containing the putative cell wall binding domains, the *cwpV* gene in the plasmid pOECwpV was mutated by inverse PCR to produce a truncated version of the *cwpV* gene in a new plasmid pOECwpVNter. This truncation is depicted in Figure 5.3A. Analysis of surface layer extracts from WT, Δ*cwpV,* Δ*cwpV*(pOECwpV) and Δ*cwpV*(pOECwpVNter) by SDS-PAGE followed by

Coomassie staining and anti-CwpVNter Western blot was carried out, and the results are shown in Figure 5.3B and C. The truncated CwpV is processed to form the same N-terminal fragment as the full-length CwpV, which is expressed in the *C. difficile* S-layer. This suggests that the N-terminal fragment of CwpV alone is sufficient for stable association with the cell wall. The same amount of the N-terminal CwpV fragment appears to be present in surface extracts from *Δcwp*V(pOECwpV) and *Δcwp*V(pOECwpVNter) surface extracts, showing that the truncated version of CwpV is equally well expressed on the cell surface as the full-length CwpV protein. The N-terminal fragment of CwpV has therefore been named the cell wall anchoring (CWA) fragment, and this is the first reported experimental evidence that PF04122 motifs are responsible for cell wall anchoring.

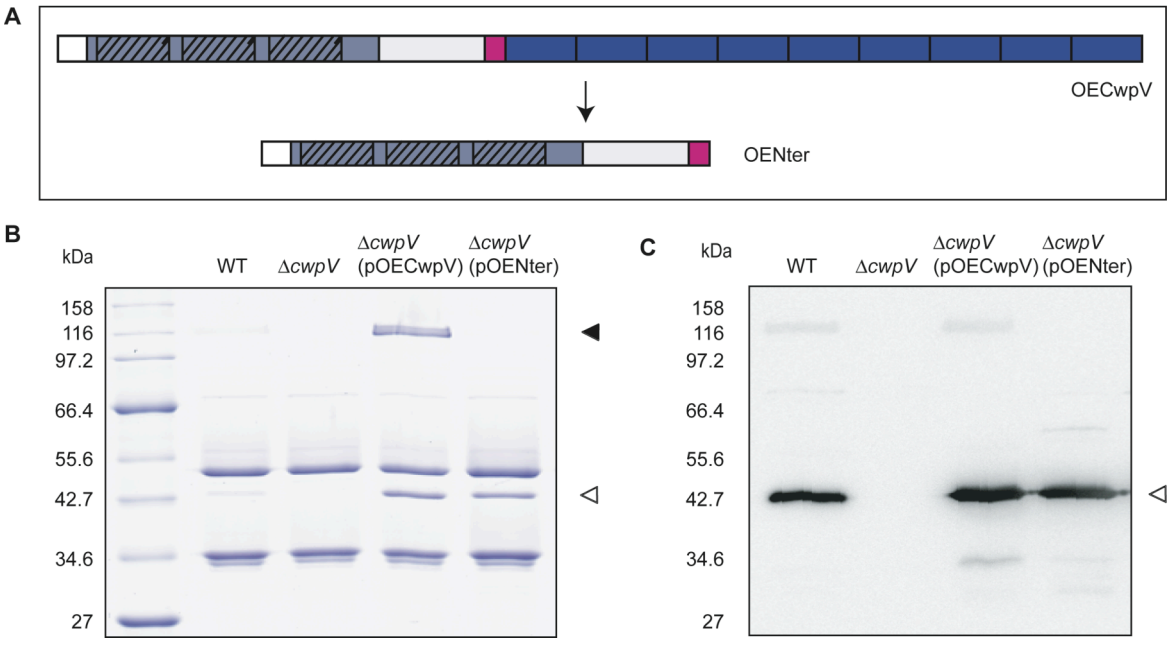


Figure 5.3 The CwpV N-terminal fragment mediates cell wall binding.

A. Illustration of the truncated version of CwpV expressed by Δ*cwpV*(pOECwpVNter). **B.** Expression of truncated CwpV in Δ*cwpV* visualized by SDS-PAGE and Coomassie staining, compared to WT, Δ*cwpV* and Δ*cwpV*(pOECwpV). Truncation of CwpV does not affect the expression of the CwpV N-terminal fragment in the *C. difficile* S-layer. **C.** Western blot analysis of the S-layer extracts from WT, Δ*cwpV,* Δ*cwpV*(pOECwpV) and Δ*cwpV*(pOECwpVNter) using anti-CwpVNter antibody.

5.2.4 The CwpV N-terminal (CWA) fragment and C-terminal fragment form a complex

Having determined the CwpV cleavage site and that the N-terminal fragment is sufficient for cell wall binding of CwpV it was postulated that the two CwpV fragments form a complex to allow anchoring of the C-terminal domain to the cell wall. In order to test this, an immunoprecipitation reaction using anti-CwpVrptI was carried out with a surface extract from *C. difficile* Δ*cwpV*(pOECwpV). This antibody only recognises the C-terminal fragment of CwpV, as shown in Figure 5.2C. However, as can be seen in

Figure 5.4 both the CWA and C-terminal fragments of CwpV are immunoprecipitated from the surface layer extract using this antibody.

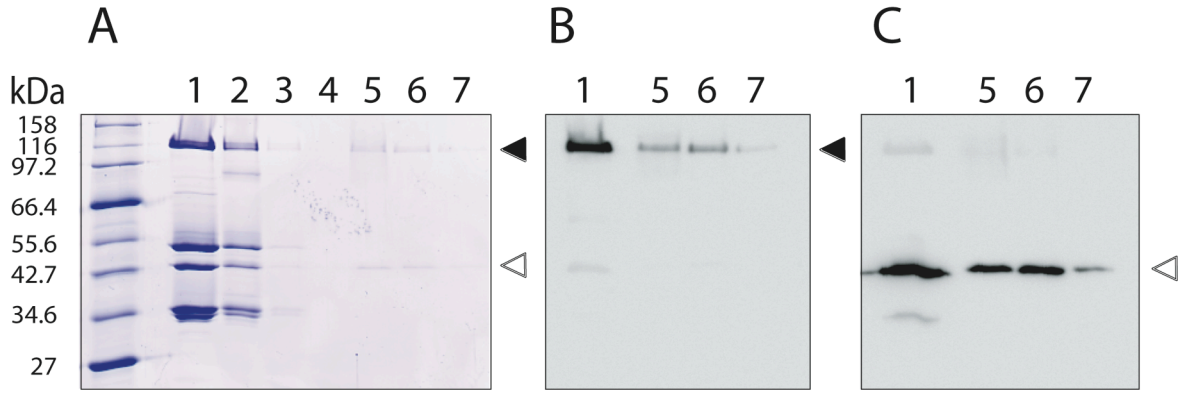


Figure 5.4 The two CwpV fragments form a complex.

Immunoprecpitation on a surface extract from Δ*cwpV*(pOECwpV) using anti-CwpVrptI antibody was carried out. **A.** SDS-PAGE gel of immunoprecipitation samples visualized by Coomassie staining. Lane 1: Δ*cwpV*(pOECwpV) surface extract; 2: antibody column flow through; 3: first column wash; 4: final column wash; 5-7: First three column elutions. **B.** Δ*cwpV*(pOECwpV) surface extract (1) and immunoprecipitation elutions (5-7) were analysed by anti-cwpVrptI Western blot. **C.** Δ*cwpV*(pOECwpV) surface extract (1) and immunoprecipitation elutions (5-7) were analysed by anti-cwpVNter Western blot. The CwpV C-terminal fragment is indicated by a black triangle, and the N-terminal fragment by an open triangle.

This suggests that the two fragments do form a complex. No other proteins were immunoprecipitated from the surface layer extract using this antibody, as can be seen from the Coomassie stained SDS-PAGE gel in

Figure 5.4A, and probing with anti-HMW SLP and anti-LMW SLP antibodies also did not detect immunoprecipitation of these proteins (data not shown). The two CwpV fragments are co-immunoprecipitated with a 1:1 ratio of fragments. This suggests that the two CwpV fragments form a stable 1:1 complex, with no interactions between CwpV and any other S-layer proteins detected.

5.2.5 Expression and purification of StrepII-tagged CwpV from *C. difficile* surface extracts

The *cwpV* gene from the plasmid pOECwpV was modified to encode a C-terminal StrepII-tag (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys), resulting in the plasmid pOEI, illustrated in Figure 5.5A. This plasmid was introduced to Δ*cwpV* by conjugation and the resulting strain Δ*cwpV*(pOEI) used to produce StrepII-tagged CwpV. Surface extracts were prepared from *ΔcwpV* and *ΔcwpV*(pOEI) and analysed by SDS-PAGE and Coomassie staining as shown in Figure 5.5B. These extracts were further analysed by Western blotting with a commercial anti-StrepII tag antibody, which detected the ~116 kDa C-terminal CwpV fragment (Figure 5.5C). Western blotting with anti-CwpVNter detected the CwpV CWA fragment at ~42 kDa (Figure 5.5D). Having validated the expression of StrepII-tagged CwpV in *C. difficile* surface extract from *ΔcwpV*(pOEI) was incubated with Streptactin resin to bind the StrepII-tagged C-terminal fragment of CwpV. The resin was washed and then the StrepII-tagged protein eluted by addition of desthiobiotin. The proteins yielded from a Streptactin binding/elution assay of a *ΔcwpV*(pOEI) glycine extract are shown in Figure 5.5E, visualized by SDS-PAGE and Coomassie staining. It

can been seen that the ~42 kDa and ~116 kDa fragments of CwpV are co-eluted (lanes 8-11), despite the presence of the strep-tag only in the larger fragment. This provides further evidence that the two CwpV fragments form a complex.

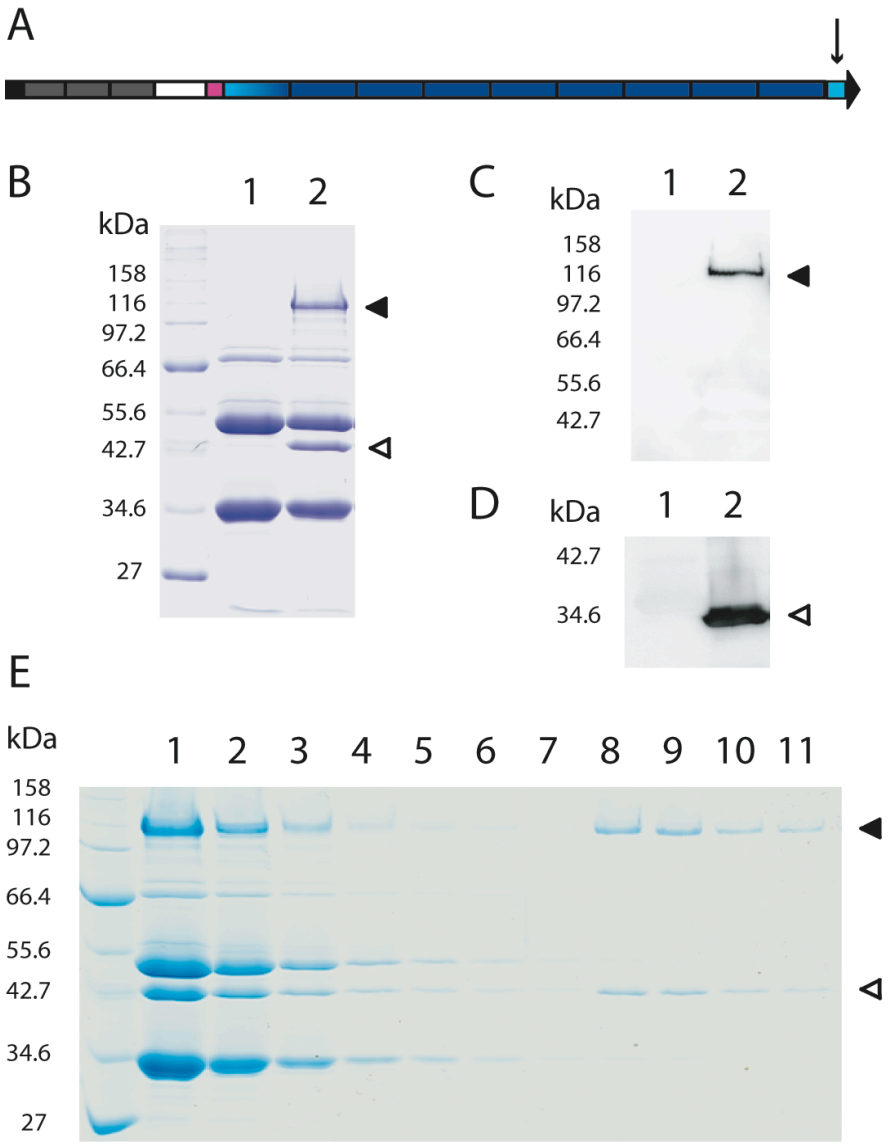


Figure 5.5 Expression and purification of StrepII-tagged CwpV from *C. difficile* surface extracts.

A. The *cwpV* gene in pOECwpV was modified to contain a C-terminal StrepII-tag. The StrepII-tag is illustrated here by a turquoise box at the C-terminus of CwpV highlighted with an arrow. The resulting plasmid was named pOEI. **B-D** Surface extracts from Δ*cwpV* (lane 1) and Δ*cwpV*(pOEI) (lane 2) were analysed by SDS-PAGE followed by Coomassie staining (**B**), anti-StrepII tag Western blot (**C**), and anti-CwpVNer Western blot (**D**). The CwpV C-terminal fragment is indicated with a black triangle, and the N-terminal fragment with an open triangle. **E.** A streptactin purification from a Δ*cwpV*(pOEI) surface extract was carried out. Samples were analysed by SDS-PAGE followed by Coomassue staining. Lane 1: Δ*cwpV*(pOEI) surface extract; 2: Streptactin column flow through; 3-7: Column washes; 8-11: Column elutions.

5.2.6 Mutagenesis of the CwpV cleavage region

To further investigate the CwpV cleavage region six deletion mutants were designed. Cleavage Domain Mutants (CDMs) were created by inverse PCR with pOEI as the template. I acknowledge that this was carried out by Marcin Dembek under my supervision. CDM1-5 each have a ten amino acid deletion extending upstream and downstream of the cleavage site and CDM6 has a 50 amino acid deletion encompassing the entire region as illustrated in Figure 5.6. The plasmids containing these deletions (pOEI CDM1-6) were then transferred to Δ*cwpV.*

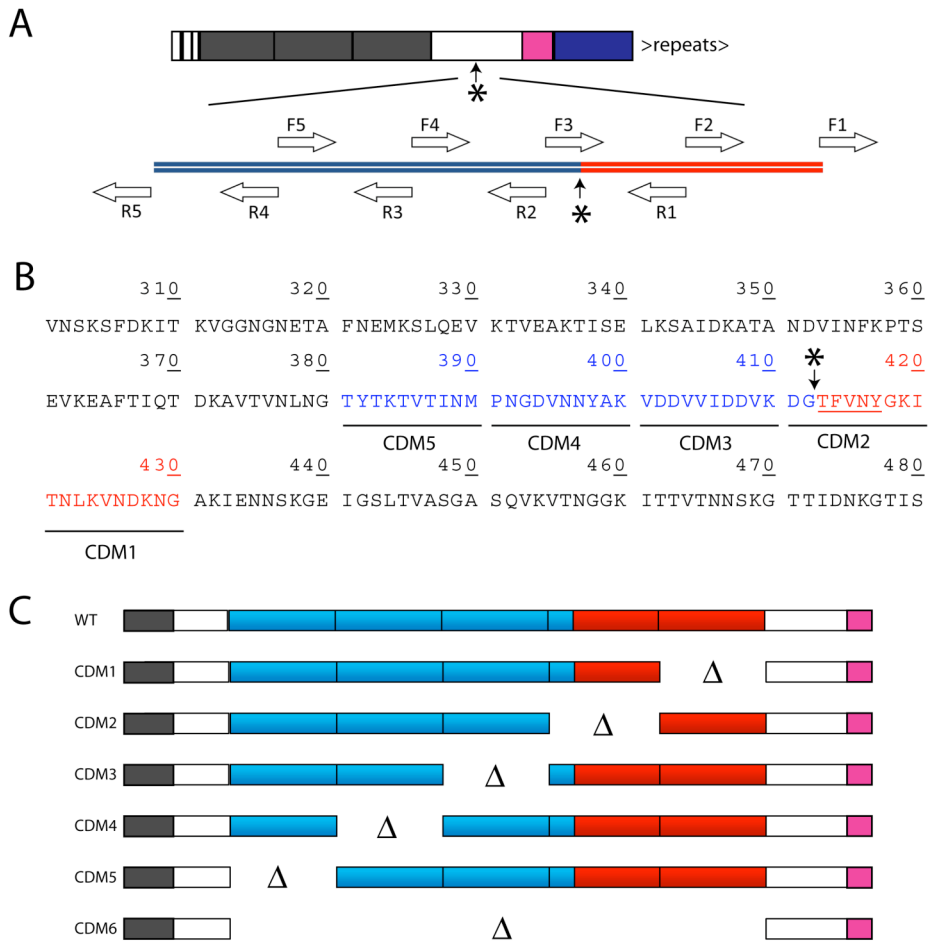


Figure 5.6 Production of CwpV Cleavage Domain Mutants.

A. The region surrounding the cleavage site (*) of CwpV was mutated by inverse PCR. The plasmid pOEI was used as a template and the primers F1-F5 and R1-R5 used to create mutants. F1+R1 were used to create CDM1, F2+R2 - CDM2. F3+R3 - CDM3, F4+R4 - CDM4, F5+R5 - CDM5 and finally F1+R5 - CDM6. **B.** The amino acids deleted in CDM1-5 are indicated. The cleavage site is designated by * at the TFNVY sequence determined by N-terminal sequencing. **C.** Cartoon representation of the six mutants made.

Strains Δ*cwpV*(pOEI CDM1-6) were grown overnight in BHI and analysed for CwpV expression in the surface extracts and supernatants. The results of this analysis are shown in Figure 5.7. In Figure 5.7A and B surface extracts are analysed by Coomassie staining and anti-CwpVrptI Western blotting, respecitively. It can be seen that the CwpV C-terminal fragment is not present in any of the mutant samples (Figure 5.7A/B lanes 6-11). This suggests that complex formation between the CWA and C-terminal fragments is abrogated by these mutations, and therefore the C-terminal fragment is not anchored to the cell surface. Large amounts of the C-terminal fragment can be seen in Figure 5.7C and D lanes 6-11, which contain supernatants of the *C. difficile* cultures (from which surface extracts were prepared for Figure 5.7A and B) analysed by Coomassie staining and anti-CwpVrptI Western blotting respectively. This further suggests that the CwpV C-terminal fragments from CDM1-6 are expressed but not anchored to the cell surface and therefore lost into the supernatant. The CWA fragment of CwpV is only seen in surface extracts from Δ*cwpV*(pOEI CDM1-3), suggesting that the CDM4-6 mutations affect the stability of the CwpV CWA fragment in the S-layer, and therefore neither CwpV fragment is detected in surface extracts from Δ*cwpV*(pOEI CDM4-6). The CwpV N-terminal fragment present in the surface extract from Δ*cwpV*(pOEI CDM1) (Figure 5.7A lane 6) appears larger than the normal CwpV N-terminal size (Figure 5.7A lanes 3-5). This suggests that cleavage is not occurring at the native site in this mutated CwpV protein. Similarly the N-terminal fragment from Δ*cwpV*(pOEI CDM3) (Figure 5.7A lane 8) appears smaller than the normal CwpV N-terminal size (Figure 5.7Alanes 3-5). This also suggests that the cleavage is not occurring at the native site. Taken together the data from these experiments suggest that a CwpV cleavage domain, consisting of at least 50 residues, is required for correct

cleavage and complex formation of CwpV, and that 10 amino acid deletion mutations are not tolerated at any point in this domain. This suggests that the intact domain is required for native processing and expression of CwpV.

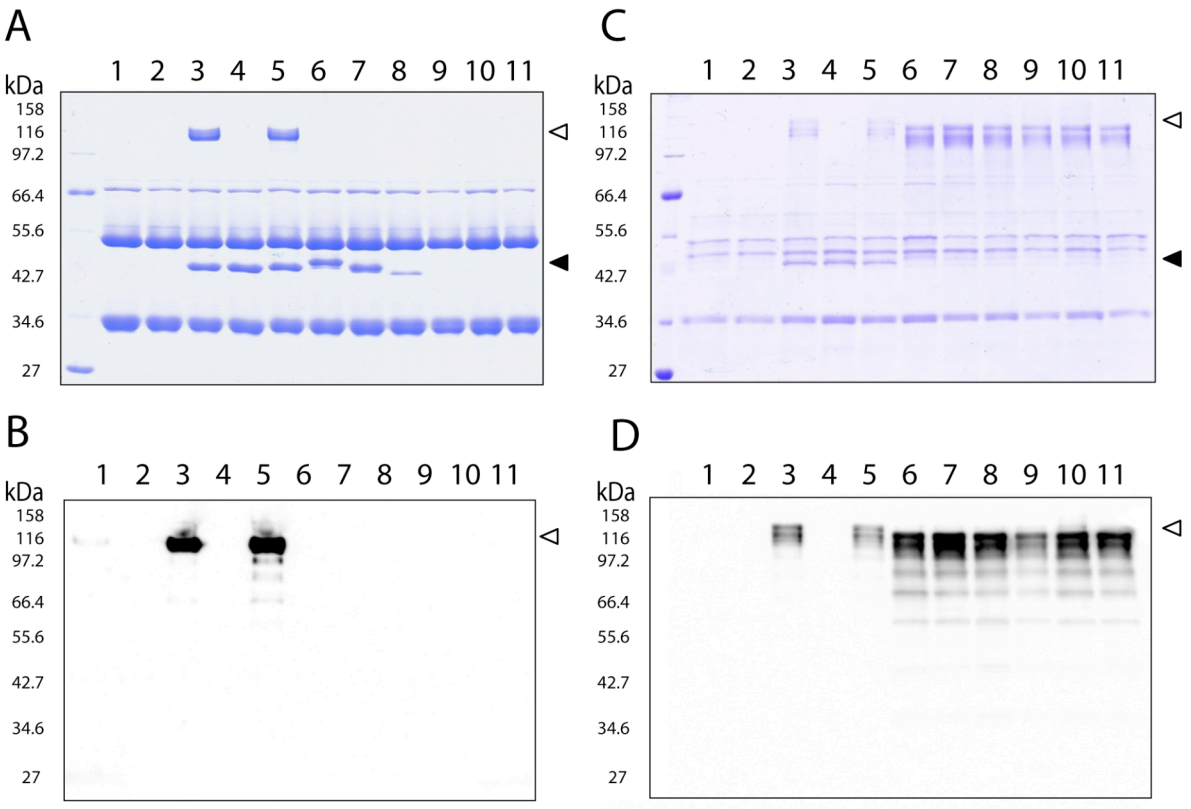


Figure 5.7 Analysis of CwpV Cleavage Domain Mutants.

Samples from *C. difficile* cultures 1 - WT, 2 - Δ*cwpV,* 3 - Δ*cwpV*(pOECwpV), 4 - Δ*cwpV*(pOECwpVNter), 5 - Δ*cwpV*(pOEI), 6-11 - Δ*cwpV*(pOEI CDM1-6). **A.** Surface extracts analysed by SDS-PAGE and Coomassie staining. **B.** Surface extracts analysed by anti-CwpVrptI Western blotting. **C.** Culture supernatants analysed by SDS-PAGE and Coomassie staining. **D.** Culture supernatants analysed by anti-CwpVrptI Western blotting. The CwpV C-terminal fragment is indicated by an open triangle. The N-terminal fragment is indicated by a black triangle.

None of these CwpV cleavage domain mutants led to the detection of full-length CwpV; even if the native cleavage site is deleted the CwpV protein still fragments. A panel of protease inhibitors, including inhibitors against all major classes of protease, was tested for activity in preventing CwpV cleavage, but none were found to have any effect (data not shown). Cwp84, the cysteine protease responsible for SlpA cleavage is not responsible for CwpV cleavage, as CwpV is

still cleaved in a *C. difficile* Δ*cwp84* background (Kirby et al., 2009). There is another protease member of the Cwp family, Cwp13. CwpV is still cleaved in a *C. difficile* Δ*cwp13* background, indicating that this is also not responsible for CwpV cleavage (Lucia de la Riva, personal communication). Therefore the mechanism responsible for CwpV cleavage remains uncharacterized, but it is clear that cleavage and complex formation between the CWA and C-terminal fragments of CwpV is necessary for stable expression of CwpV on the *C. difficile* surface.

5.2.7 Overall Model for CwpV processing and expression

Having characterised the CwpV cleavage site and complex formation between the two CwpV fragments, a model for CwpV processing and expression of the *C. difficile* cell surface could be presented. Given what is known about SlpA processing and expression on the *C. difficile* cell surface an analogy between the two proteins could be made. The proposed model is illustrated in

Figure 5.8. For both SlpA and CwpV the proteins are exported across the cell membrane due to the presence of a signal peptide at the N-terminus of the protein, which is removed by signal peptidase cleavage (step 1). Following export to the *C. difficile* cell surface both SlpA and CwpV are cleaved (step 2). In the case of SlpA, it is known that the protease Cwp84 is required for cleavage. For CwpV the mechanism of cleavage has not yet been characterised. Following cleavage the two fragments (for SlpA HMW and LMW SLP, for CwpV the N- and C-terminal fragments) associate to form a complex (step 3). This allows the fragment containing the cell wall binding domains (HMW SLP or CwpV N-terminal fragment) to anchor the other fragment to the cell surface. It is clear that the *C. difficile* S-layer, although composed primarily of SlpA, can accommodate a large amount of CwpV, as seen in Δ*recVON* cells (likely to reflect the natural expression

level of CwpV in WT *C. difficile* phase ON cells). This suggests that SlpA and CwpV must interact in such a way that maintains the integrity of the S-layer. Other members of the Cwp family are also likely to be involved in these S-layer interactions. Such interactions merit further investigation, and are likely to rely on the correct functioning of the three-step processing pathway of these proteins that has been outlined in this chapter.

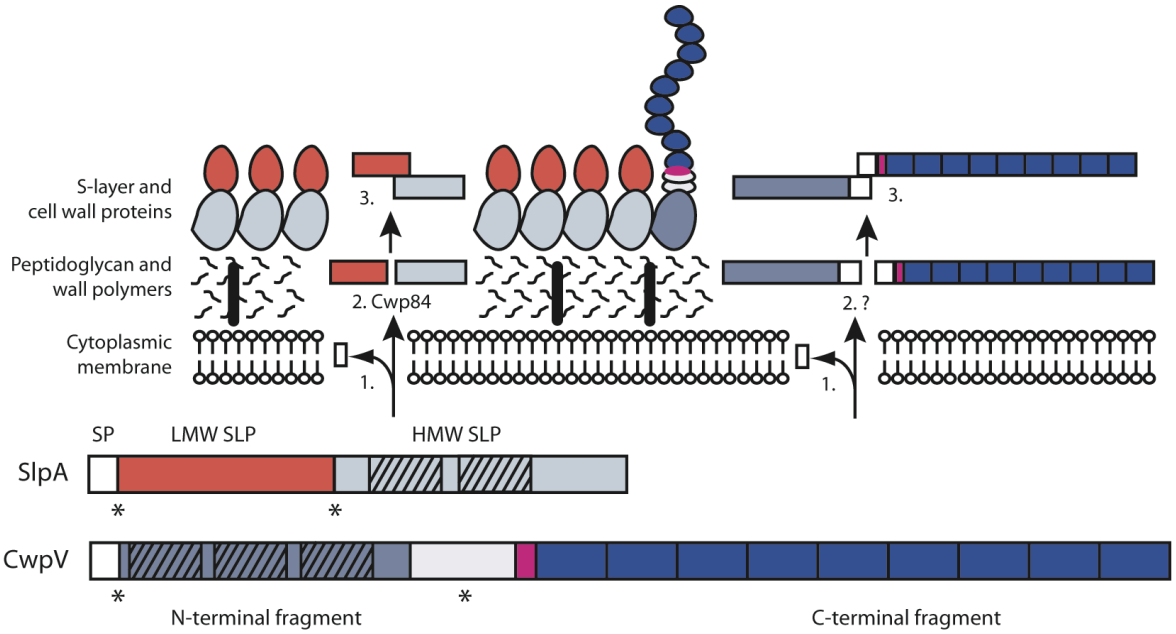


Figure 5.8 Overall Model for CwpV processing and expression.

A three-step model is proposed for CwpV processing, in a manner analogous to SlpA processing. SlpA is shown in red and grey. CwpV is shown in dark grey, white, pink and blue. Cleavage sites are indicated by *. Pfam04122 domains are indicated by hactching. Step 1: Export of the proteins across the cell membrane, accompanied by signal peptide cleavage. Step 2: Cleavage of the protein to yield two fragments. In the case of SlpA the protease Cwp84 is responsible, in the case of CwpV the cleavage mechanism is unknown. Step 3: Formation of a complex of the two fragments, which allows anchoring of both fragments to the cell surface via the cell wall binding domains of the HMW SLP or the CwpV N-terminal fragment.

5.2.8 CwpV C-terminal repeat diversity across strains

As mentioned in the introduction, prior to the start of this project a number of different CwpV C-terminal repeat types had been identified and their DNA

sequences determined by Dr. Jenny Emerson. However, the consequences of this sequence diversity had not been investigated. In order to investigate the five different known CwpV repeat types one representative *C. difficile* strain of each type was chosen for further study. The 630 CwpV repeats were termed type I, and this is the CwpV protein that has been characterised and discussed so far. A ribotype 027 strain isolated from an outbreak in Canada named R20352 encodes CwpV with the N-terminus up to the S/G region highly conserved compared to 630 CwpV, but containing eight 78-79 amino acid repeats at the C-terminus, termed type II repeats. The sequences of these repeats and their ClustalW2 alignment, accompanied by a cartoon illustration of the protein is shown in Figure 5.9A. As can be seen from this alignment the different repeats are virtually identical, with the first repeat exhibiting slightly more divergence from the consensus than the other repeats.

C. difficile strain CDKK167 was found to encode a CwpV protein with the N-terminal domain up to the S/G linker well-conserved compared to both 630 and R20352, with six copies of third type of repeat, termed type III, at the C-terminus. A cartoon representation of this protein is shown in Figure 5.9B accompanied by an alignment of the six ~94 amino acid type III repeats. All six repeats align well, however the first two repeats are slightly more divergent from the last four repeats than they are from each other. This is reflected in the shading of the repeats.

C. difficile strains M9 and AY1 encode the conserved CwpV N-terminus followed by mosaics of type III then type IV and V repeats respectively. This is illustrated in the diagram in Figure 5.9C, accompanied by an alignment of all the type III repeats from all three strains (CDKK167, M9, AY1). All the type III repeats show a

high degree of conservation, with the repeats from M9 and AY1 aligning best with the first two repeats from CDKK167. This is reflected in the colouring of the diagrams. The seven 98-100 amino acid type IV repeats from M9 are very well conserved, with very few divergent residues between repeats, as shown in the alignment in Figure 5.9D. Similarly, the five type V repeats from AY1 are all very similar in sequence, as the alignment in Figure 5.9E illustrates.

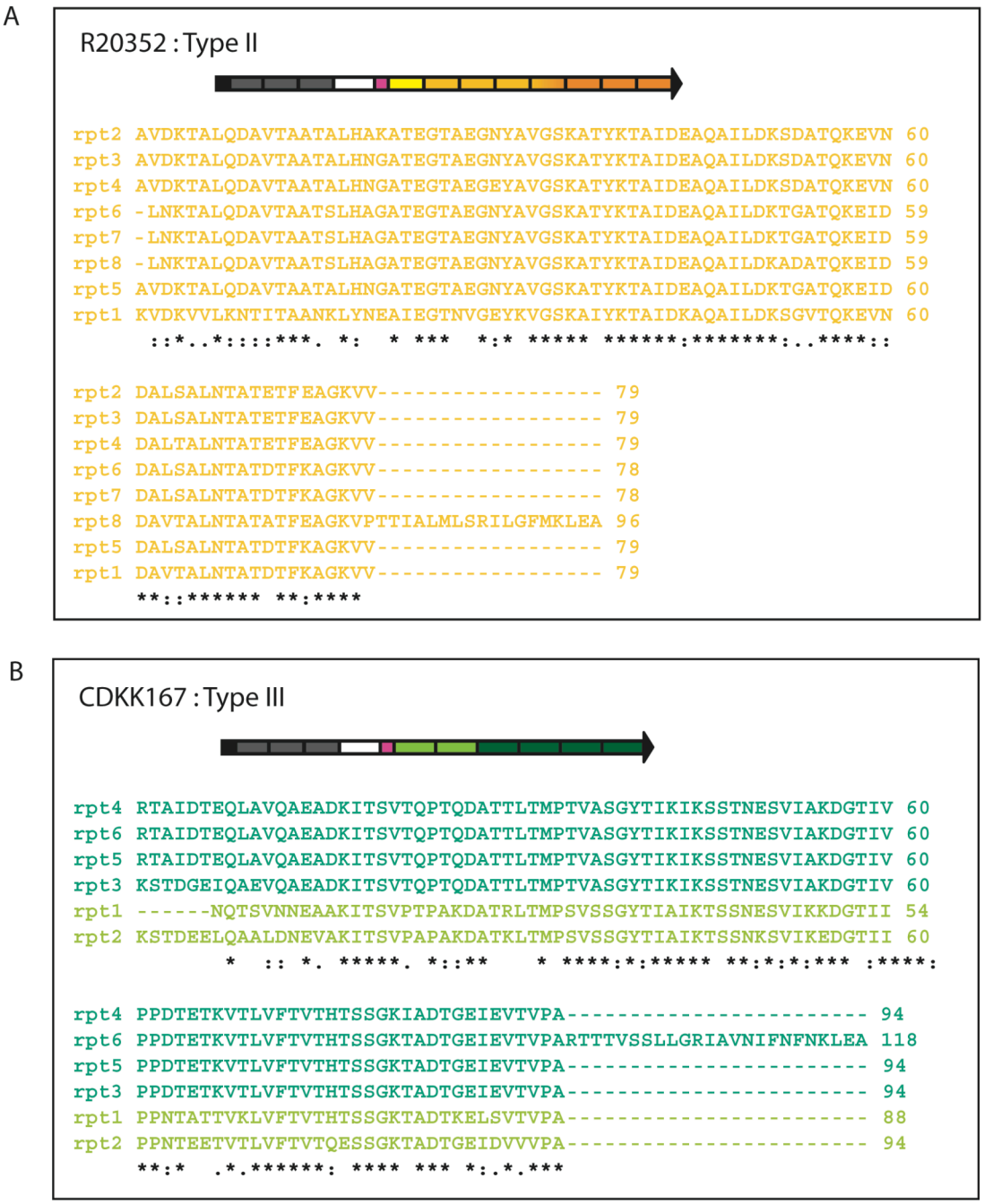


Figure 5.9 CwpV C-terminal repeat types II-V.

A. *C. difficile* strain R20352 encodes CwpV with type II repeats. Cartoon representation and ClustalW2 mutliple sequence alignment of repeats is shown. **B.** *C. difficile* strain CDKK167 encodes CwpV with type III repeats. Cartoon representation and ClustalW2 mutliple sequence alignment of repeats is shown. **C.** *C. difficile* strains M9 and AY1 encode CwpV with mosaics of type III and Type IV or V repeats. Cartoon representations of these mosaic CwpV proteins and ClustalW2 mutliple sequence alignment of all type III repeats is shown. **D.** ClustalW2 mulitple sequence alignment of M9 type IV repeats. **E.** ClustalW2 multiple sequence alignment of AY1 type V repeats is shown.

C




```
167rpt4   RTAIDTEQLAVQAEADKITSVTQPTQDATTLTMPTVASGYTIKIKSSTNESVIAKDGTIV 60
167rpt6   RTAIDTEQLAVQAEADKITSVTQPTQDATTLTMPTVASGYTIKIKSSTNESVIAKDGTIV 60
167rpt5   RTAIDTEQLAVQAEADKITSVTQPTQDATTLTMPTVASGYTIKIKSSTNESVIAKDGTIV 60
167rpt3   KSTDGEIQAEVQAEADKITSVTQPTQDATTLTMPTVASGYTIKIKSSTNESVIAKDGTIV 60
167rpt1   ------NQTSVNNEAAKITSVPTPAKDATRLTMPSVSSGYTIAIKTSSNESVIKKDGTII 54
AY1rpt1   ------NQTSVNNEAAKITSVPTPAKDATRLTMPSVSSGYTIAIKTSSNESVIKKDGTII 54
M9rpt1    ------NQTSVNNEAAKITSVPTPAKDATKLTMPSVSSGYSIAIKTSSNESVIKKDGTII 54
167rpt2   KSTDEELQAALDNEVAKITSVPAPAKDATKLTMPSVSSGYTIAIKTSSNKSVIKEDGTII 60
AY1rpt2   KSTDEELQAALDNEVAKITSVPAPAKDATKLTMPSVSSGYTIAIKTSSNKSVIKEDGTII 60
                 *  :: *. *****. *::*** ****:*:***:* **:*:*:*** :*****:

167rpt4   PPDTETKVTLVFTVTHTSSGKIADTGEIEVTVPA------------------------ 94
167rpt6   PPDTETKVTLVFTVTHTSSGKTADTGEIEVTVPARTTTVSSLLGRIAVNIFNFNKLEA 118
167rpt5   PPDTETKVTLVFTVTHTSSGKTADTGEIEVTVPA------------------------ 94
167rpt3   PPDTETKVTLVFTVTHTSSGKTADTGEIEVTVPA------------------------ 94
167rpt1   PPNTATTVKLVFTVTHTSSGKTADTKELSVTVPA------------------------ 88
AY1rpt1   PPNTATTVKLVFTVTHTSSGKTADTKELSVTVPA------------------------ 88
M9rpt     PPNTATTVKLVFTVTHTSSGKTADTKELSVTVPA------------------------ 88
167rpt2   PPNTEETVTLVFTVTQESSGKTADTGEIDVVVPA------------------------ 94
AY1rpt2   PPNTEETVTLVFTVTQESSGKTADTGEIDVVVPAKSTK-------------------- 98
          **:*  .*.******: **** *** *:.*.***
```

D

```
M9typeIVrpt6  --KDVVAAVSTVNATNGTLTIVLDKTPTVDPVEGDFTAKKSIDGGQESELTLSNFAYNKE 58
M9typeIVrpt7  --KDVVAAVSTVNATNGTLTIVLDKTPTVDPVEGDFTAKKSIDGGQESELTLSNFAYNKE 58
M9typeIVrpt5  --KDVVAAVSTVNATNGTLTIVLDKTPTVDPVEGDFTAKKSIDGGQESELTLSNFAYNKE 58
M9typeIVrpt4  KGQDVVAAVSTVNATNGTLTIVLDKTPTVDPVEGDFTAKKSIDGGQESELTLSNFAYNKE 60
M9typeIVrpt2  KGQDVVAAVSTVNATNGTLTIVLDKTPTVDPVEGDFTAKKSIDGGQESELTLSNFAYNKE 60
M9typeIVrpt3  KGQDVVAAVATVNATNGTLTIVLDKTPTVDPVEGDFTAKKSIDGGQESELTLSNFVYNKE 60
M9typeIVrpt1  KSTEVQAAVSTVNATNGTLTIVLDKTPTVDPVEGDFTAKKSIDGGQESELTLSNFAYNKE 60
                 :* ***:*************************************************.****

M9typeIVrpt6  SKTVTYSFVPIEQTELEQSVVVGVDYKGENTKAAAFKVDA----------------------- 98
M9typeIVrpt7  SKTVTYSFVPIEQTELEQSVVIGVDYKGENTKAAAFIVNGNSVPTRTLKTNLTNVAKSVINLEA 122
M9typeIVrpt5  SKTVTYSFVPIEQTELEQSVVVGVDYKGENTKAAAFKVDA----------------------- 98
M9typeIVrpt4  SKTVTYSFVPIEQTELEQSVVVGVDYKGENTKAAAFKVDA----------------------- 100
M9typeIVrpt2  SKTVTYSFAPIEQTELEQSVVVGVDYKGENTKAAAFKVEA----------------------- 100
M9typeIVrpt3  SKTVTYSFTPIEQTELEQSVVVGVDYKGENTKAAAFKVEA----------------------- 100
M9typeIVrpt1  SKTVTYSFTPIEQTELEQSVVVGVDYKGENTKAAAFKVEA----------------------- 100
              ********.************:*************** *:.
```

E

```
AY1typeVrpt3  VIIGEGSVSGATAGDKEITGLTSGKIYKVTMDGVEKYTKADGTLGEEGEKAALTGTSITG 60
AY1typeVrpt4  VIIGEGSVSGATAGDKEITGLTSGKIYKVTMDGVEKYTKADGTLGEEGEKAALTGTSITG 60
AY1typeVrpt1  VIIGEGSVSGATAGDKEITGLTSGKIYKVTMDGVEKYTKADGTLGEEGEKAALTGTSITG 60
AY1typeVrpt2  VIIGEGSVSGATAGDKEITGLTSGKIYKVTMDGVEKYTRADGTLGEEGEKAALTGTSITG 60
AY1typeVrpt5  VIIGEGSVSGATAGDKEITGLTSGKIYKVTMDGVEKYTKADGTLGEEGEKAALTGTSITG 60
              ************************************:**********************

AY1typeVrpt3  LVNGKTYKVEEYVAPVVNE--------------- 79
AY1typeVrpt4  LVNGKTYKVEEYVAPVVNE--------------- 79
AY1typeVrpt1  LVNGKTYKVEEYVAPVVNE--------------- 79
AY1typeVrpt2  LVNGKTYKVEEYVAPVVNE--------------- 79
AY1typeVrpt5  LVNGKTYKVEEDLSAISTIGRLMLKFTNLVKLEA 94
              *********** ::.: .
```

Given the similarity of amino acid sequence of repeats from within a specific type it is likely that each repeat is antigenically very similar to another repeat of its type. However, a key question is whether or not there is any sequence similarity between different CwpV repeat types. One repeat of each type was selected based on having the amino acid sequence closest to the repeat-type consensus sequence, indicated in Figure 5.10A. These five different CwpV repeat amino acid sequences were aligned in a ClustalW2 multiple sequence alignment and it can be seen in Figure 5.10B that there is no significant alignment to a consensus sequence of amino acids between different CwpV repeat types.

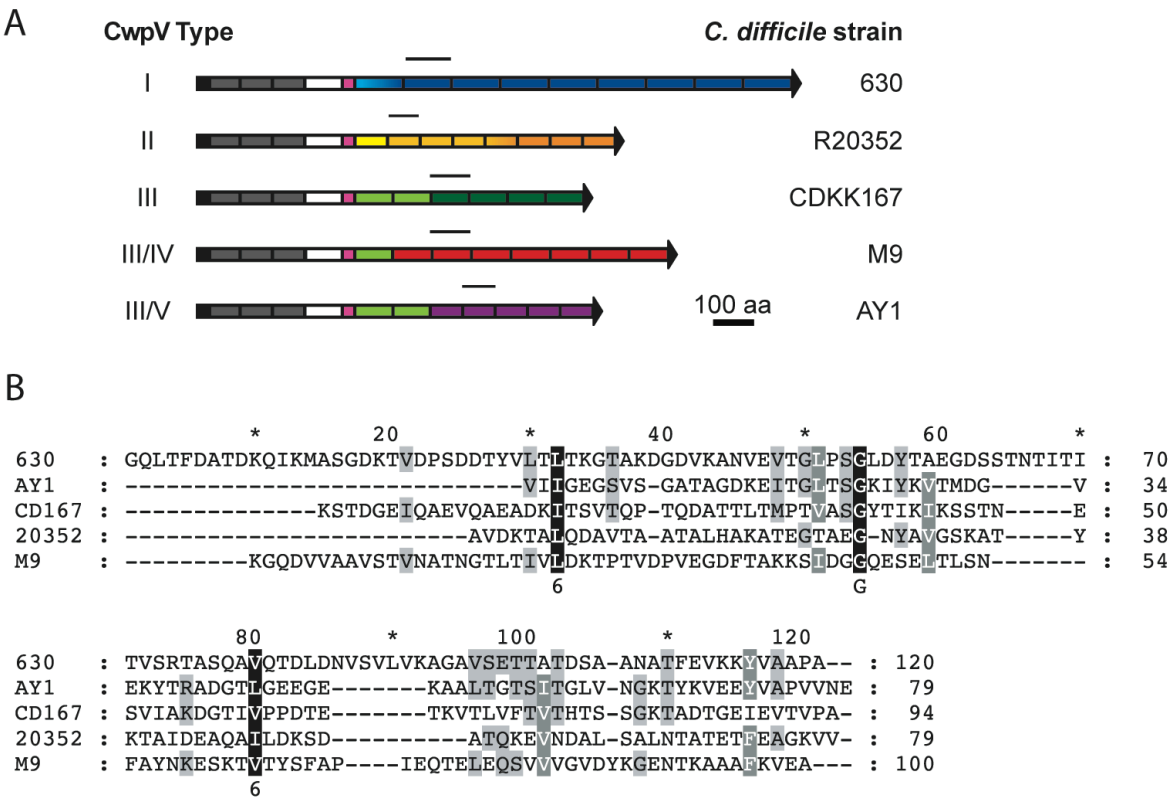


Figure 5.10 Different CwpV types have distinct amino acid sequences

A. Cartoon representation of the CwpV proteins from each type-representative strain chosen for further study. One repeat of each type, indicated by a black line above the cartoon, was selected and amino acid sequences from each type aligned by ClustalW2 sequence alignment. **B.** The ClustalW2 multiple sequence alignment of all CwpV repeat types indicates that there is no significant amino acid conservation between CwpV repeat types.

5.2.9 Different CwpV repeat types are antigenically distinct

Given the completely different amino acid sequences for each repeat type we hypothesized that the repeats may be antigenically distinct. Investigation into this hypothesis was desired in order to understand the significance of CwpV repeat type diversity. In order to test this hypothesis antibodies against each repeat type were required. Antibodies against the 630 CwpV N-terminus (anti-CwpVNter) and the first 630 C-terminal repeat (anti-CwpVrpt1) had been raised prior to the start of this project and have been used extensively in experiments reported so far. Rabbit antibodies against repeat types II-V were obtained commercially by immunization with one or two recombinant His-tagged CwpV repeat domains of a specific type purified from *E. coli*. The recombinant CwpV domains used as antigens are indicated by black lines above the CwpV protein cartoons in Figure 5.11A. An example of the affinity purification carried out for each of the recombinant proteins is shown in Figure 5.11B, where CwpV type V repeat purification is shown. CwpV repeat domains of types II-IV were purified in the same way (data not shown).

I acknowledge the technical assistance of Allie Shaw in carrying out this work.

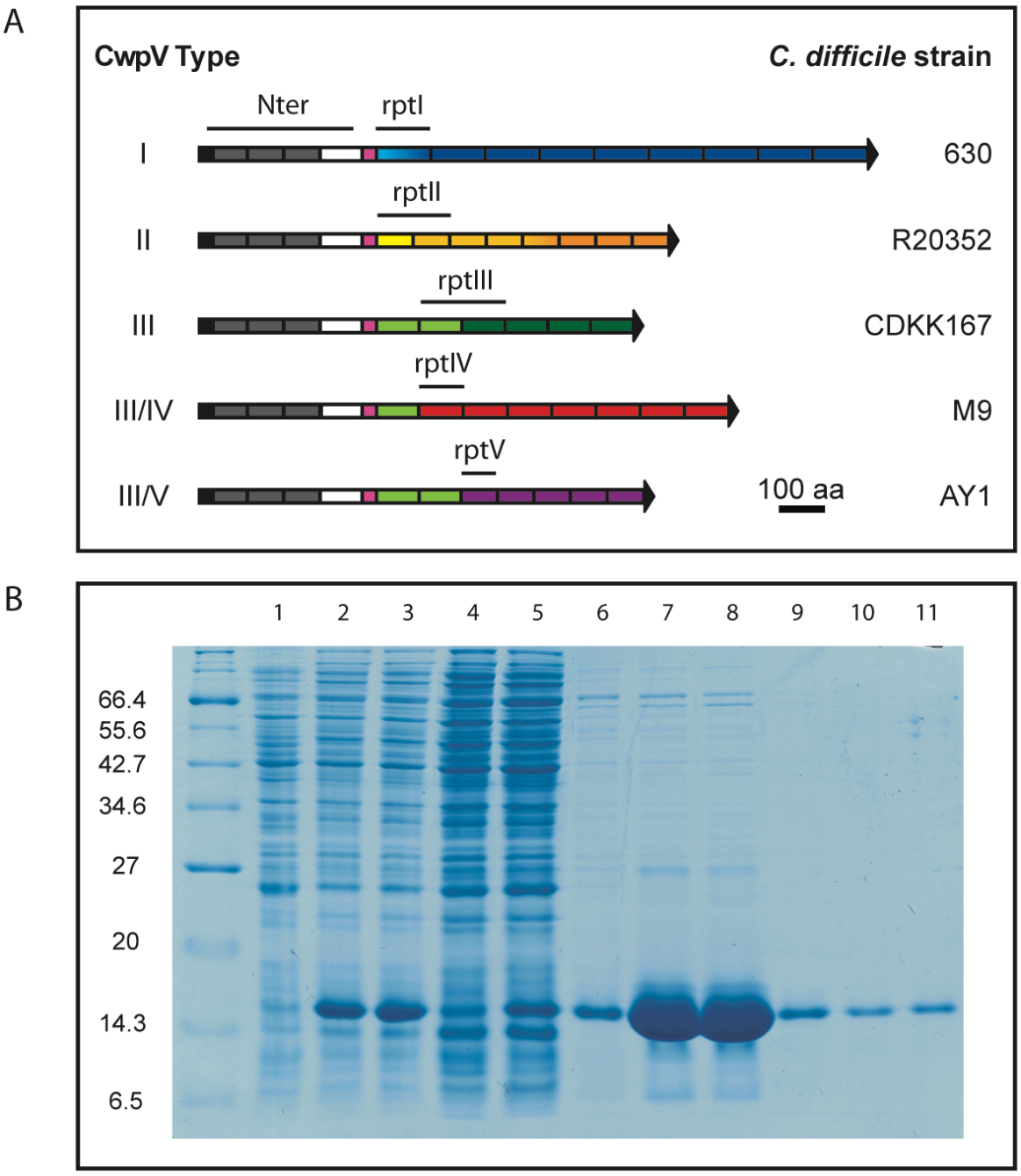


Figure 5.11 Raising CwpV repeat-specific antibodies.

A. Diagrammatic representation of the CwpV domains used to raise anti-CwpV antibodies. The domains used as antigens are indicated by black lines above the cartoons. Anti-CwpVNter and anti-CwpVrpt1 antibodies were raised prior to this study, but are included in this diagram for completeness. The type II-V repeat domains indicated were cloned into pET28a to encode a C-terminal his-tag. Proteins were then expressed in Rosetta *E. coli* and purified by Nickel affinity chromatography. **B.** Analysis of CwpV type V repeat recombinant expression and purification by SDS-PAGE and Coomassie staining. Lanes contain the following samples: 1; Whole cell lysate of uninduced Rosetta *E. coli*, 2; Whole cell lysate of induced cells, 3; Soluble fraction of induced cells, 4; Nickel column flow through, 5; First column wash, 6; Final column wash, 7-11 Column elution fractions. This analysis is shown as an example, the same process was carried out for CwpV repeat types II-IV.

In order to test the resultant antibodies for reactivity against CwpV proteins of all five types, S-layer extracts were prepared for the representative strain of each type; 630 – type I, R20352 – type II, CDKK167 – type III, M9 – type III/IV, and AY1 – type III/V. S-layer extracts were normalized to contain the same level of HMW SLP, and were visualized by SDS-PAGE and Coomassie staining shown in Figure 5.12A. The significant variation in SlpA across these strains can be seen, as has been described across *C. difficile* strains previously (Calabi & Fairweather, 2002a). The S-layer extracts were then analysed by Western blotting with anti-CwpVNter antibody. A single band at ~42 KDa was seen in each strain, shown in Figure 5.12B, confirming that antibodies raised against the 630 CwpV N-terminal domain recognized this domain in all the other strains, as would be expected by the high level of conservation of the CwpV N-terminus across strains. This also confirmed that CwpV is present in all these *C. difficile* S-layer extracts, and therefore CwpV is expressed by all these strains.

S-layer extracts were then analysed by Western blotting using antibodies raised against the specific repeat types I-V. It is clear from Figure 5.12C that anti-CwpVrptI detects only type I CwpV repeats in strain 630. Anti-CwpVrptII detects only type II CwpV repeats in strain R20352. Anti-CwpVrptIII detects CwpV repeats in CDKK167, M9 and AY1, and all these three strains contain at least one type III repeat. Detection of M9 CwpV with anti-CwpVrptIII is weak, this may be due to the presence of only one type III repeat in this CwpV protein. Anti-CwpVrptIV only detects CwpV in M9, the only strain with type IV repeats. Finally, anti-CwpVrptV only detects CwpV from AY1. These results demonstrate that, as expected based on distinct amino acid sequences, the CwpV repeat types are antigenically distinct.

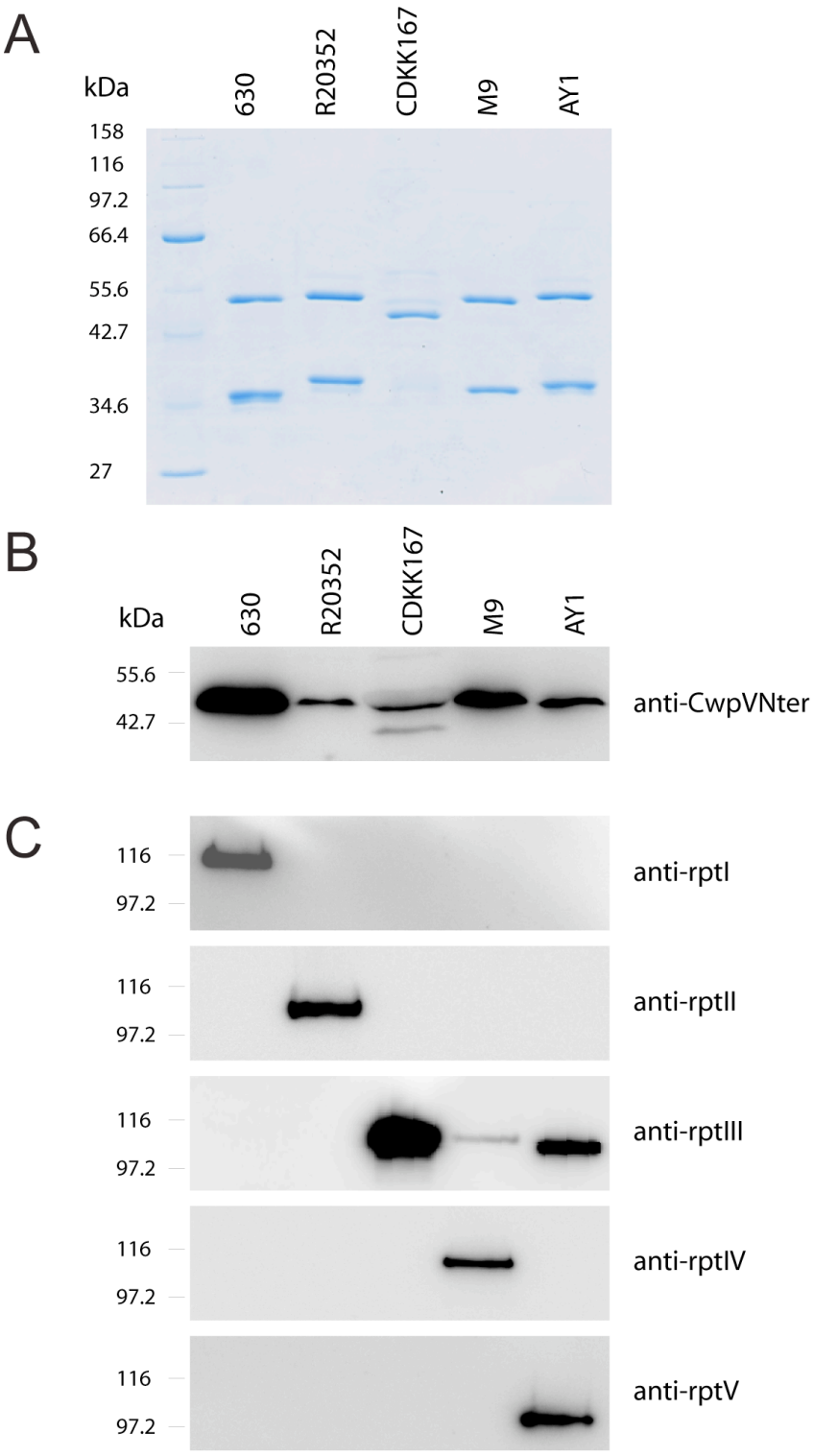


Figure 5.12 Different CwpV repeat types are antigenically distinct.

A. S-layer extracts from *C. difficile* strains analysed by SDS-PAGE and Coomassie staining illustrates the variability of S-layer proteins across strains, as previously described (Calabi & Fairweather, 2002a). **B.** Western blot analyis of S-layer extracts using anti-CwpVNter shows that the CwpV N-terminus from all strains is recognized, as expected due to highly conserved N-terminal sequences. **C.** Western blot analysis of S-layer extracts using anti-CwpVrptI-V shows that each repeat-specific antibody only detects CwpV proteins containing at least one repeat of its own repeat type.

5.2.10 The post-translational cleavage site is conserved for all CwpV proteins

In order to further characterize the four distinct CwpV proteins expressed by R20352, CDKK167, M9 and AY1 the *cwpV* genes from these strains were amplified by PCR and cloned into a *C. difficile* expression vector encoding a C-terminal StrepII-tag. The resulting plasmids pOEII-V were introduced by conjugation to *ΔcwpV.* Transconjugants were single colony purified and the resultant strains designated *ΔcwpV*(pOEII-V), for strains over-expressing CwpV types II-V respectively. *C. difficile* surface extracts were prepared from *ΔcwpV* and *ΔcwpV*(pOEI-V) and analysed as shown in Figure 4.13A. A high level of CwpV expression can be seen for each CwpV protein, with CwpV constituting a significant proportion of the total S-layer extract. All 5 proteins are cleaved into a ~42 kDa N-terminal fragment and C-terminal fragments of ~90-120 kDa. These extracts were further analysed by Western blotting with a commercial anti-StrepII tag antibody, which detected the ~90-120 kDa C-terminal CwpV fragments (Figure 4.13B). Western blotting with anti-CwpVNter detected the CwpV N-terminal fragment in each extract at ~42 kDa (Figure 4.13C). In order to determine the cleavage site for each CwpV protein, each StrepII-tagged C-terminal fragment was transferred to PVDF membrane and subjected to N-terminal sequencing by Edman degradation. This identified the first 5 amino acids of all the C-terminal CwpV fragments as TFVNY, unambiguously locating the cleavage site in the domain of CwpV between the CWA domains and the serine/glycine rich region, as depicted in Figure 4.13D. This domain is well-conserved between CwpV proteins of different types, and suggests that the CwpV post-translational cleavage mechanism is conserved across CwpV types.

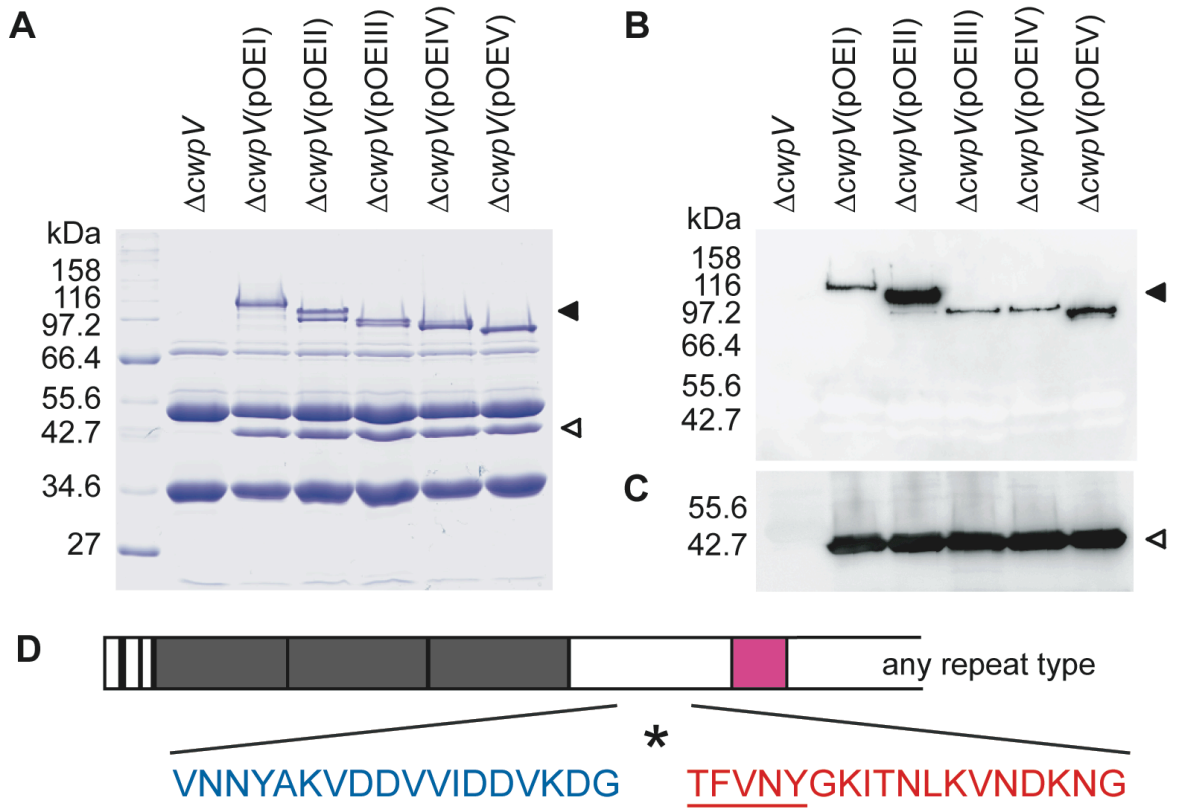


Figure 5.13. There is a conserved post-translational cleavage site for all CwpV proteins.

A. S-layer extracts from *ΔcwpV* and *ΔcwpV* overexpressing CwpV types I-V with a C-terminal strep-tag *(ΔcwpV*(pOEI-V)) analysed by SDS-PAGE gel and Coomassie staining. Open triangle, CwpV N-terminus. Black triangle, CwpV C-terminus **B.** Western blot analysis of S-layer extracts using anti-Strep tag antibody detects the strep-tagged CwpV C-termini. **C.** Western blot analyis of S-layer extracts using anti-CwpVNter antibody detects the CwpV N-termini. **D.** N-terminal sequencing of all StrepII-tagged CwpV C-termini yielded the sequence TFVNYG, revealing the conserved cleavage site for all CwpV types located between the cell wall anchoring domains (grey) and the serine/glycine-rich region (pink).

5.2.11 All CwpV types form a CWA/C-terminal fragment complex

Considering that the cleavage site of CwpV was conserved across CwpV proteins of different types, and that the cleavage domain thought to be involved in CwpV complex formation is also conserved, it seemed likely that CwpV proteins of all repeat types form complexes of CWA and C-terminal fragments. To test this hypothesis glycine extracts from *ΔcwpV*(pOEI-V) strains were incubated with Streptactin resin to bind the StrepII-tagged C-terminal fragment of CwpV. The resin was washed and then the StrepII-tagged protein eluted. This Streptactin binding/elution assay was described in detail earlier (see Figure 5.5). The complete S-layer extracts (S) and first elutions (E) from each assay are visualized in Figure 5.14. Both CwpV fragments co-eluted from S-layer extracts containing CwpV of all repeat types. This suggests that all types of CwpV form a complex of the two CwpV cleavage products.

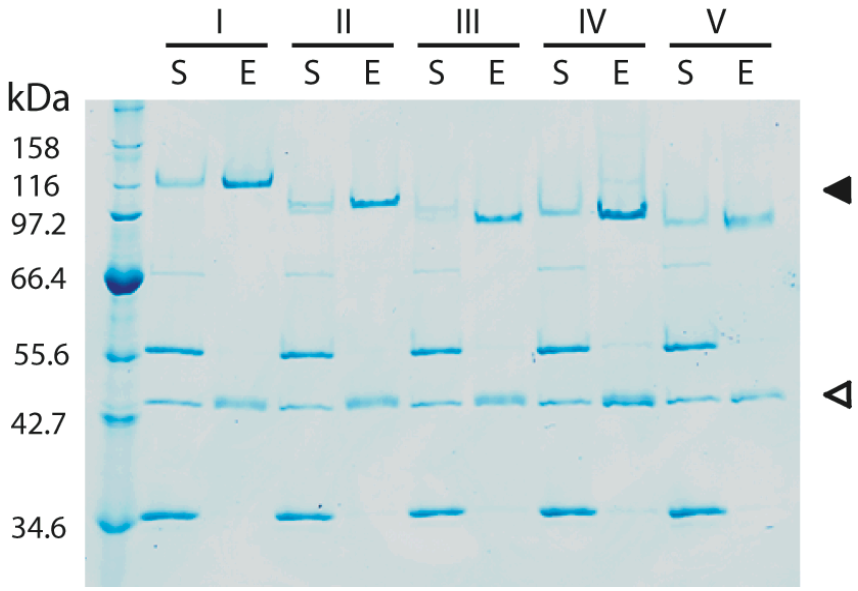


Figure 5.14 All CwpV proteins form a complex of N- and C-terminal fragments.

S-layer extracts (S) and first elutions (E) from ΔcwpV(pOEI-V) Streptactin binding/elution assays. In each E lane the CwpV N-terminal fragments (open triangle) and CwpV C-terminal fragments (closed triangle) can be seen, indicative of complex formation by the two fragments.

5.2.12 Phase variable regulation of CwpV expression by *C. difficile* strains of all CwpV repeat type

Given the known conservation of *recV* and the *cwpV* switch across *C. difficile* strains it seemed possible that CwpV would be expressed in a phase variable manner in strains expressing all types of CwpV. Overnight *C. difficile* cultures of representative strains for each type were analysed by *cwpV* switch orientation PCR. It can be seen in Figure 5.15A that all strains exhibited *cwpV* switch inversion. The same cultures were stained using the appropriate anti-CwpVrpt antibody for analysis by immunofluorescent microscopy. As can be seen in Figure 5.15B all strains were shown to express CwpV in a subset of cells, indicative of phase variable expression. In all strains CwpV was expressed in a minority of cells, ranging from 0.1 - 10% of cells being CwpV-positive. The factors determining the proportion of CwpV-positive cells are yet to be determined, but it is clear that under standard laboratory culture conditions it is a minority of cells that are CwpV-positive in all strains tested.

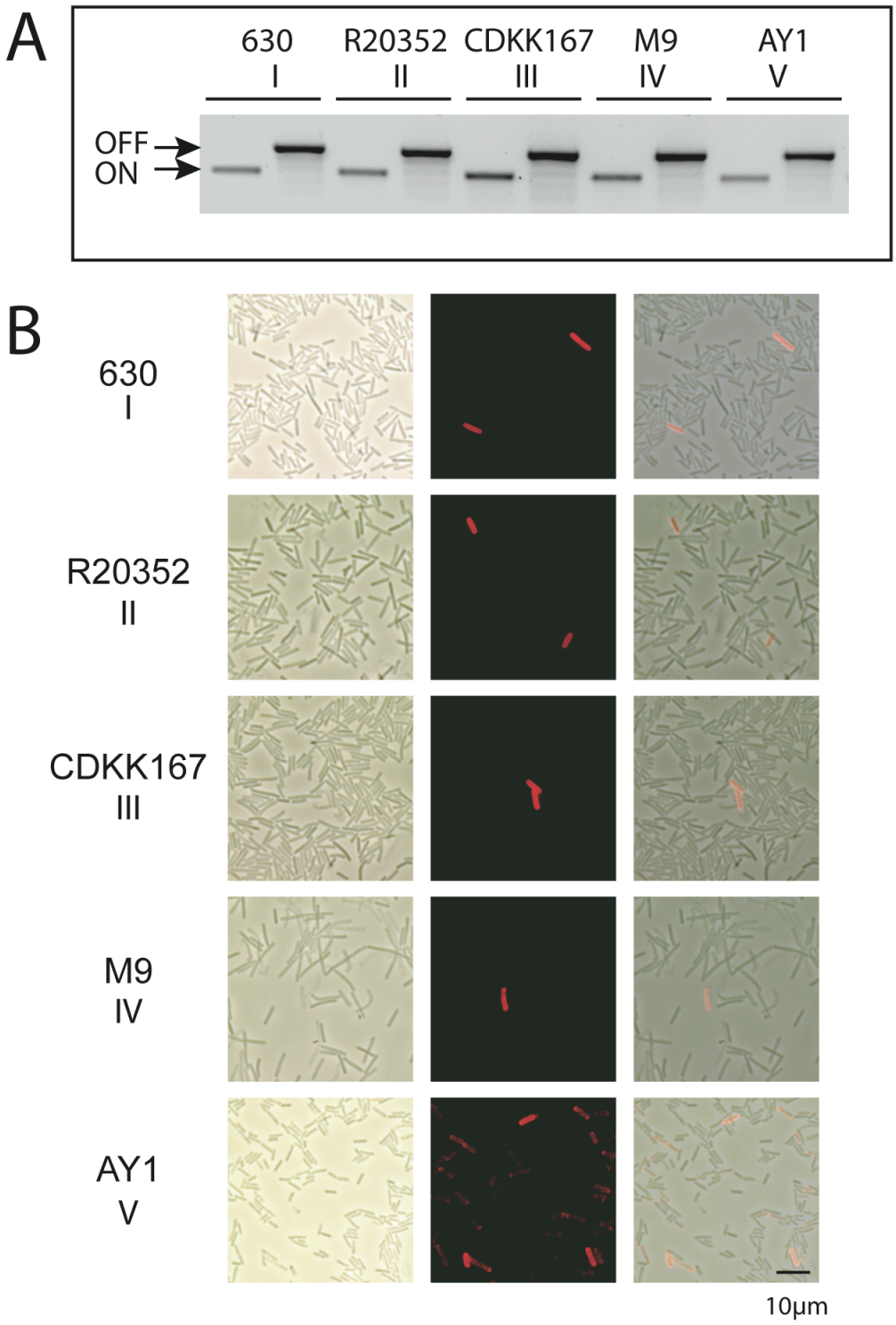


Figure 5.15 Phase variable regulation of CwpV expression by *C. difficile* strains of all CwpV repeat type.

A. Orientation PCRs amplified from overnight cultures of *C. difficile* strains 630, R20352, CDKK167, M9 and AY1 indicating that all strains exhibit *cwpV* switch inversion. **B.** Overnight *C. difficile* cultures for each strain were stained with the appropriate anti-CwpVrpt antibody (630 – rptI, R20352 – rptII, CDKK167 – rptIII, M9 – rptIV, AY1 – rptV), then with anti-rabbit rhodamine red. From left to right – Phase image, fluorescent image, overlay.

5.3 Discussion

The high degree of sequence conservation within repeats of a specific type, exhibited for all repeat types, suggests that the repetitive repeat domain structure may have arisen by genetic duplication events. Without knowing the function of CwpV it is unclear why a repetitive repeat domain structure would be maintained, however one obvious explanation is that if these repeats are involved in a ligand binding event then by increasing the number of repeats present in each CwpV protein the avidity of the interaction would be increased. This is discussed further in Chapter 6, which focuses on the function of CwpV

C-terminal variation of CwpV across different *C. difficile* strains, with antigenically distinct types encoded by unrelated sequences, suggests that these sequences have been acquired by horizontal gene transfer, rather than having diverged from a common genetic ancestor. Any known mechanism of chromosomal sequence acquisition; homologous recombination, illegitimate recombination or additive integration could account for introduction of these sequences to the *C. difficile.* chromosome (Thomas & Nielsen, 2005). Without knowing the origin of these sequences, the mechanism of transfer into *C. difficile* cells and incorporation into the *C. difficile* genome is not clear. Searching of current nucleotide databases does not provide any clear evidence as to the origin of these sequences. Much of the human microbiome remains uncharaterised, but is currently targeted for high-throughput sequencing (Peterson et al., 2009), which may shed light on the origins of CwpV sequences. However, if acquisition of these sequences by *C. difficile* occurred a long time ago, sequence divergence may hide links between CwpV sequences and the sources. The presence of mosaic gene structure in M9

and AY1 strains, where they have type III repeats combined with type IV and V repeats respectively, suggests that there have been at least two independent horizontal *cwpV* repeat acquistition events in these strains, unless the source of the repeats already contained a mosaic. Our current knowledge of the repeat types found in different *C. difficile* strains is shown in Table 5.1. Typing of the *C. difficile* strains listed here was carried out by Dr. Jenny Emerson, Zoe Seager and Stina Linden.

This is by no means an exhaustive examination of *C. difficile* strains, and it seems likely that as more strains are investigated further novel CwpV repeat types will be discovered. Based on the strains analysed so far it seems that CwpV type exchange is not occurring at a very high frequency, as all strains tested from within one ribotype contain the same CwpV type. For example, all 078 strains tested, despite diverse geographic sources and different animal hosts, all have type I CwpV repeats. Current understanding of the phylogeny of the diversity of *C. difficile* strains is limited, with the most detailed study only including 5 different ribotypes (He et al., 2010). However, we can say that current data suggests a single CwpV type is associated with each clade; the HY (027) clade identified appears to all have type II CwpV, the A-B+ clade type IV CwpV, the HA2 (078) clade type I CwpV, and the HA1 clade type I CwpV. The locations of type III and V CwpV strains within this phylogenetic tree of *C. difficile* are not yet known. As more *C. difficile* strains are sequenced and our understanding of evolutionary relationships between strains improves it will become more clear as to when different CwpV types were acquired, and whether single or multiple independent acqusition events for each type have occurred.

Whatever the mechanism for CwpV type acquisition, it seems likely there is positive selection pressure for CwpV variablilty, as genetic changes are only maintained in a population if they produce successful genomes. Host immune pressure is one possible selection pressure that could promote CwpV variability. This selection pressure seems to have been at work promoting variability in other *C. difficile* surface proteins SlpA and Cwp66 (Lemee et al., 2005) and from this study it is clear that one key difference between CwpV types is their antigenicity. Variation in surface proteins may be involved in competition between different C. *diffiicle* strains. It has been reported that 60-70% of healthy infants are colonized by *C. difficile* and that there is an immune response to infection that can be protective in later life (Jangi & Lamont, 2010). This implies that there may be inherent differences in susceptibility of individuals to different strains depending on their histories of *C. difficile* infection and resulting immune memory. Interestingly, Type II CwpV has thus far only been found in ribotype 027 strains, which have exhibited recent worldwide spread and caused multiple hospital outbreaks (McDonald et al., 2005). It is possible that the susceptibility of individuals to 027 strains, that were previously rarely isolated, may be higher than to other antigenically different *C. difficile* strains that were more likely to have colonized individuals as infants. Further work is required to understand the relationships between prior *C. difficile* exposure and susceptibility. If susceptibility differences do exist CwpV may play a role.

Conserved post-translational processing of CwpV across different CwpV types suggests that this process is important for optimal CwpV expression and function. We have shown that cell wall proteases Cwp84 and Cwp13 are not responsible for CwpV cleavage (data not shown) and therefore the mechanism of CwpV

cleavage is not known. It is possible that CwpV undergoes protein autoprocessing, which has been reported for a number of families of proteins (Paulus, 2000, Tajima *et al.*, 2010, Dautin *et al.*, 2007). Given that we show a threonine residue to be directly downstream of the cleavage site CwpV, a plausible mechanism of autocleavage is N, O-acyl migration with the oxygen of the hydroxy group of threonine acting as the nucleophile. Such a reaction requires extreme chemical conditions (Crawhall & Elliott, 1955), unless catalysed by the tertiary structural organisation of the protein. This mechanism of peptide bond cleavage is deployed by a number of different protein families, which undergo autocleavage, for example the N-terminal nucleophile hydrolases (Brannigan et al., 1995), glycosylasparginases (Xu et al., 1999) and SEA proteins (Johansson et al., 2008). Targeted mutagenesis of specific residues at the CwpV cleavage site is likely to provide evidence for or against an autocleavage mechanism. Mutation of the threonine residue at the cleavage site to a serine or a cysteine may still allow cleavage to proceed due to the presence of a nucleophilic side chain, although cleavage may be less efficient. However, mutation to an alanine residue would abolish cleavage if it does rely on this autocleavage mechanism. The cleavage domain should also cleave in a heterologous system, for example in *E. coli*, if no exogenous protease is required for cleavage. Further experiments are required to test this hypothesis. The functional importance of this CwpV cleavage is currently unclear. If the two CwpV fragments do not disassociate after cleavage, there may be no real consequence to cleavage compared to if CwpV remained a single protein. There may however be conditions were the two fragments disassociate, which may have functional importance.

It has been shown that inhibition of SlpA cleavage leads to shedding of uncleaved SlpA into a *C. difficile* culture supernatent, suggesting that this unprocessed form of SlpA is not stable in the S-layer (Dang et al., 2010). By analogy it seems likely based on the mutagenesis study reported in this chapter that cleavage of CwpV is important for stable incorporation of CwpV into the *C. difficile* S-layer. This domain, the cleavage site and subsequent complex formation between the CWA and C-terminal domains are conserved in all CwpV proteins, again indicating the importance of this processing for optimal CwpV expression and function. Given the high proportion of the total S-layer that CwpV constitutes in *ΔrecVON* cells it seems likely that CwpV interacts positively with other S-layer proteins to maintain the integrity of S-layer packing. The loss of the CwpV CWA domain from the cell surface in the CDM4-6 mutants suggests that this domain could no longer be anchored in the S-layer. This may be due to loss of cell wall anchoring, although the PF04122 domains were not directly affected by the mutation. Interactions between the CwpV CWA domain and the HMW SLP may have been affected. Whether or not such interactions are important for *C. difficile* S-layer integrity is an interesting area for future investigation. If found to be important, such interactions may rely on the mature cleaved forms of these proteins.

The experimental demonstration that the CwpV CWA domain alone can anchor to the cell wall is the first direct evidence of the function of PF04122 domains. For S-layer proteins from some other bacterial species the mechanism of surface anchoring has been characterised (Sara, 2001), however the mechanism underpinning cell wall anchoring via PF04122 domains has not been characterised. It should be noted that such studies with SlpA have not been possible as it has not been possible to create an *slpA* knock-out mutant, perhaps

suggesting that *slpA* is an essential gene. The experimental setup developed in this study, expressing the CwpV CWA domain in *C. difficile* from a plasmid, will facilitate investigation into the details of the PF04122 cell wall anchoring mechanism, which will improve our understanding of the whole CWP family of proteins.

All types of CwpV being phase variable, suggests that this complex regulatory mechanism is important for optimal function of CwpV. Phase variation of bacterial surface proteins has been reported to be involved in a number of processes including immune evasion and colonization (van der Woude, 2006). Many such explanations behind CwpV phase variable regulation can be suggested, and this will be discussed further in Chapter 6, which focuses on CwpV function.

Table 5.1 CwpV characteristics of *C. difficile* strains examined

Strain	CwpV architecture		Other characteristics		Reference or source
	Type	No. of repeats	Ribo-type	Toxin status	
CDKK371	I	6 [b]	001	+ve	(Calabi & Fairweather, 2002b)
R8366	I	4 [b]	001	A+, B+	(Calabi *et al.*, 2001a)
R13537	I	6[a]	001	+ve	(Eidhin et al., 2006)
R12879	I	6 [a]	001	+ve	(Eidhin et al., 2006)
R14628	I	6 [a]	001	+ve	(Eidhin et al., 2006)
Y	I	4 [b]	010	A-, B-	(Calabi & Fairweather, 2002b)
630	I	9 [b, c]	012	A+, B+	(Calabi et al., 2001a)
CDKK959	I	4 [b]	053	+ve	(Calabi & Fairweather, 2002b)
CDKK101	I	4 [b]	053	+ve	(Calabi & Fairweather, 2002b)
CDKK291	I	4 [b]	053	+ve	(Calabi & Fairweather, 2002b)
M120	I	4 [a]	078	A+, B+	B. Wren
C22	I	4 [a]	078	A+, B+	S. d'Arc
C36	I	4 [a]	078	A+, B+	S. d'Arc
H57	I	4 [a]	078	A+, B+	S. d'Arc
H94	I	4 [a]	078	A+, B+	S. d'Arc
R20291	II	8 [b, c]	027	A+, B+; CDT+	(Stabler *et al.*, 2006a)
R20352	II	8 [b]	027	A+, B+; CDT+	(Stabler et al., 2006a)
R20928	II	8 [a]	027	A+, B+; CDT+	(Stabler et al., 2006a)
CD196	II	8 [a]	027	A+, B+; CDT+	(Popoff et al., 1988)
R12628	II	8 [a]	027	+ve	J. Brazier
R16760	II	8 [a]	027	+ve	J. Brazier
CDKK167	III	6 [b]	016	+ve	C. Kelly
NF2029	III	6 [a]	106	A+, B+	M. Wilcox
R7404	III/IV	1/7 [a]	017	A–, B+	(Calabi et al., 2001a)
M9	III/IV	1/7 [b]	017	A–, B+	D. Drudy
AY1	III/V	2/5 [b]	N.D.	A–, B-	D. Gerding
AY2	III/V	2/5 [b]	N.D.	A–, B-	D. Gerding
AY3	III/V	2/5 [b]	N.D.	A–, B-	D. Gerding
AY4	III/V	2/5 [b]	N.D.	A–, B-	D. Gerding
AY5	III/V	2/2 [b]	N.D.	A–, B-	D. Gerding
B-one (B1)	I	9 [a]	N.D.	A+, B+	Gill Douce; (Merrigan *et al.*, 2003a)

The types of repeats and their number were determined by [a]PCR (this study), [b]DNA sequencing (this study) or [c] genome sequencing studies http://www.sanger.aC.uk/resources/downloads/bacteria/*Clostridium-difficile*.html; N.D., not determined; +ve, toxin A+ve but toxin B unknown.

6 Investigating the Function of CwpV

6.1 Introduction

Chapter 4 described the discovery of the phase variable regulation of CwpV, elucidation of the mechanism underpinning phase variation and determination of the essential role of the recombinase RecV in CwpV phase variation. Chapter 5 concerned CwpV protein domain structure, post-translational processing and C-terminal repeat sequence variability exhibited across *C. difficile* strains. So far it has been discussed that many aspects of CwpV are conserved across diverse sets of *C. difficile* strains; phase variable regulation, the CWA domain, the cleavage domain, formation of a CWA/C-terminal fragment complex and repetitive C-terminal structure. However, we have shown that there are distinct types of CwpV repeat, which appear to have been acquired by independent horizontal gene transfer events. These repeats are antigenically distinct. All this information hints at a complex evolutionary history for CwpV but does not provide any direct information as to the function of CwpV in the *C. difficile* life cycle. The previous chapters also describe the production of mutant *C. difficile* strains that are ideal for subjecting to functional assays in order to investigate the function of CwpV, which is the focus of this final results chapter.

Clues as to protein function can be gained from searching the databases of known proteins. Bioinformatic analysis of CwpV identifies the presence of the PF04122 domains in the CWA domain, shown in the previous chapter to be responsible for cell wall anchoring. The cleavage domain does not contain any Pfam domains and using BLASTP does not show significant homology to any well-characterised proteins in the protein database. The repeat domains have

also been analysed by bioinformatics searches. Type I repeats do not contain any Pfam motifs, nor show significant sequence homology to any well-characterised proteins. The best hit is with a hypothetical protein from the bacterium *Shuttleworthia satelles* but this only shows 29% amino acid identity with an e value of 2.7. Type II repeats do contain a Pfam motif, the PF07554 FIVAR motif, which is an uncharacterized sugar binding domain. This Pfam domain is found in proteins with varying functions. Of note it is present in a fibronectin-binding surface protein from *Staphylococcus epidermis.* This Pfam domain shows a lack of conserved catalytic residues, which can be characteristic of a ligand binding domain. However this Pfam family is not characterised to the extent that the binding ligand can be identified by sequence analysis. The best match from a BLASTP search for a type II repeat is a putative surface expressed glycosyl hydrolase from *Clostridium beijerinckii* which shows 49% sequence identity, with an e value of $8e^{-9}$. This protein is however not well characterised, so little can be said about the relevance of this homology to type II repeat function. Type III repeats do not contain any Pfam motifs. The most significant homology match found by BLASTP is a predicted alpha-L-fucosidase S-layer protein (containing an SLH domain) from Paenibacillus sp., which is predicted to have sugar binding function. This shows 52% sequence identity with a type III repeat, with an e value of $2e^{-14}$. This protein and none of the other significant matches by BLASTP have been well-characterised though, therefore direct interpretation of the homology to type III repeat function is not possible. Similarly the CwpV type IV repeat contains no Pfam motifs, with the best BLASTP hit being a hypothetical putative S-layer protein from *Syntrophomonas wolfei* which shows 44% sequence identity with an e value of $5e^{-10}$. Finally the type V CwpV repeat contains no Pfam motifs, and its closest homolog in the database is a hypothetical protein from *Clostridium*

carboxidivorans with a predicted signal peptide, which has 53% sequence identity to type V repeat and gives an e value of $3e^{-8}$. Therefore bioinformatic analysis of the CwpV repeats does not give any clear indication as to their likely function.

Most of these proteins to which the repeats show a level of homology do not themselves contain repeats of the region of homology. This was determined by analysis of the homologous sequences using RADAR (Rapid Automatic Detection and Alignment of Repeats) (Heger & Holm, 2000). An interesting exception to this is a predicted S-layer domain protein from Paenibacillus sp., which contains a repetitive domain with 54% sequence identity to the six type III repeats found in CDKK167, with an e value of $1e^{-59}$. The function of this homologous protein has not been investigated, therefore little can be said about the significance of this homology. Overall, the evidence suggests that the generation of repeats has occurred since the sequences have been acquired by *C. difficile*. This is further supported by the discovery that different numbers of repeats of a given type have been found in different strains, for example 630 contains 9 type I repeats, whilst CDKK371 contains contains only six. This is indicative of the numbers of CwpV repeats increasing or decreasing since the introduction of the sequence to the species.

It has been reported that in prokaryotes for proteins of >1000 amino acids in length ~ 20% contain repeats (Marcotte et al., 1999). Therefore repetition of domains appears to be a successful way of generating large proteins in prokaryotes. This high prevalence of repeating proteins is thought to be due to the fact that repetition of protein domains is a good way to generate proteins of novel function, in a similar way to how gene duplication and genome duplication have

been thought to allow diversification of genes and genomes (Wolfe, 2001). In proteins where the repeats can be clearly identified it is likely that repeat generation is a relatively recent evolutionary event. This is further confirmed in the case of CwpV, as *C. difficile* is known to be an ancient species with an estimated age of 1.1-85 billion years (He et al., 2010), but the existence of these distinct repetitive *cwpV* genes in different strains suggests recent evolutionary change.

Generation of repeats can occur during DNA replication via recombination events or slip-strand mispairing. For long repeats, as seen in CwpV, recombination is the more likely mechanism for generation of repeats (Marcotte et al., 1999). Whatever the mechanism of repeat generation, a selective advantage conferred by domain repetition is required to maintain the repetitive sequence. One obvious such advantage conferred by repetition is increased avidity to ligands via multiple copies of the binding domain. This avidity is conferred by a much reduced overall dissociation rate when multiple interactions rather than a single interaction must dissociate simultaneously in order for two molecules to dissociate fully, which has been shown experimentally using engineered antibody molecules (Zhou, 2003).

Given the surface expression of CwpV, its repetitiveness and homology of repeat domains with various ligand-binding proteins, it seemed plausible that CwpV may have a ligand binding function. In other pathogenic bacteria repetitive surface proteins are known to function as adhesins, for example the MSCRAMM family of extracellular matrix binding proteins (Schwarz-Linek et al., 2004) and therefore it seemed possible that CwpV may also act as an adhesin. However, as previously mentioned, large bacterial proteins are often repetitive and therefore this characteristic in itself is not indicative of a specific functional role.

In this chapter the role of CwpV in the *C. difficile* life cycle is investigated using available *in vitro* assays. CwpV is shown to be an aggregation-promoting factor, and exhibits fibronectin binding activity but does not appear to directly affect adhesion of *C. difficile* to Caco-2 cells. The expression level of CwpV under different conditions is also investigated, as this may provide insight into its function. No conditions have been identified where expression of CwpV is specifically switched ON or OFF. The implications of these findings for future investigations into CwpV function, including *in vivo* experiments, are discussed.

6.2 Results

For clarity a table of the strains used in this chapter is shown in Table 6.1.

Table 6.1 *C. difficile* strains used in Chapter 6

NF strain name	Descriptive Strain name	Strain genotype	Plasmid	CwpV repeat type	% cells expressing CwpV
2147	**WT**	630*Δerm*	pMTL960*	I	1-10
2166	***ΔcwpV***	630*ΔermΔcwpV*	pMTL960*	I	0
2151	***ΔcwpV*(pOENter)**	630*ΔermΔcwpV*	pCBR067 pOENter	-	100
2167	***ΔcwpV*(pOEI)**	630*ΔermΔcwpV*	pCBR080 pOEI	I	100
2176	***ΔcwpV*(pOEII)**	630*ΔermΔcwpV*	pCBR105 pOEII	II	100
2177	***ΔcwpV*(pOEIII)**	630*ΔermΔcwpV*	pCBR106 pOEIII	III	100
2178	***ΔcwpV*(pOEIV)**	630*ΔermΔcwpV*	pCBR107 pOEIV	III/IV	100
2179	***ΔcwpV*(pOEV)**	630*ΔermΔcwpV*	pCBR109 pOEV	III/V	100
2225	***ΔrecVON***	630*ΔermΔrecV*	pMTL960*	I	100
2226	***ΔrecVOFF***	630*ΔermΔrecV*	pMTL960*	I	0

* For all experiments comparing the full panel of strains, inclusion of empty vector pMTL960 allowed all strains to be under thiamphenicol selection. For sporulation and germination experiments WT, Δ*cwpV,* Δ*recVON* and Δ*recVOFF* without plasmids are used.

6.2.1 All *C. difficile* strains have the same exponential growth rates

The growth rates for each *C. difficile* strain were analysed. *C. difficile* overnight cultures were grown in triplicate in TYG medium supplemented with thiamphenicol to select for pMTL960-based plasmids (present in all strains, see Table 6.1). These cultures were subcultured into fresh growth medium and the OD_{600nm} was monitored. A graphical representation of the results is shown in Figure 6.1. As can be seen the exponential growth rates, represented by the gradient of the growth curves between OD_{600nm} of 1.5 - 2.5, of all strains are equal. This suggests that under standard laboratory conditions the level of CwpV expression and the type of CwpV expressed does not affect growth rate.

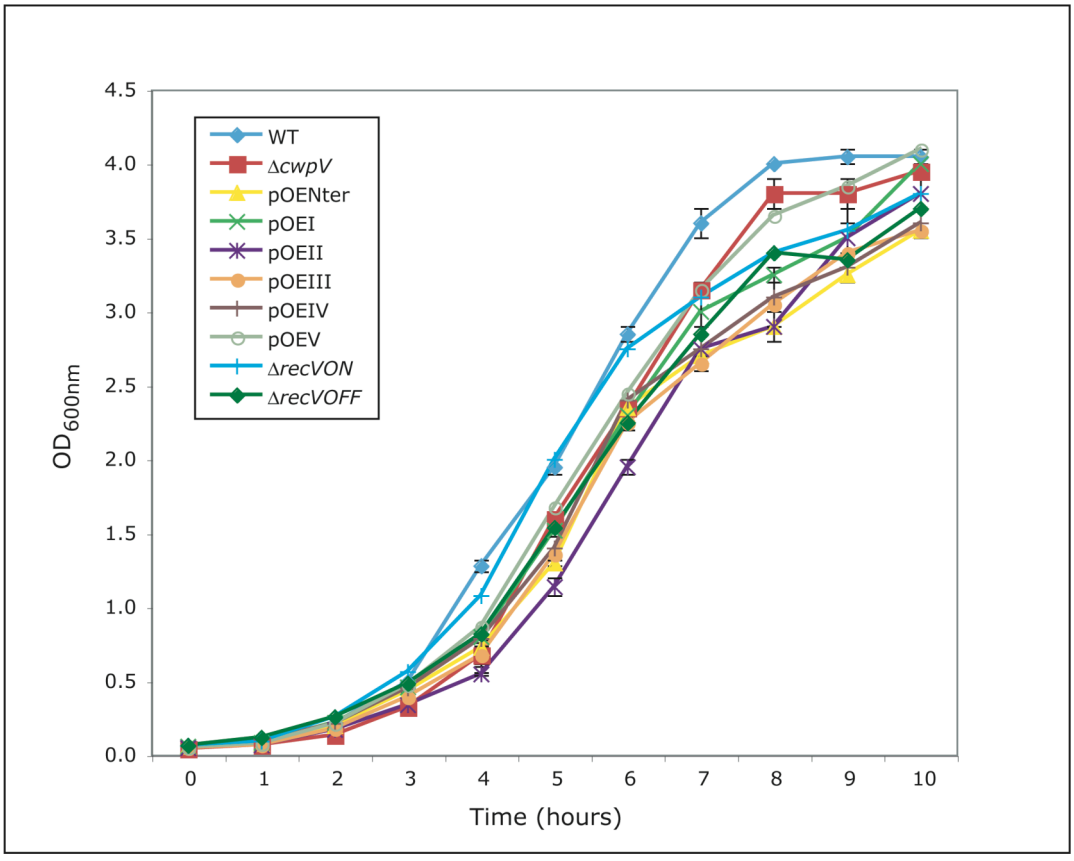


Figure 6.1 Growth curves of *C. difficile* strains.

C. difficile cultures in TYG medium supplemented with thiamphenicol were monitored for growth by OD_{600nm}. Cultures were grown in triplicate with shaking under standard anaerobic conditions. The mean OD_{600nm} value at each timepoint is plotted against time. Error bars represent the standard deviation of the triplicate measurements.

6.2.2 *C. difficile* strains exhibit colony morphology differences

The colony morphology of *C. difficile* strains was assessed by growth of *C. difficile* on solid BHI agar plates supplemented with thiamphenicol. Plates were inoculated from *C. difficile* overnight cultures at an appropriate dilution such that bacterial growth occurred as single isolated colonies. In Figure 6.2 photographic images of the topside of *C. difficile* colonies grown for 36 hours can be seen. There is a clear difference between the strains with high levels of CwpV expression (Δ*recVON,* Δ*cwpV*(pOEI-V)) and low or no CwpV expression (WT, Δ*cwpV* and Δ*recVOFF).* The strains with high levels of CwpV expression exhibit a small, round colony morphology, although this is less marked for Δ*recVON* than for strains expressing each of the five CwpV types from a plasmid Δ*cwpV*(pOEI-V). The smaller colonies also appear denser and shinier. The Δ*cwpV*(pOENter) colonies appear similar to WT, Δ*cwpV* and Δ*recVOFF* colonies, suggesting that it is the repeat domains of CwpV that confer the morphology difference.

In order to further investigate this difference in morphology strains were grown for 6 days before analysis of colony morphology. As can be seen in Figure 6.3, after this extended period of growth on solid medium WT *C. difficile* colonies exhibit sprawling colony morphologies with branching edges of growth. This is observed for WT, Δ*cwpV,* Δ*recVOFF* and Δ*cwpV*(pOENter), whilst Δ*recVON* and Δ*cwpV*(pOEI-V) have much rounder edges. The images in Figure 6.3 were taken by scanning through plates, and the increased density of the Δ*recVON* and Δ*cwpV*(pOEI-V) colonies compared to the more diffuse WT, Δ*cwpV,* Δ*recVOFF* and Δ*cwpV*(pOENter) colonies can be seen. Therefore high levels of CwpV expression cause a dense, tight, shiny pattern of colony growth on solid medium, with little branching at the edge of colonies. This colony morphology phenotype is

conferred by the repetitive C-terminal domains of CwpV rather than the CWA domain. All CwpV repeat types confer this phenotype, despite their distinct sequences.

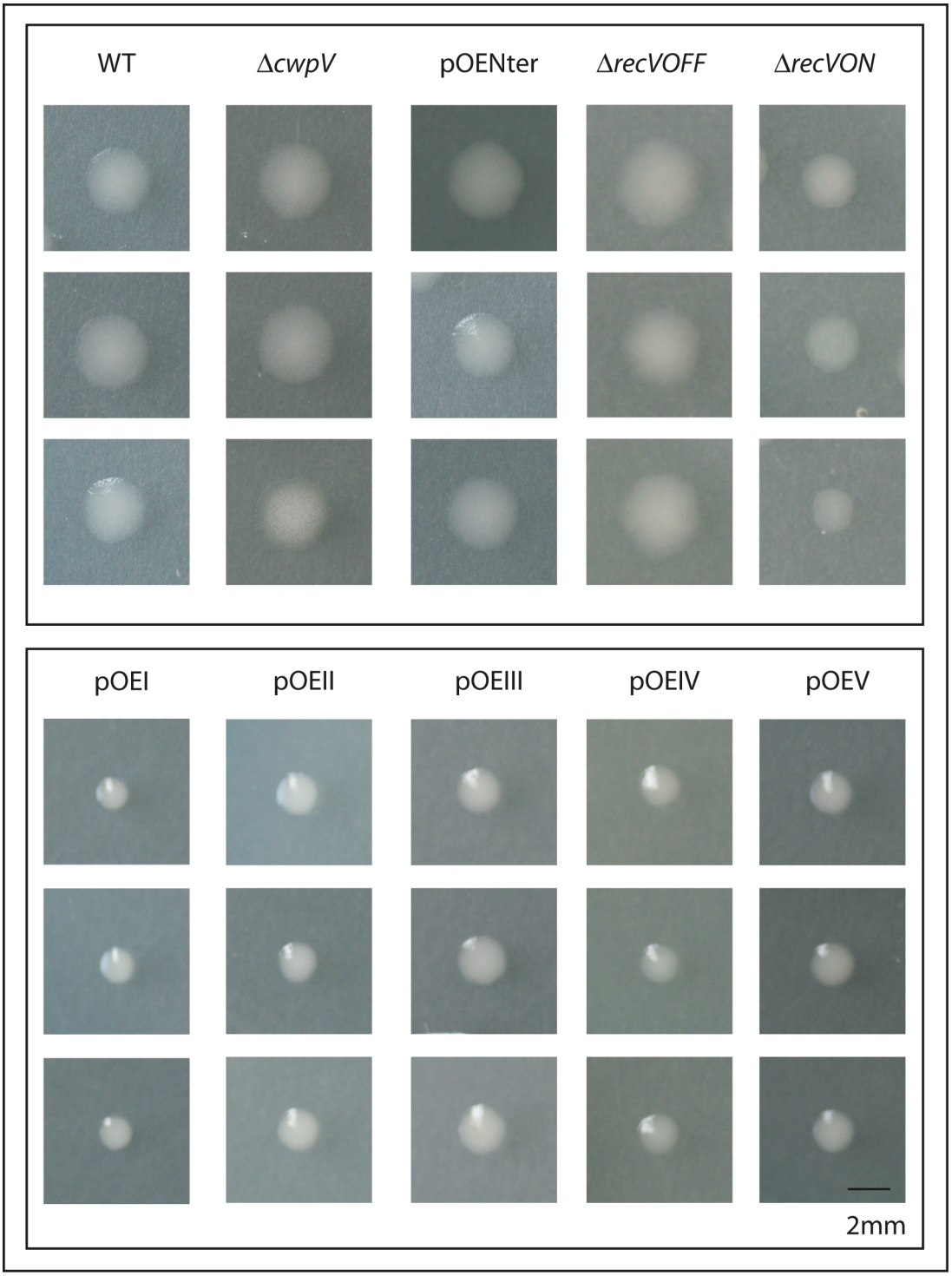


Figure 6.2 *C. difficile* mutants exhibit colony morphology differences.

Photographic images of the top-side of *C. difficile* colonies grown on BHI agar supplemented with thiamphenicol for 36 hours under standard anaerobic conditions.

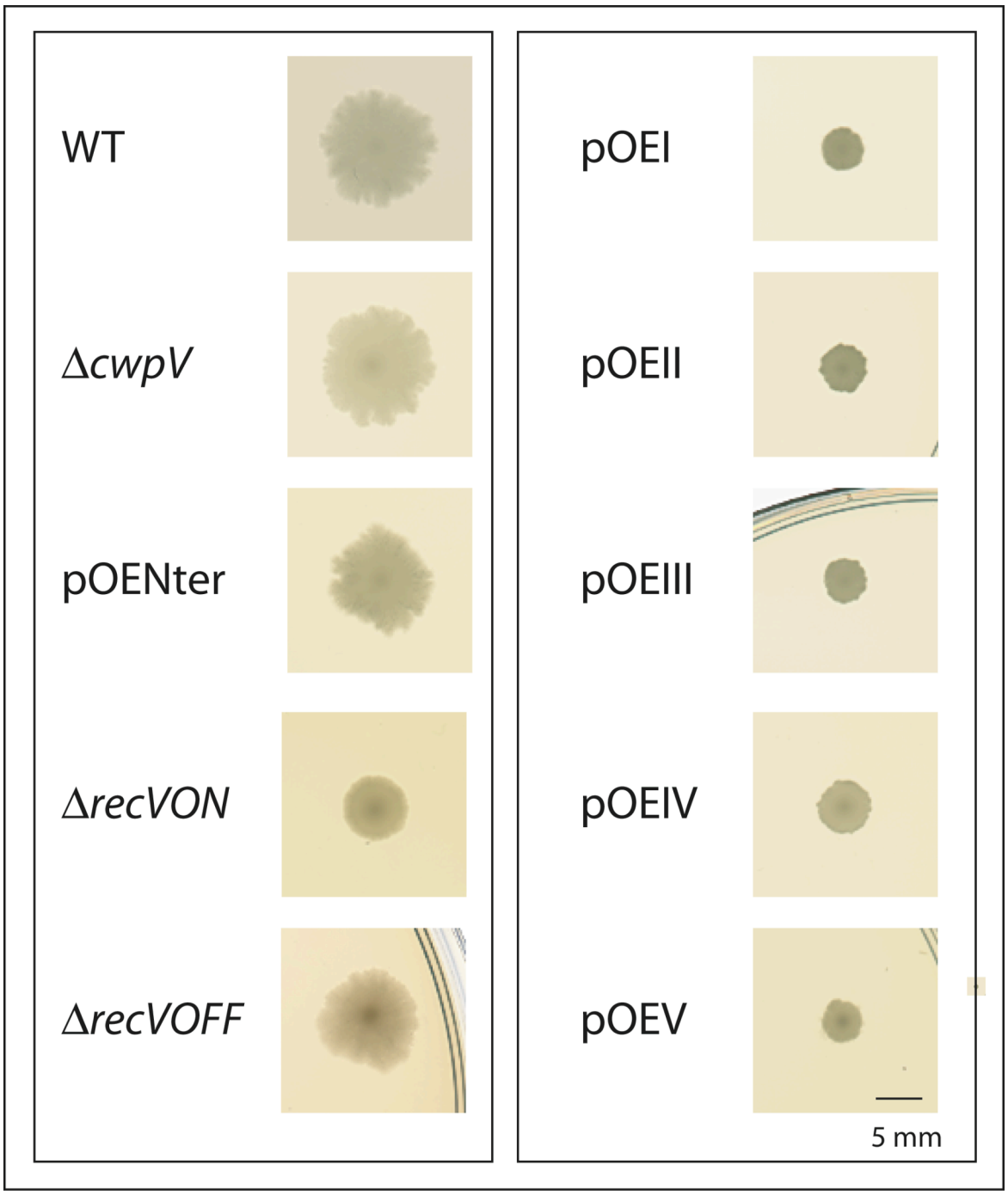
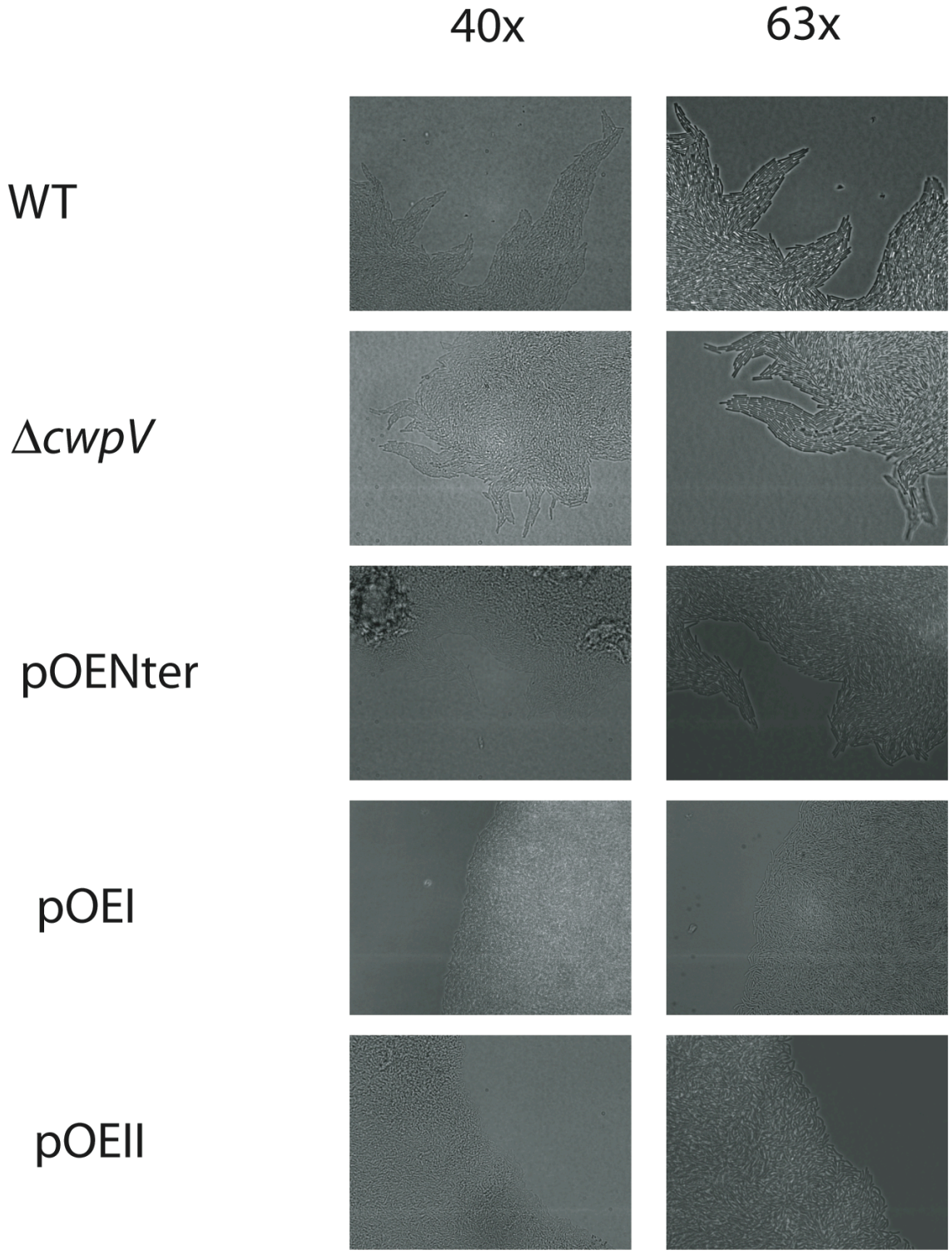


Figure 6.3 *C. difficile* strains exhibit colony morphology differences.

Scanned images of *C. difficile* colonies grown on BHI agar supplemented with thiamphenicol for 6 days under standard anaerobic conditions.

6.2.3 Microscopic analysis of *C. difficile* growth solid medium

To understand the basis of these colony morphology differences at the cellular level, colonies were grown in glass-bottomed Petri dishes, between the glass and the solid growth medium, to allow microscopic analysis of the edges of colony growth. Images of the edges of growth at 40x and 63x magnification are shown in Figure 6.4. In WT, Δ*cwpV,* Δ*recVOFF* and Δ*cwpV*(pOENter) images the growth can be seen to be quite diffuse at the edges of colonies, with a clear single layer of cells visualized. These cells appear to grow with a certain level of cellular alignment, forming directional protrusions that appear as flare-like patterns of growth. In contrast Δ*recVON* and strains expressing each of the five CwpV types from a plasmid Δ*cwpV*(pOEI-V) exhibit a more densely packed edge of growth, with a more randomly arranged packing of cells. This seems to correlate with the morphology seen macroscopically with WT, Δ*cwpV,* Δ*recVOFF* and Δ*cwpV*(pOENter) having directional protrusions leading to ruffled edges, whilst Δ*recVON* and Δ*cwpV*(pOEI-V) have no directional protrusions leading to smooth edges.


40x
63x
WT
ΔcwpV
pOENter
pOEI
pOEII

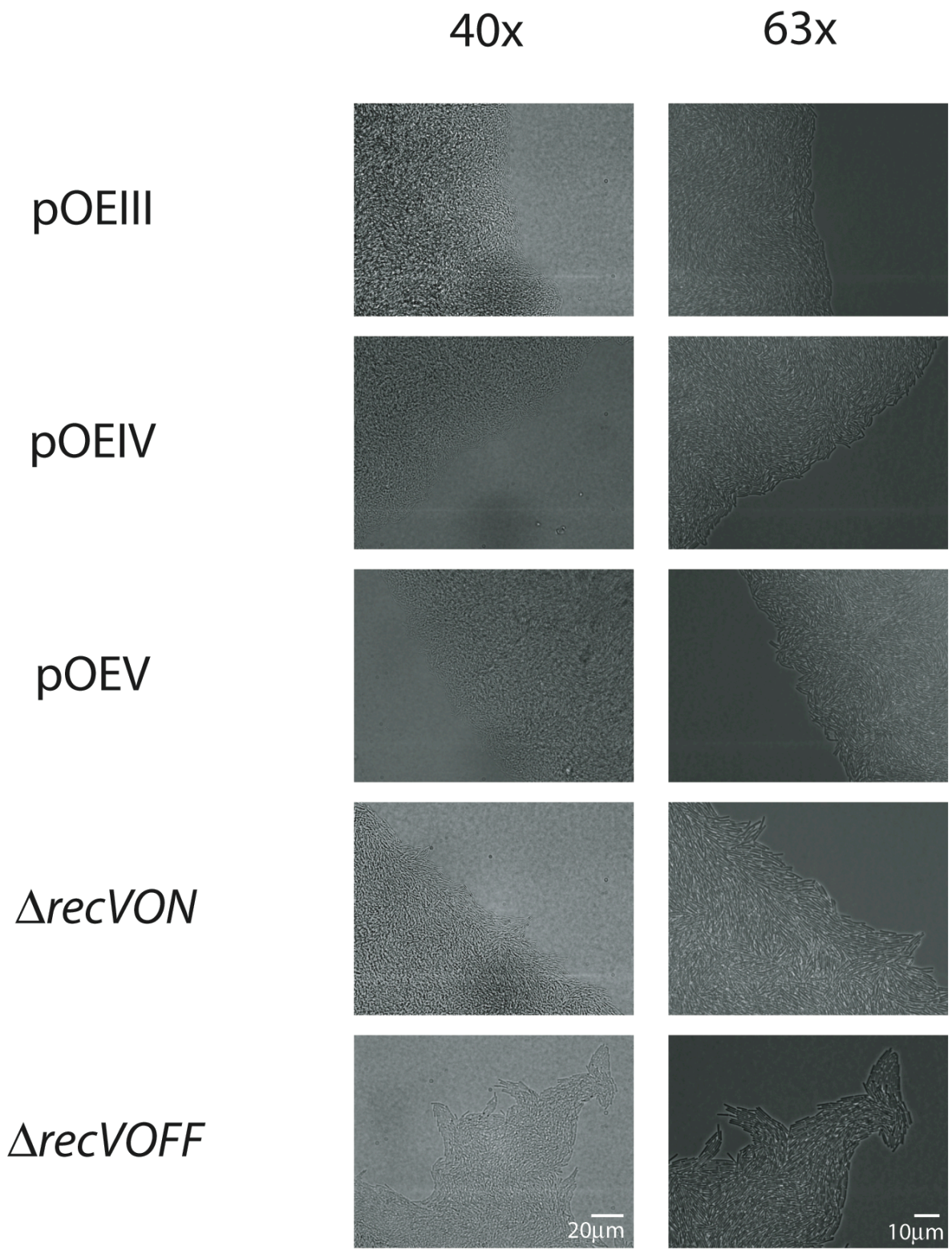


Figure 6.4 Microscopic images of the edges of *C. difficile* colony growth.

C. difficile strains were grown between glass and solid growth medium to allow analysis of growth by microscopy. Images of the edges of colony growth at 40X and 63X magnification are shown. Differences in cellular packing between strains, attributable to CwpV expression, can be seen.

6.2.4 CwpV is an Aggregation Promoting Factor

Such a colony morphology change caused by CwpV over-expression could be explained by an aggregation-promoting function of CwpV. Therefore the propensity of *C. difficile* to aggregate from an even cellular suspension was assessed. Initial assessment was carried out with 630 CwpV (type I) over-expression in Δ*cwpV*(pOEI) compared to WT, Δ*cwpV* and Δ*cwpV* (pOE Nter). Bacteria were grown overnight under standard conditions, harvested by centrifugation and resuspended to an OD_{600nm} of ~3 in round bottomed glass tubes. After 7 hours the appearance of the tubes was photographed. As can be seen from Figure 6.5A, CwpV expression did cause aggregation of cells. This was dependent on the repetitive domain, as expression of the N-terminal CWA domain alone did not cause aggregation. To assess this aggregation as a function of time, the meniscus OD_{600nm} of *C. difficile* suspensions was monitored over time. In Figure 6.5B the decreasing OD_{600nm} over time at the meniscus of the suspension is caused by Type I CwpV over-expression and reflects the aggregation of cells. Having established that Type I CwpV could promote aggregation of *C. difficile* cells from a suspension, aggregation promoting function was assessed in the same way for all CwpV types. It can be clearly seen in Figure 6.5C that over-expression of all CwpV types, except type IV, can induce aggregation of *C. difficile* from a suspension with a starting OD_{600nm} of 10. Aggregation was followed in real time, and could be detected in tubes where aggregation was occurring from as soon as 1 hour after suspension of bacteria. Tubes were left for 16 hours and photographed as shown in Figure 6.5C. No further aggregation was observed if tubes were left for 48 hours.

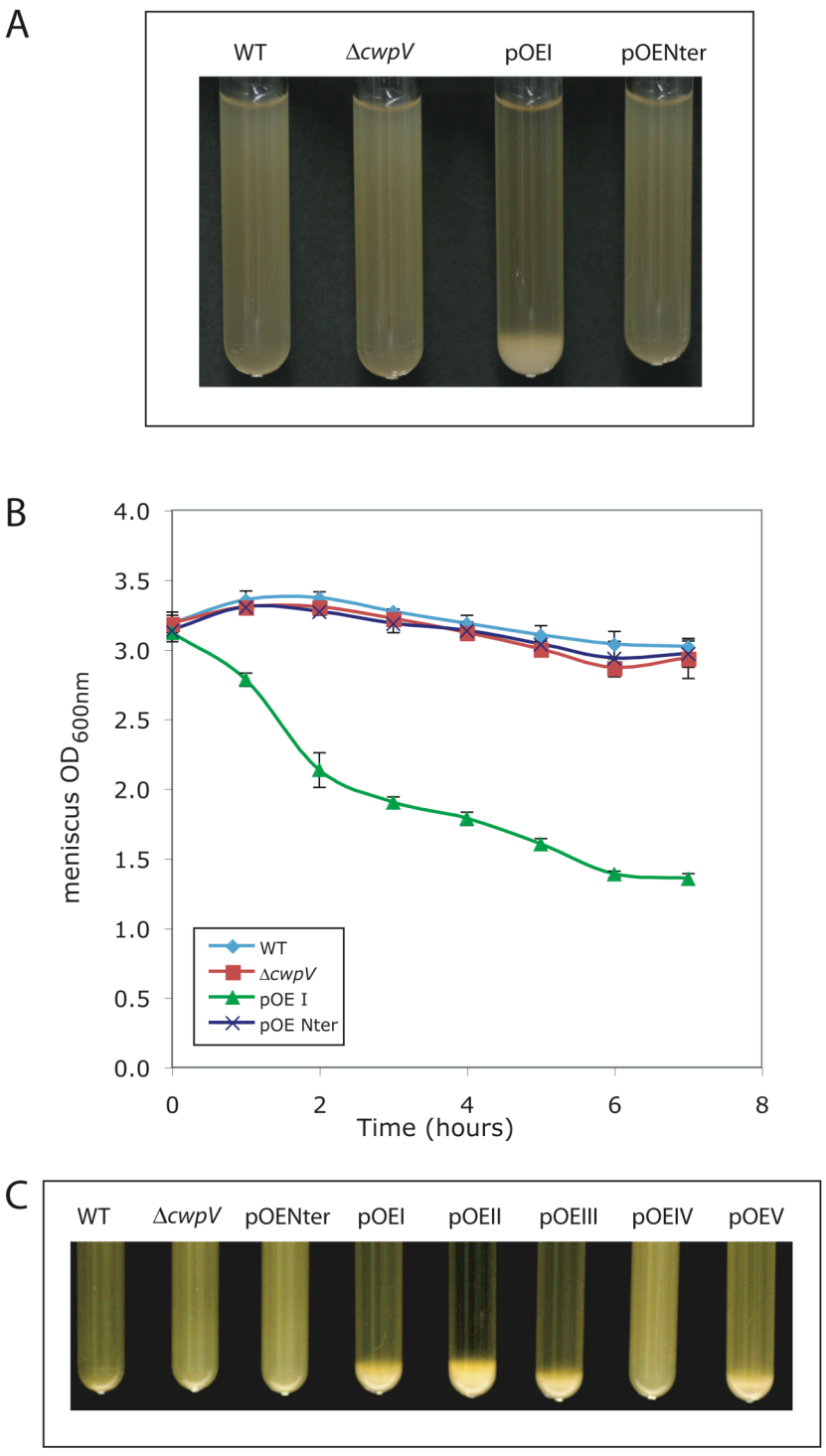


Figure 6.5 CwpV is an Aggregation Promoting Factor.

A. *C. difficile* strains were grown overnight under standard conditions. Cells were harvested by centrifugation then resuspended to an OD_{600nm} of ~3 in round bottomed glass tubes. After 7 hours the appearance of the tubes was photographed. Type I CwpV expression promoted cellular aggregation, dependent on the repetitive domain. **B.** The meniscus OD_{600nm} of *C. difficile* suspensions was monitored over time and the decreasing OD_{600nm} at the meniscus caused by Type I CwpV overexpression reflects the aggregation of cells. **C.** A large panel of *C. difficile* strains were resupended to an OD_{600nm} of 10 and after 16 hours the tubes were photographed. All CwpV types except type IV promoted aggregation of cells.

It was found that only 630 CwpV (type I) induced aggregation of *C. difficile* from suspensions with an OD_{600nm} of 3, while 630 CwpV and R20352 CwpV (type II) induced aggregation from an OD_{600nm} of 6 (data not shown). An OD_{600nm} as high as 40 did not lead to type IV-mediated aggregation, nor aggregation of WT, *Δcwp*V or *Δcwp*V(pOENter) (data not shown). Therefore it appears that ability to induce aggregation is highest for 630 CwpV, followed by R20352 CwpV, then CDKK167 and AY1 CwpV, with no observed propensity of M9 CwpV or the CwpV CWA fragment to induce aggregation of *C. difficile* from suspension.

6.2.5 CwpV expression in *C. difficile* does not affect FliC expression or swimming motility

Functional interaction between bacterial autoaggregation and motility has been reported in *E. coli,* with the autoaggregative protein antigen 43 (Ag43) and flagella having opposing functional roles (Ulett et al., 2006). *C. difficile* is known to exhibit flagellar-mediated motility (Twine et al., 2009), therefore the effect of CwpV expression on *C. difficile* swimming motility and flagellar expression was investigated. *C. difficile* cultures were analysed for flagella expression based on the release of FliC from intact cells after vigorous vortexing (Twine et al., 2009). Figure 6.6A shows the analysis of *C. difficile* flagella expression by SDS-PAGE and Coomassie staining, with 630 FliC running at ~33 kDa. No significant differences in FliC expression are caused by different levels of CwpV expression. In order to assess the swimming motility of *C. difficile* strain BHI 0.175% agar was poured into sterile tubes and allowed to cool to room temperature. Tubes were then transferred to the anaerobic cabinet and left for 4 hours to equilibrate under anaerobic conditions. Tubes were inoculated to a defined depth with the straight

end of sterile inoculation loops that had been dipped into *C. difficile* overnight liquid cultures. Tubes were incubated overnight then photographed to document motility. Images of *C. difficile* strains with varying levels of CwpV expression are shown in Figure 6.6B. No significant differences in *C. difficile* swimming motility were observed due to CwpV expression.

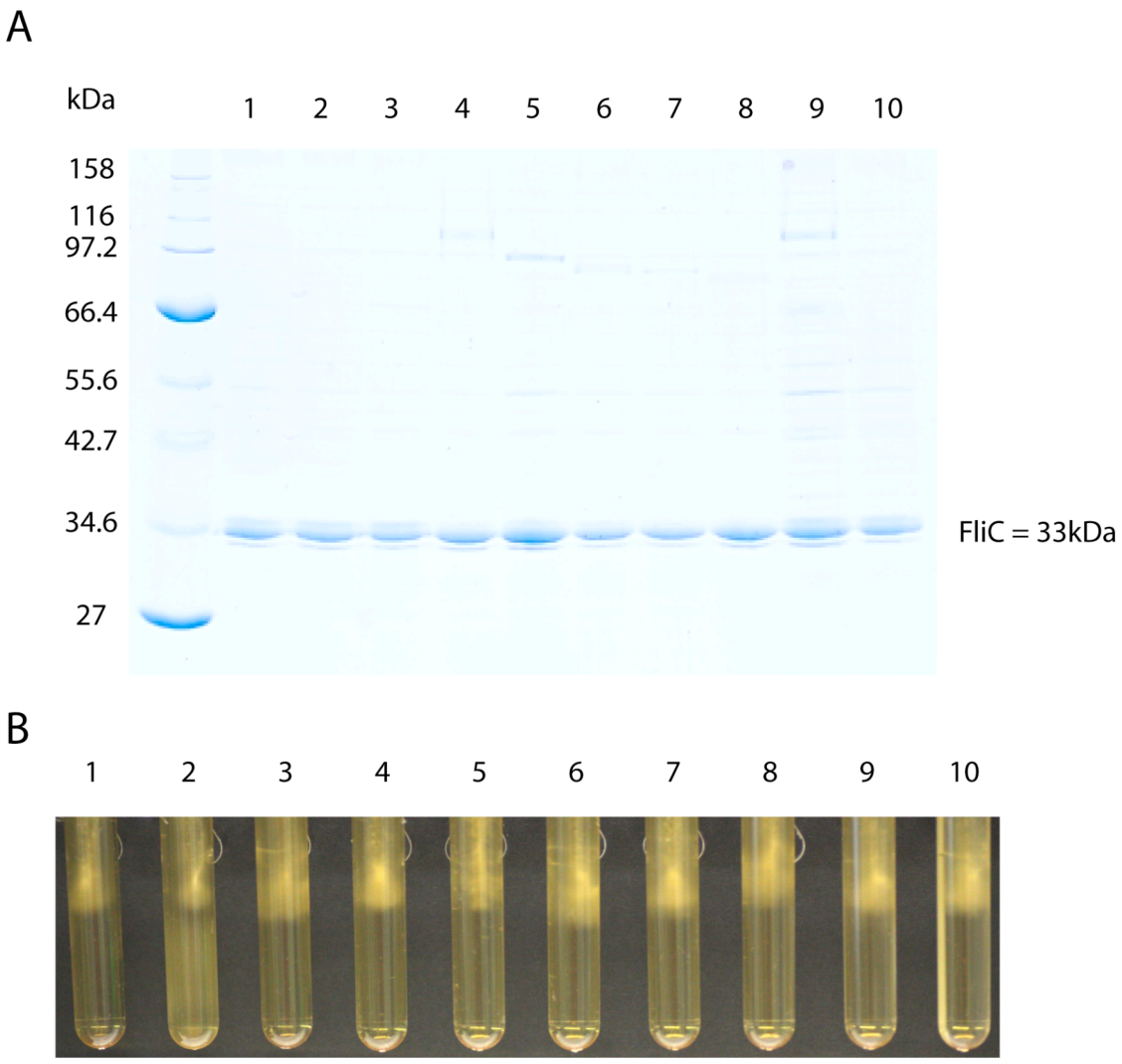


Figure 6.6 CwpV expression does not affect FliC expression or swimming motility.

A. *C. difficile* strains were analysed for FliC expression based on a published protocol (Twine et al., 2009). *C. difficile* FliC expression from the panel of strains with varying levels of CwpV expression was assessed by SDS-PAGE and Coomassie staining, with 630 FliC running at ~33 kDa. Lanes 1: WT, 2: Δ*cwpV,* 3: pOENter, 4-8: pOEI-V, 9: Δ*recVON,* 10: Δ*recVOFF.* **B.** To assess swimming motility of *C. difficile* strains BHI 0.175% agar was inoculated to a defined depth with *C. difficile* from overnight liquid cultures. Tubes were incubated overnight then photographed to document motility. Tubes are numbered in the same way as the lanes in A.

6.2.6 CwpV exhibits fibronectin binding activity

Given the recognition that CwpV may act as an adhesin, as well as promoting *C. difficile* aggregation, purified repeats of CwpV were tested for binding to commercially available extracellular matrix proteins. DNA fragments encoding one or two CwpV repeats were cloned as indicated in Figure 6.7A and were expressed and purified via a C-terminal His-tag. Seven commercially available mammalian extracellular matrix proteins were tested for interaction with these CwpV repeats by ELISA. 96-well plates were coated with the extracellular matrix proteins, which were then incubated with four different concentrations of CwpV repeat proteins. Detection of CwpV repeat binding to the immobilized extracellular matrix protein was achieved using an anti-His tag-HRP conjugate antibody and a colorimetric HRP substrate. The results of these seven ELISAs are shown in Figure 6.7A-H. The only interaction detected was between plasma fibronectin and CwpV repeat types I and IV. To follow up this result, a more stringent ELISA was set up to assess plasma fibronectin binding. All repeat types were tested in triplicate against plasma fibronectin and BSA (as a control), using a range of concentrations of CwpV repeats. Purified recombinant LMW SLP was also assayed to act as a further negative control. The results of these assays (Figure 6.8A and B) indicate that CwpV repeats types I and IV interact with plasma fibronectin in a concentration dependent manner, but do not interact with BSA. In order to study this interaction further it was decided to focus on CwpV type I repeats. A final ELISA with 6 replicate wells per concentration of repeat was carried out to compare CwpV type I binding to plasma fibronectin and BSA, which further confirmed this interaction (Figure 6.8C).

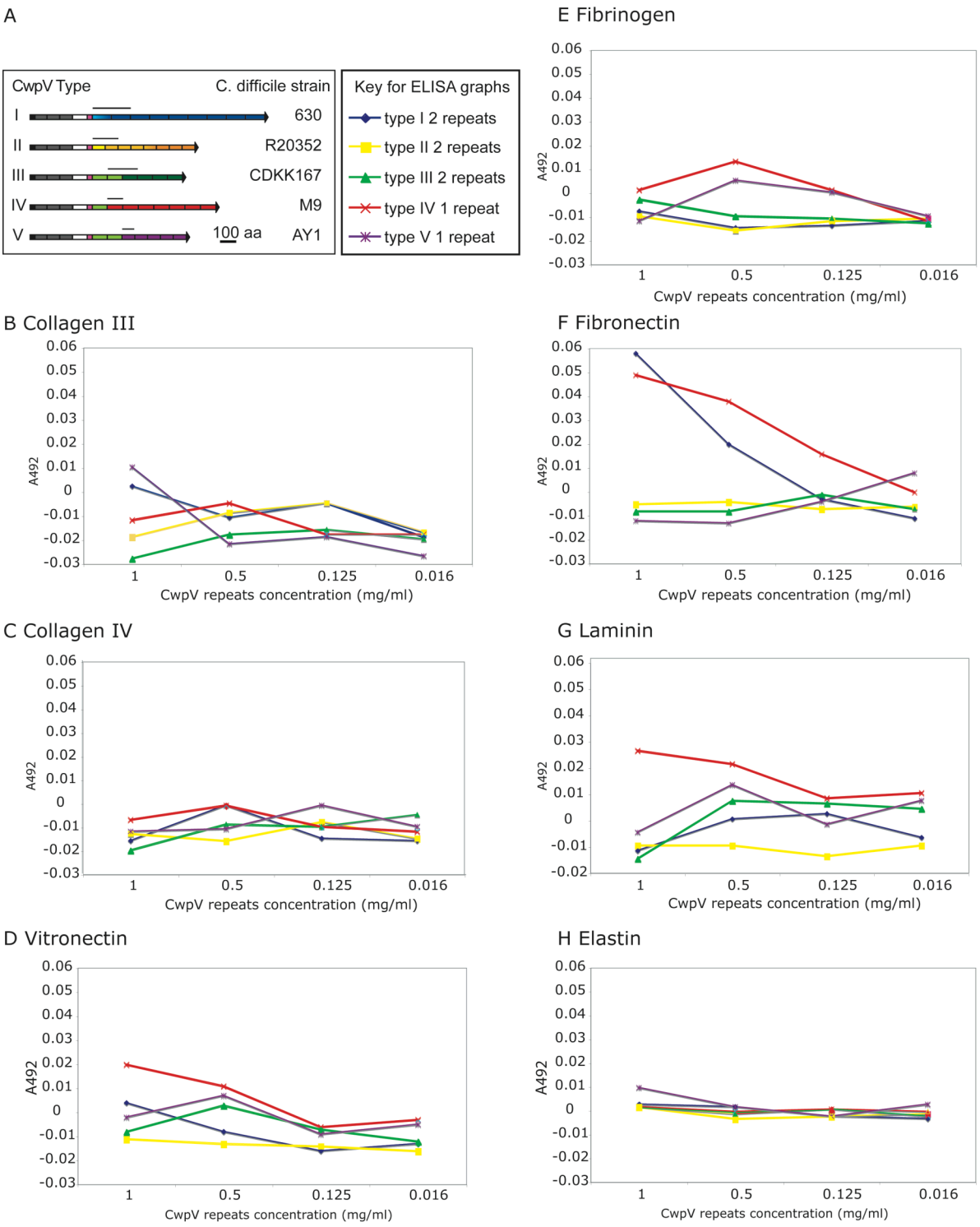


Figure 6.7 Screening for interactions between CwpV repeats and extracellular matrix proteins by ELISA.

A. Diagrammatic representation of repeat fragments cloned and purified from *E. coli* for use in ELISAs with extracellular matrix proteins. The legend for all ELISA graphs is also shown. **B-H.** ELISAs using four different concentrations of CwpV repeats binding to collagen III, collagen, IV, vitronectin, fibrinogen, fibronectin, laminin, elastin respectively. Binding was detected using an anti-his-HRP antibody conjugate with a colorimetric HRP substrate. A492 represents the amount of CwpV repeat binding detected.

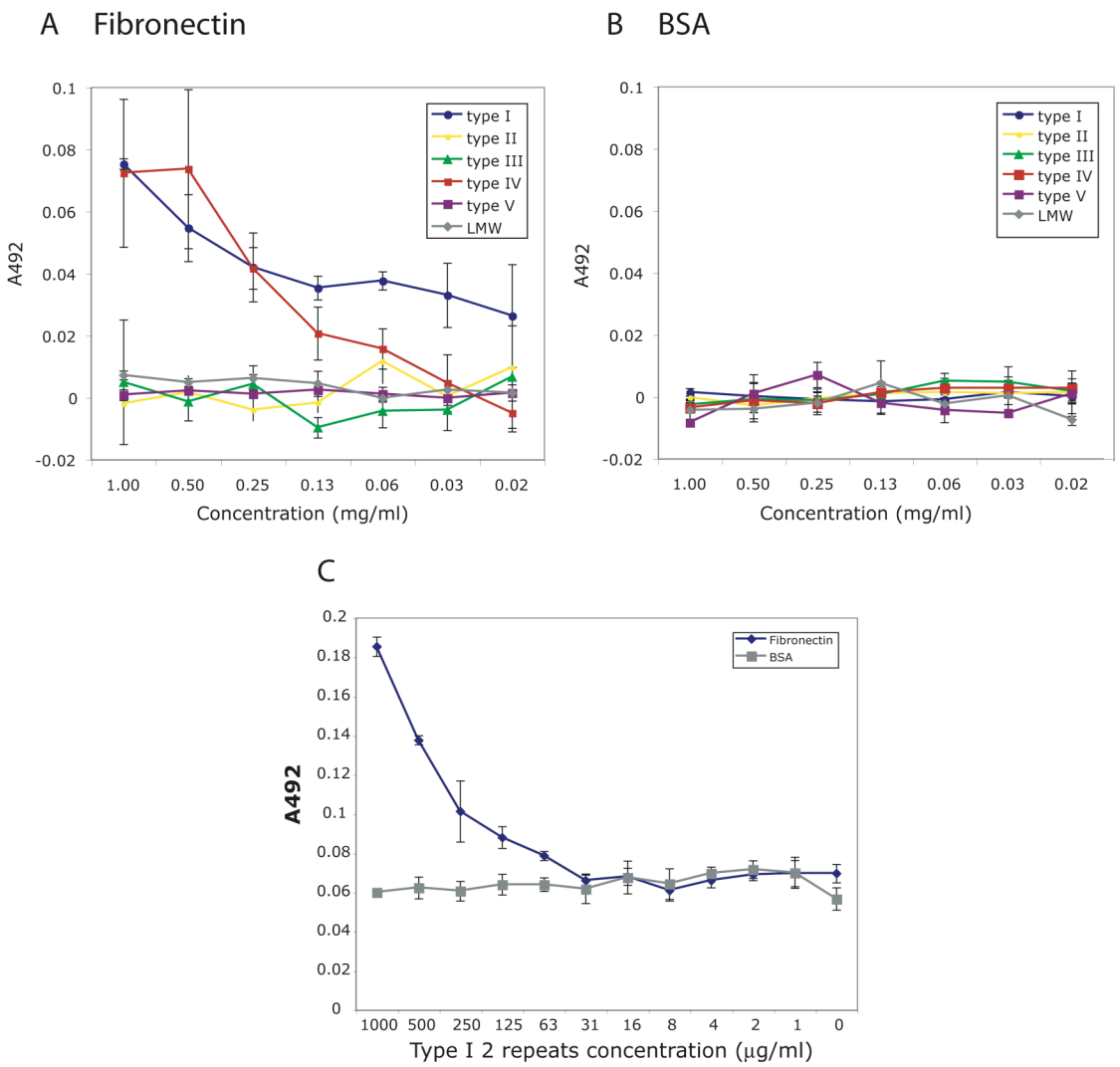


Figure 6.8 Confirmation of interaction between CwpV repeats types I and IV with fibronectin.

A. CwpV repeats ELISA assaying binding to immobilized plasma fibronectin. Concentrations of CwpV repeats were assayed in triplicate. Mean values are plotted with standard deviation error bars. **B.** CwpV repeats ELISA assaying binding to immobilized BSA. Concentrations of CwpV repeats were assayed in triplicate. Mean values are plotted with standard deviation error bars. **C.** Direct comparison of CwpV type I repeats binding to fibronectin and BSA. Concentrations of CwpV repeats were assayed with six replicates. Mean values are plotted with standard deviation error bars.

Fibronectin purified from plasma or from cells is commercially available. These two sources of fibronectin differ in size, with cellular fibronectin being larger due to the presence of additional polypeptide segments, accounted for by alternative splicing of fibronectin mRNA (Kornblihtt et al., 1985). Thus far CwpV binding had only been assessed using plasma fibronectin, therefore levels of binding to

plasma fibronectin and cellular fibronectin were directly compared. Purified CwpV repeats were assayed for cellular and plasma fibronectin binding activity by ELISA and the results are shown in Figure 6.9A and B respectively. Binding activity is exhibited for type I and type IV repeats with both forms of fibronectin, but there is increased binding activity with cellular fibronectin. Therefore for all future assays cellular fibronectin was chosen for use.

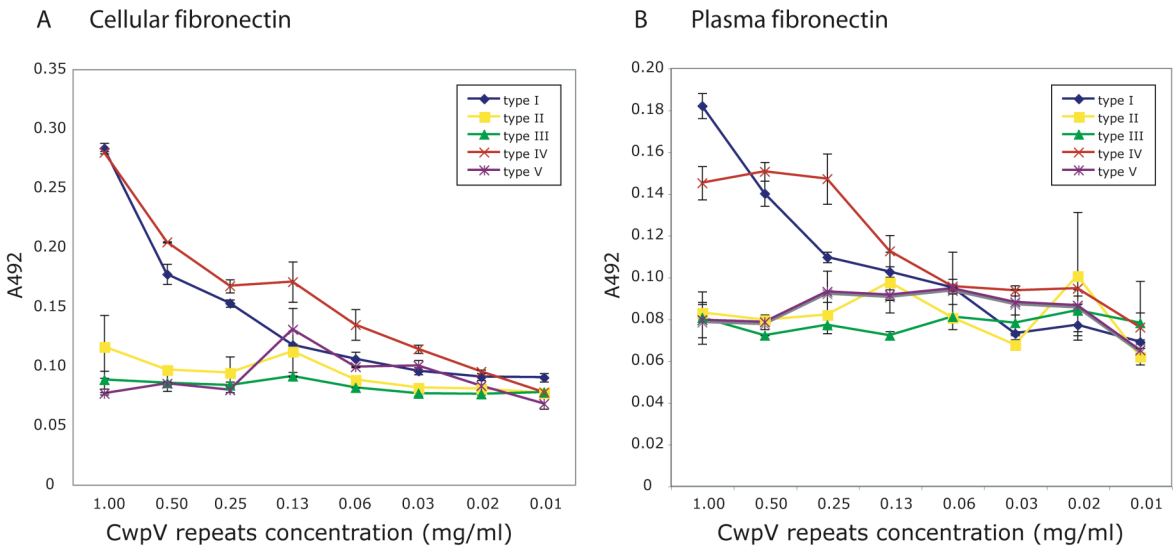


Figure 6.9 Comparison of cellular and plasma fibronectin interactions with CwpV repeats.

A. CwpV repeats ELISA assaying binding to immobilized cellular fibronectin. **B.** CwpV repeats ELISA assaying binding to immobilized plasma fibronectin. Concentrations of CwpV repeats were assayed in triplicate. Mean values are plotted with standard deviation error bars.

In order to assess the avidity effect due to increasing numbers of CwpV repeats a recombinant protein containing nine type I repeats was expressed and purified, and its binding affinity for cellular fibronectin compared to protein containing only two type I repeats (used in previous fibronectin-binding experiments). The two different recombinant proteins are shown diagrammatically in Figure 6.10A. An ELISA was carried out to assess the relative binding activity of the two proteins. As can be seen in Figure 6.10B increasing the number of CwpV type I repeats does increase the binding activity. However, there is only ~ 2-fold increase in the

signal detected by ELISA, reflecting a modest increase in binding activity with the increase from two repeats to nine repeats.

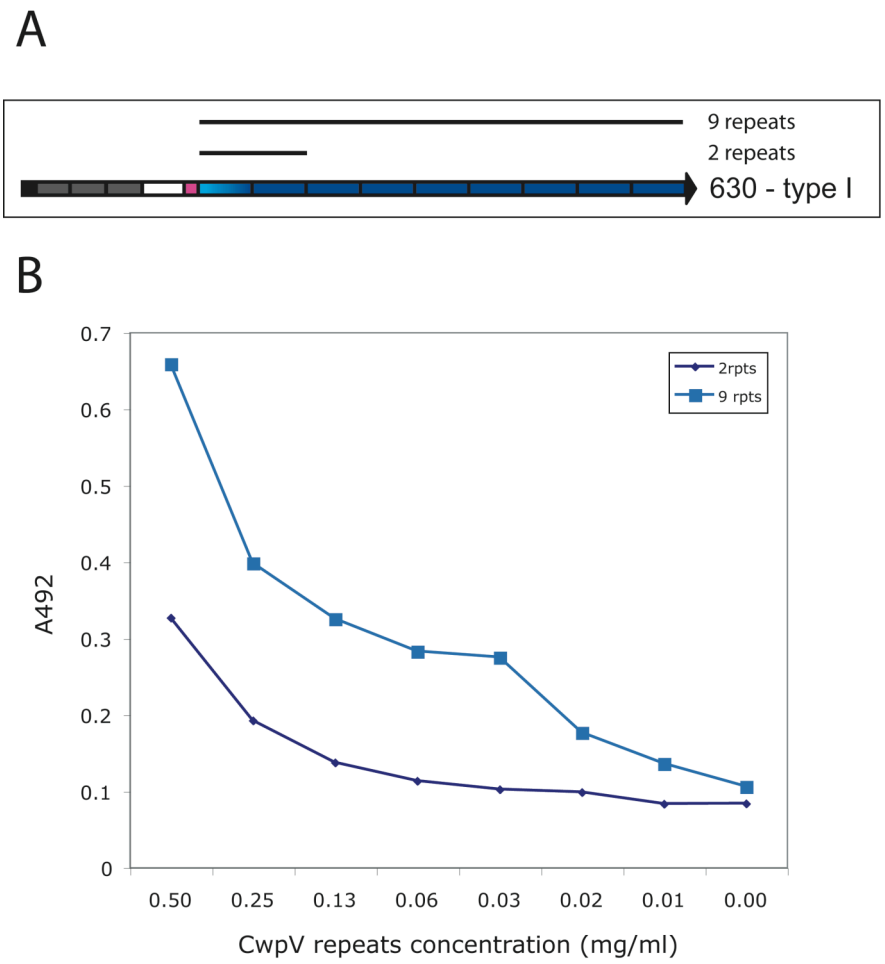


Figure 6.10 Cellular fibronectin binding activity of 2 and 9 CwpV type I repeat domains.

A. Diagrammatic representation of the two recombinant proteins compared, containing 2 and 9 CwpV type I repeat domains. **B.** ELISA assaying binding of recombinant protein to immobilized cellular fibronectin.

Having assayed fibronectin binding activity of purified recombinant protein containing CwpV repeats, the cellular fibronectin binding activity of *C. difficile* cells expressing high levels of CwpV on the surface was investigated. *C. difficile* overnight cultures were harvested by centrifugation then resuspended to an OD_{600} of 10. A two-fold dilution series of the bacterial suspension was carried out in PBS. Samples of bacterial suspensions of the appropriate OD_{600} were then added

to fibronectin-coated wells and incubated for 1 h at 37ºC. Following incubation wells were washed and bound cells detected using anti-HMW SLP and anti-LMW SLP antibodies. The results of an assay using WT, Δ*cwpV*, Δ*cwpV*(pOENter) and Δ*cwpV*(pOEI) are shown in Figure 6.11A. It can be seen that the cells from Δ*cwpV* (pOEI), which express high levels of CwpV, exhibit increased cellular fibronectin binding activity compared to WT, Δ*cwpV* and Δ*cwpV*(pOENter) cells. This suggests that the CwpV repeats expressed on the surface of *C. difficile* cells confer cellular fibronectin binding activity. *C. difficile* strains expressing the different types of CwpV repeats were then assessed for cellular fibronectin binding activity. The results of an assay using Δ*cwpV* and Δ*cwpV*(pOEI-V) *C. difficile* cells is shown in Figure 6.11B. It can be seen that all CwpV types appear to confer some cellular fibronectin binding activity. However, the binding activity conferred by CwpV over-expression was less than two-fold that exhibited by Δ*cwpV* suggesting that CwpV does not confer a high affinity interaction between *C. difficile* cells and immobilized fibronectin. A number of different variations to this assay format were tried, including shorter and longer incubations with cells, cross-linking of cells to the plate after incubation, and increasing the bacterial densities used. However, none of these assays increased the sensitivity of the assay without specificity being compromised. Thus a greater difference between the binding activities of Δ*cwpV* and Δ*cwpV*(pOEI-V) than shown in Figure 6.11 was not observed (data not shown). Therefore, it appears that CwpV expression on the *C. difficile* surface confers a low affinity fibronectin binding activity.

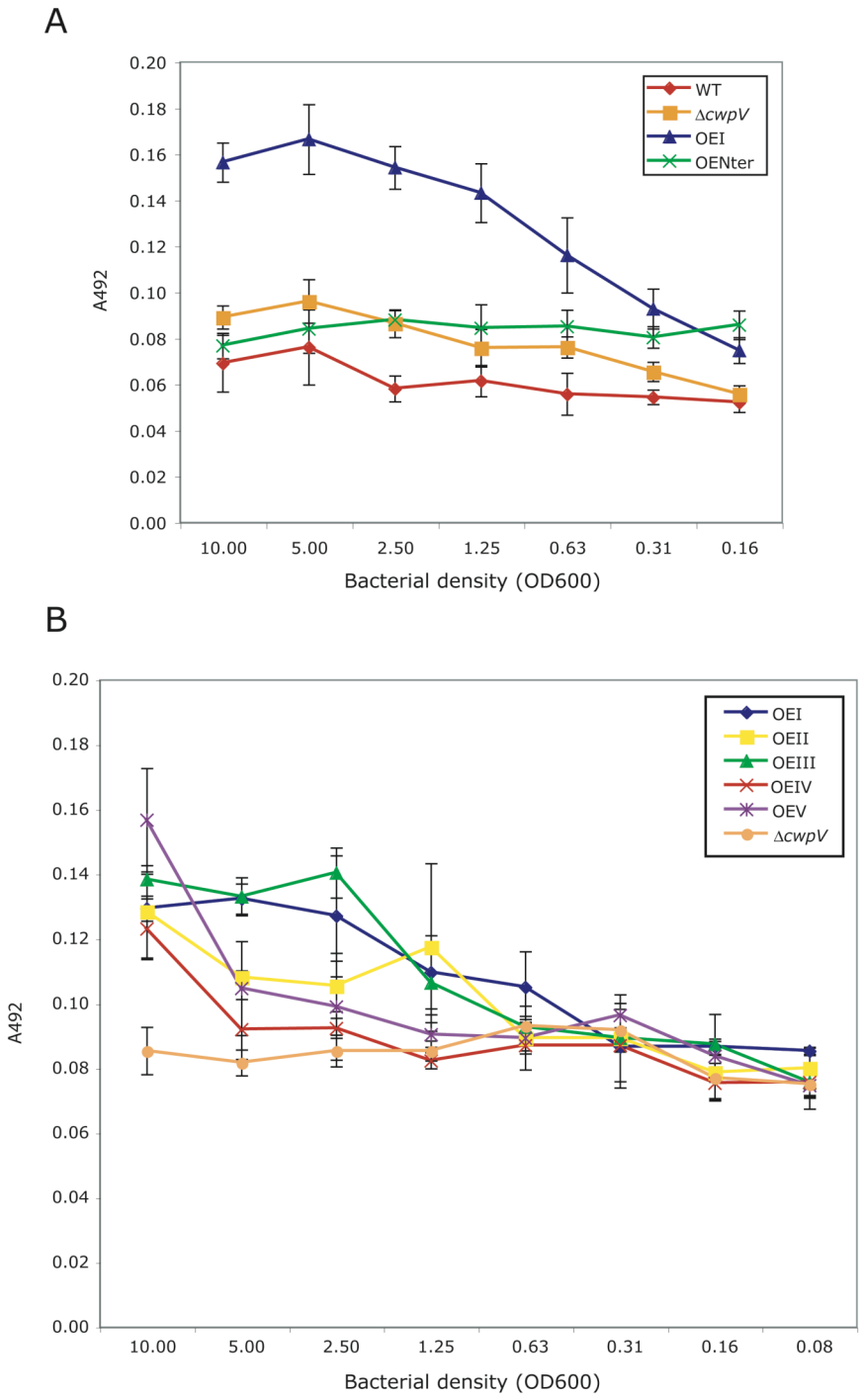


Figure 6.11 *C. difficile* binding to cellular fibronectin.

A. WT, Δ*cwpV*, Δ*cwpV*(pOENter) and Δ*cwpV*(pOEI) bacteria were harvested by centrifugation then resuspended to an OD_{600} of 10. A two-fold dilution series of the bacterial suspension was carried out in PBS. 50 µl samples of bacterial suspensions of the appropriate OD_{600} were then added to fibronectin-coated wells and incubated for 1 h at 37ºC. Following incubation wells were washed, then anti-HMW SLP (1:15,000 dilution) and anti-LMW SLP (1:30,000 dilution) antibodies, followed by anti-rabbit HRP and colorimetric HRP substrate, were used to detect bacteria bound to the plate. **B.** Δ*cwpV* and Δ*cwpV*(pOEI-V) *C. difficile* cells were harvested by centrifugation and assayed for cellular fibronectin binding as described for A.

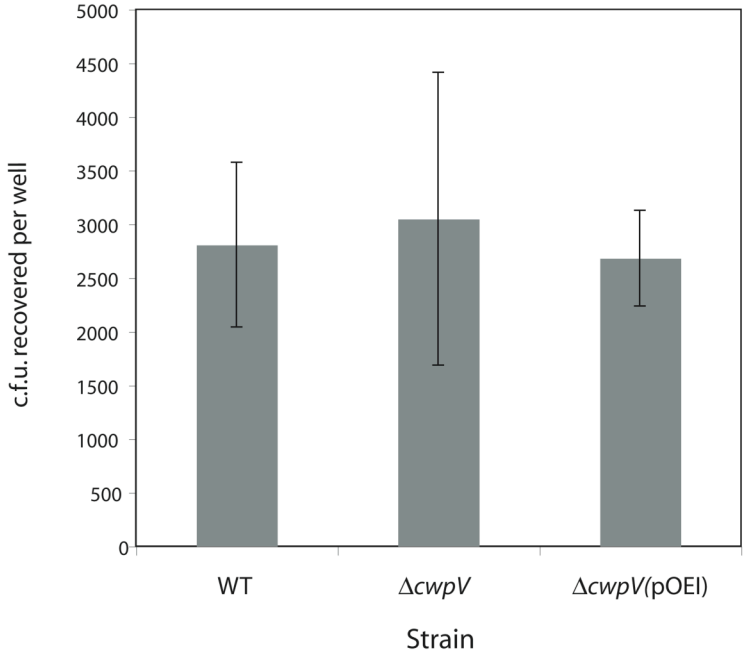
6.2.7 CwpV (type I) expression does not affect *C. difficile* binding to Caco-2 BBE1 cells

Having found that CwpV confers fibronectin binding activity it was postulated that CwpV may facilitate adhesion of *C. difficile* to host cells. *C. difficile* binding to cells *in vitro* has been reported in a number of studies, with Caco-2 cells (Cerquetti *et al.*, 2002, Drudy *et al.*, 2001, Eveillard *et al.*, 1993, Dingle *et al.*, 2010, Schwan et al., 2009), Vero cells (Waligora et al., 1999), HT-29 cells (Eveillard et al., 1993, Drudy et al., 2001) and primary cells isolated from biopsies (Drudy et al., 2001) having been used. The Caco-2 cell line is derived from human epithelial colorectal adenocarcinoma cells, and differentiates *in vitro* into a polarized monolayer mimicking the *in vivo* morphology and functionality of the intestinal epithelium (Hidalgo et al., 1989). The most recent *C. difficile in vitro* adhesion studies have employed Caco-2 cells (Schwan et al., 2009, Dingle et al., 2010), which are known to express fibronectin (Levy et al., 1994). It was therefore decided to test the effect of CwpV on *C. difficile* Caco-2 cell adhesion.

In Figure 6.12A it can be seen that CwpV expression does not affect *C. difficile* binding to Caco-2 cells as WT, Δ*cwpV* and Δ*cwpV*(pOEI) bacteria are all recovered to the same level. To validate the numbers of bacteria found to be adherent, the assay was repeated and the cells fixed prior to analysis by immunofluorescent microscopy. Low levels of adhesion were observed, with approximately one bacterium present per 200 Caco-2 cells. An example of a field of view showing two individual *C. difficile* cells adhered to the Caco-2 monolayer is shown in Figure 6.12B. This level of adhesion observed corresponds well to the number of bacteria determined by viable counts: It was determined that there are $5x10^5$ Caco-2 cells per well of a 12-well plate (data not shown), therefore binding

of 1 bacterium per 200 cells would mean a total of 2500 counts per well, which is the approximate number of counts recovered per well as shown in Figure 6.12A. This level of adhesion observed can be compared to a recently published paper describing a Caco-2 adhesion assay where levels of adhesion are measured by qPCR (Dingle et al., 2010). In this assay an MOI of 1:1 is used, which for a 12-well plate would equate to 5 x 10^5 bacterium added per well. Adhesion is reported to be ~ 0.7% of the total innoculum, which would equate to 3500 counts per well. The adhesion assay we used was adapted from (Schwan et al., 2009), and the levels we observed are comparable to the levels observed in this study (C. Schwan, personal communication). Therefore the levels of adhesion observed in my assay are comparable to those observed in published studies. A number of variations to the Caco-2 binding assay were carried out, including wounding the monolayer prior to incubation, use of different incubation media, longer and shorted incubation times, but no significant differences in adhesion level were observed (data not shown). As no difference in Caco-2 adhesion was conferred by type I CwpV, the effect of expression of the other types of CwpV on adhesion was not tested. It remains possible that these other CwpV types affect *C. difficile* adhesion to Caco-2 cells.

A



B

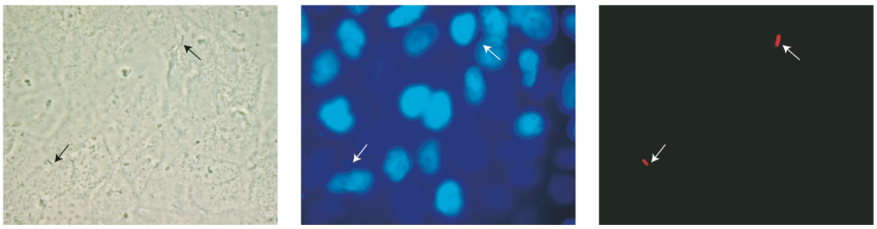


Figure 6.12 Effect of CwpV expression on *C. difficile* binding to Caco-2 BBE1 cells.

A. WT, Δ*cwpV* and Δ*cwpV*(pOEI) bacteria were incubated with Caco-2 BBE1 cells and the number of adherent bacteria determined by plating onto BHI plates and determining the *C. difficile* colony forming units (c.f.u.) recovered per well. **B.** Visualization of WT *C. difficile* adherence to Caco-2 cells. Images shown from left to right: Phase contrast image, Hoechst DNA staining (bacterial cells and Caco-2 cell nuclei) and rhodamine-red stained *C. difficile* cells. Locations of bacterial cells are indicated by arrows.

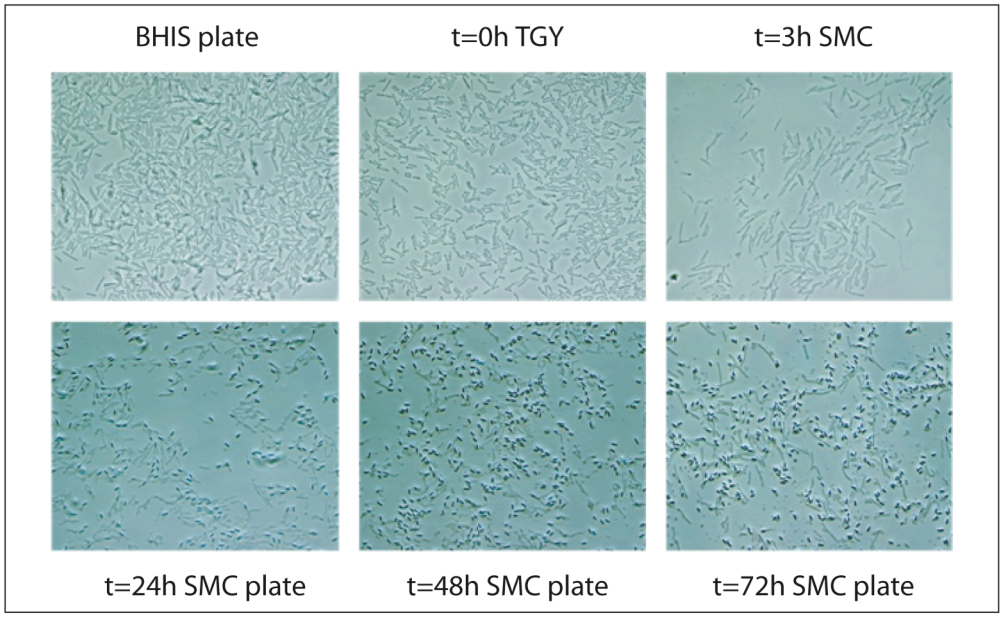
6.2.8 CwpV expression and *C. difficile* sporulation

Sporulation is a key stage in the *C. difficile* life cycle and comparative genomics to the well-characterised sporulation process of *Bacillus subtilis* and other Clostridial sp. has allowed identification of putative factors involved in *C. difficile* sporulation (Paredes et al., 2005). Sporulation has not been studied intensively in *C. difficile*, but links between sporulation and toxin production have been shown (Underwood et al., 2009). This suggests interlinking of sporulation and virulence signaling networks, which has been observed in other spore-forming bacterial species (Perego & Hoch, 2008). Given the importance of sporulation in the *C. difficile* life cycle, and links between regulatory networks governing sporulation and other pathways, it seemed possible that CwpV expression may be regulated during sporulation. The spore proteome has recently been described (Lawley et al., 2009b) and the data showed CwpV is present in the spore. One main question to address was whether the proportion of cells expressing CwpV alters during sporulation. In order to do this sporulation was induced using a protocol kindly provided by Simon Cutting at Royal Holloway London, an expert in *Bacillus subtilis* sporulation. Six timepoints from the sporulation protocol are documented in Figure 6.13A by phase microscopy images. The formation of spores (dark blue under our microscopy conditions) can be seen to occur once on the SMC plates, with many spores present after 24 h and maximum numbers of spores after 48 h, with no increase in spore numbers after 72 h. These observations were validated by counting numbers of heat resistant colony forming units present over time (data not shown). Samples from four time points during the sporulation protocol were taken and cells assayed for CwpV expression by staining with anti-CwpVrptI and anti-rabbit rhodamine red-X followed by immunofluorescent microscopy. As can be seen from Figure 6.13B CwpV is expressed in ~ 5% of cells at all stages of

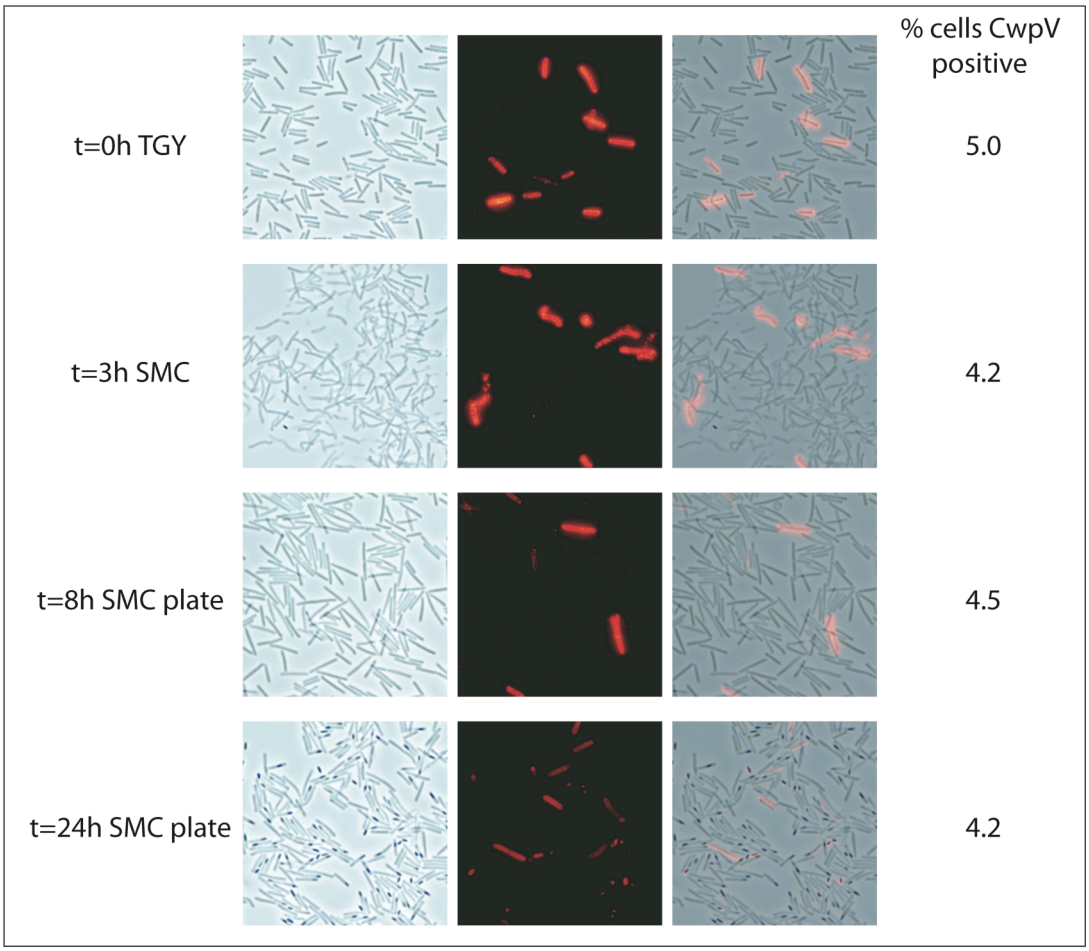
the sporulation process. This suggests that CwpV expression is not specifically turned ON or OFF during sporulation. Some staining of spores was observed using anti-CwpVrpt1, however this was shown to be non-specific as the pre-immune serum also stained spores and this staining was observed in Δ*cwpV* (data not shown).

It is possible that CwpV and/or RecV expression level affects *C. difficile* sporulation directly. Therefore sporulation efficiency in the available mutant strains (WT vs. Δ*cwpV,* Δ*recVON* and Δ*recVOFF*) was investigated. Spores were produced using the same protocol and once harvested from the SMC plates they were washed and purified by density centrifugation. This purification method removed the vast majority of vegetative cell debris, as can be observed in the phase contrast images shown in Figure 6.13C. The percentage of material recovered by density centrifugation as determined by OD_{600} was used as a measure of sporulation efficiency. In Figure 6.13D it can be seen that there is no significant difference in sporulation efficiency between WT, Δ*cwpV,* Δ*recVON* and Δ*recVOFF,* with approximately 70% of material recovered by density centrifugation for all strains. This suggests that CwpV and RecV do not play key roles in the sporulation process.

A


BHIS plate
t=0h TGY
t=3h SMC
t=24h SMC plate
t=48h SMC plate
t=72h SMC plate

B


% cells CwpV positive
t=0h TGY
5.0
t=3h SMC
4.2
t=8h SMC plate
4.5
t=24h SMC plate
4.2

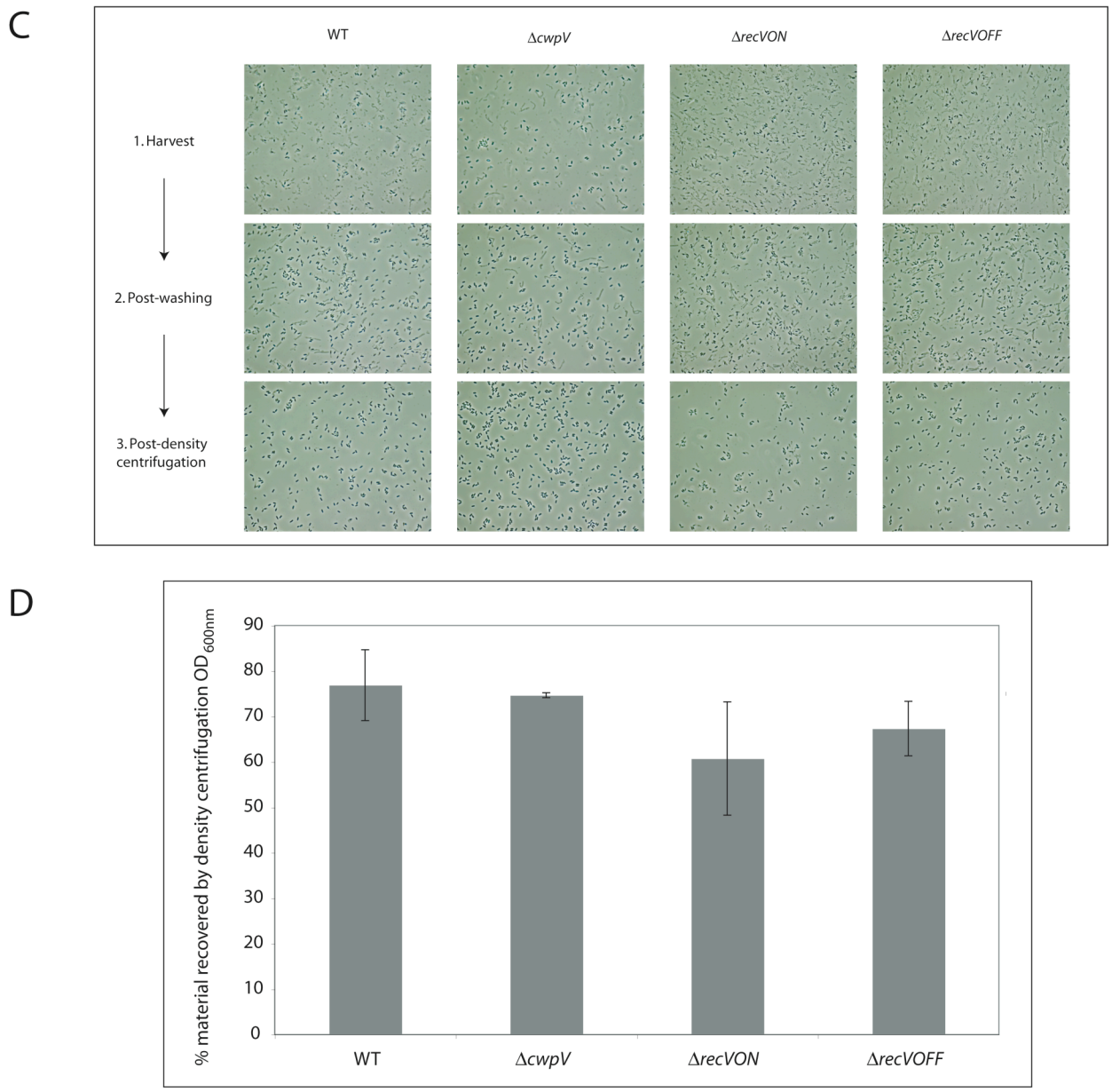


Figure 6.13 CwpV and *C. difficile* sporulation.

A. Phase contrast images of WT *C. difficile* samples taken at timepoints during the sporulation process. Spores can be identified by their dark blue appearance. **B.** Assessment of CwpV expression level in WT *C. difficile* at various sporulation timepoints by immunofluorescent staining. Images shown are (left to right) phase contrast, red fluorescence, overlay. % cells CwpV positive was calculated by counting 5 fields containing a total of at least 500 cells. **C.** Purification of spores by density centrifugation was followed for WT, Δ*cwpV,* Δ*recVON* and Δ*recVOFF* by phase contrast microscopy. **D.** Sporulation efficiency measured as % material recovered by density centrifugation for WT, Δ*cwpV,* Δ*recVON* and Δ*recVOFF.* Values shown are mean values from three independent sporulation preparations and purifications, with standard deviation error bars.

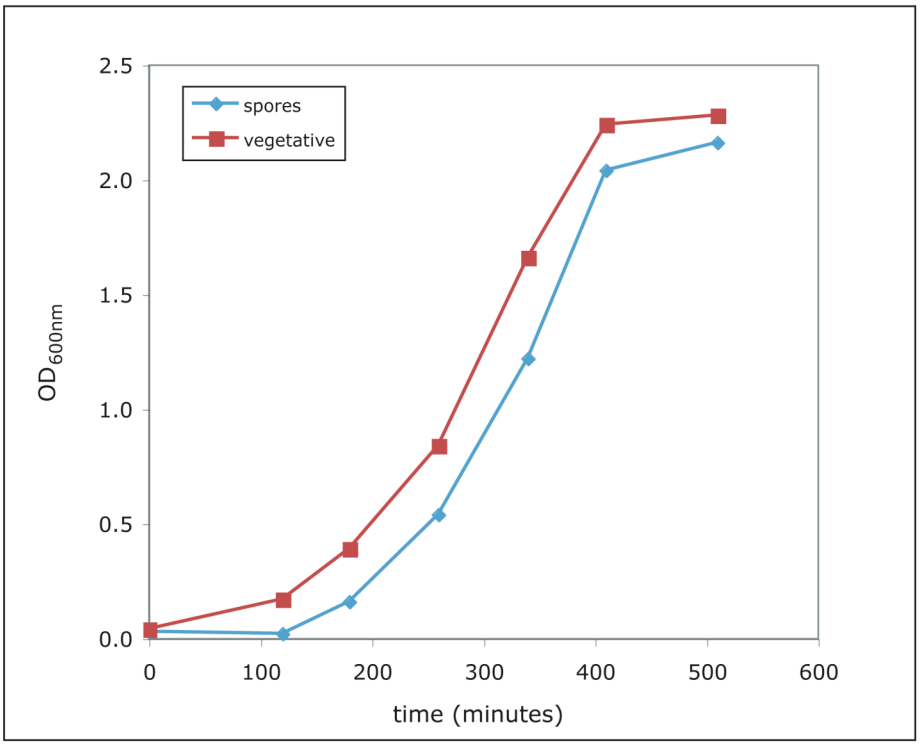
6.2.9 CwpV expression and *C. difficile* germination

Ingestion of *C. difficile* spores leads to germination of spores and subsequent vegetative proliferation of *C. difficile*. This is the necessary first stage of *C. difficile* infection of the enteric system. Bile salts sodium taurocholate and sodium cholate stimulate spore germination (Wilson et al., 1982, Wilson, 1983) and more recently glycine was shown to act as a co-germinant (Wheeldon et al., 2008a, Sorg & Sonenshein, 2008). Given the key importance of germination to the *C. difficile* life cycle, I investigated whether or not CwpV expression was regulated during germination, and if CwpV expression affected germination efficiency. Inoculation of BHIS medium containing the germinant sodium taurocholate (0.1%) with either spores (purified by Histodenz) or vegetative cells allowed determination of the growth kinetics of WT *C. difficile* directly after germination, compared to vegetative cells. As can be seen in Figure 5.14A inoculation with spores follows the same growth kinetics as inoculation with vegetative cells, but growth from spores exhibits a lag of ~ 60 minutes. This germination growth curve was also followed by phase microscopy of samples taken at timepoints during growth from spores, as shown in Figure 5.14B. It can be seen that as the growth progresses the proportion of vegetative cells, but not spores (dark blue) increases. The time-point 410 minutes after inoculation was chosen for assessment of CwpV expression level by immunofluorescence, as at this time-point cells have divided 5-6 times following initital germination. Therefore the majority of material is in vegetative cell form but represents a stage of growth likely to mimic the early colonization phase. In Figure 5.14C it can be seen that at this time-point 5.9% of cells are CwpV-positive. This level of CwpV expression is similar to that seen during normal vegetative cell growth, and therefore suggests that CwpV expression at this point shortly after germination is not switched ON or OFF. Earlier time-points were also

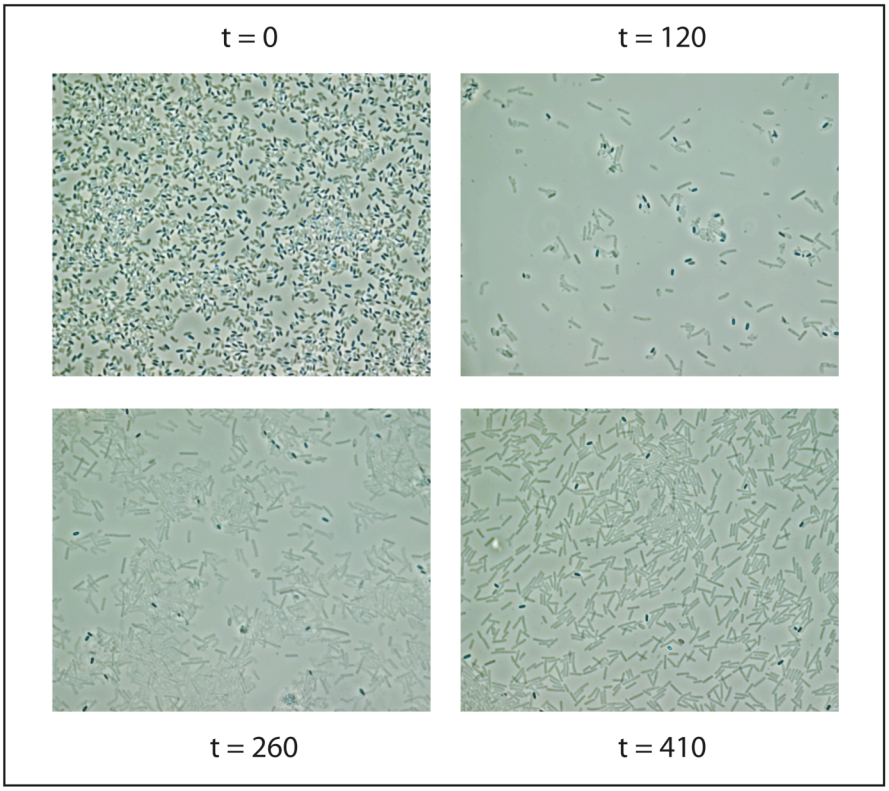
analysed in this way, but were complicated by non-specific binding of the anti-CwpVrptI antibody to spores. It seems unlikely that CwpV expression would recover to 5% after 5-6 cell divisions if its expression was specifically switched ON or OFF during germination, however without further investigation into S-layer protein expression during germination this cannot be conclusively ruled out.

To assess the role of CwpV and/or RecV in sporulation efficiency the number of c.f.u/ml of a purified spore preparation with an OD_{600} of 1.0 was measured for WT, ∆*cwpV,* ∆*recVON* and ∆*recVOFF.* Counts were measured for three independent spore preparations of each strain. As shown in Figure 5.14D all four strains gave values of ~ 1 x 10^8 c.f.u./ml(OD_{600} = 1.0). This suggests that CwpV and RecV expression do not affect germination efficiency *in vitro*.

A


spores
vegetative
OD600nm
2.5
2.0
1.5
1.0
0.5
0.0
0
100
200
300
400
500
600
time (minutes)

B


t = 0
t = 120
t = 260
t = 410

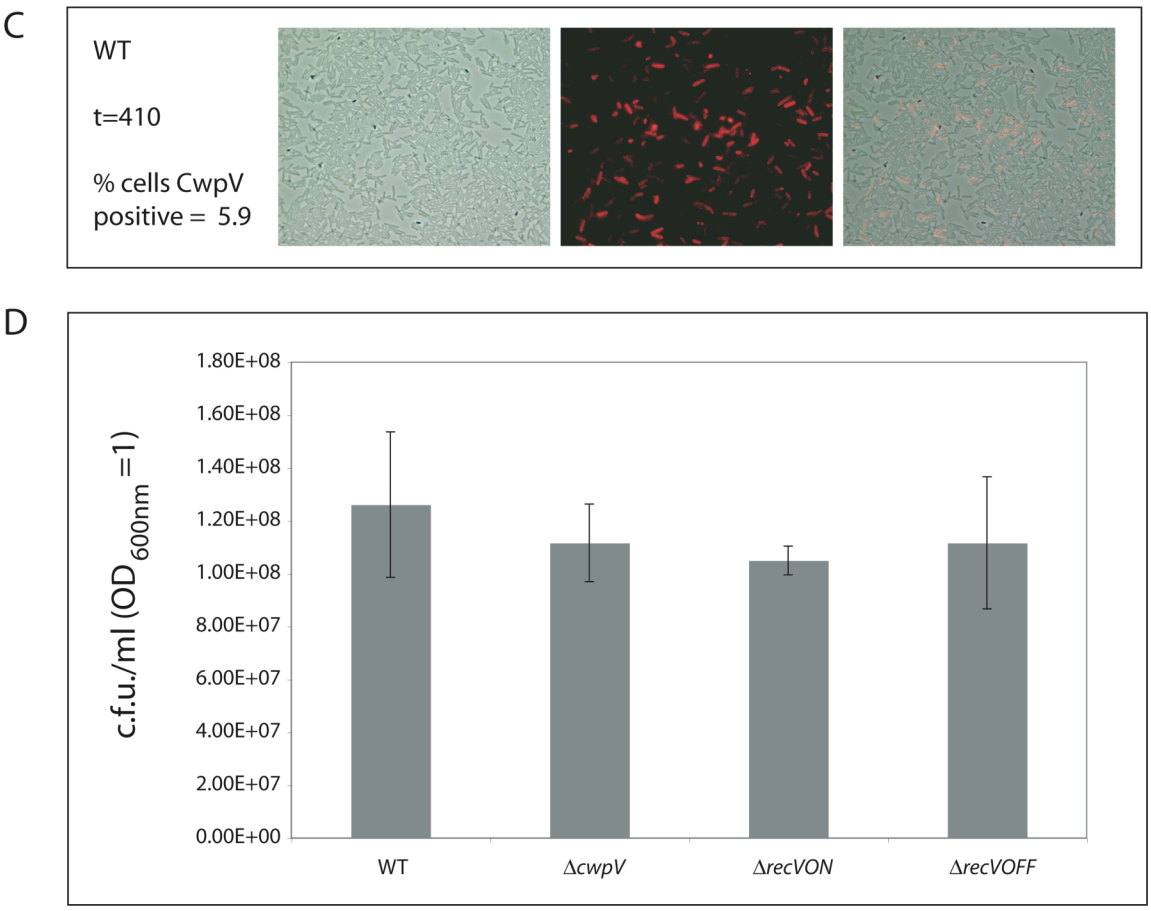


Figure 6.14 CwpV and *C. difficile* spore germination.

A. Comparison of the growth kinetics of WT *C. difficile* cultures in BHIS 0.1% sodium taurocholate inoculated with vegetative cells or spores to a starting OD_{600} of 0.05**. B.** Phase contrast images of samples taken at various timepoints (given in minutes) throughout germination. **C.** Assessment of CwpV expression level in WT *C. difficile* at 410 minutes after germination by immunofluorescent staining. Images shown are (left to right) phase contrast, red fluorescence, overlay. % cells CwpV positive was calculated by counting 5 fields containing a total of at least 500 cells. **D.** Germination efficiency of purified spore preparations of WT, Δ*cwpV,* Δ*recVON* and Δ*recVOFF C. difficile* strains measured as C.f.u./ml(OD_{600}=1) grown on BHIS 0.1% sodium taurocholate plates. Values are means calculated from counts of three independent spore preparations. Error bars represent the standard deviation.

6.3 Discussion

A high level of expression of all known CwpV types was found to alter colony morphology. This effect of CwpV expression is conferred by the repetitive domain, as high level of expression of the CwpV N-terminal CWA domain does not affect colony morphology. This suggests that, despite being entirely distinct at the amino acid sequence level, the different types of CwpV repeat exhibit a functionally conserved activity. Although single species colony growth *in vitro* is somewhat removed from natural growth conditions, *in vitro* colonies are organized populations of bacteria and differences in colony morphology reflect differences in the cellular organization of a population (Shapiro, 1995) that can translate to important differences affecting natural growth dynamics. The difference in the cellular organization of colonies can be seen clearly in the microscopic images in Figure 5.4, with high levels of CwpV expression causing denser more randomly packed cellular organization than the diffuse, directional patterns of cellular packing seen at the edge of WT colonies. Such differences can be explained by an aggregative function for CwpV, but could also be caused by a decrease in cellular motility.

The CwpV-mediated aggregation of *C. difficile* cells from suspension suggests that CwpV does have a direct aggregation-promoting function. All CwpV types, apart from type IV, were shown to cause aggregation from cellular suspensions of *C. difficile.* Propensities of different CwpV types to cause aggregation differed, observed by the different starting densities of cellular suspensions required in order to lead to aggregation. This is likely to reflect different affinities of the different CwpV proteins for their ligand(s). Type IV CwpV does confer the same

colony morphology phenotype as all the other CwpV types, but perhaps the affinity of type IV CwpV for its aggregating ligand(s) is lower than for the other CwpV types, which could explain why aggregation from suspension was not observed for this type of CwpV. The colony morphology phenotype observed for type IV CwpV does however suggest that it carries out the same general function as the other types.

Functional interaction between bacterial autoaggregation and motility has been reported in *E. coli,* with the autoaggregative protein antigen 43 (Ag43) and flagella having opposing functional roles (Ulett et al., 2006), but in the case of *C. difficile* CwpV expression does not appear to affect swimming motility. The diffuse directional organization of WT colony edges seen in this study is reminiscent of twitching motility mediated by type IV pili in *Clostridium perfrigens* (Varga et al., 2006). *C. difficile* does have type IV pili clusters, which are conserved across strains (Janvilisri et al., 2009). Type IV pili may therefore be responsible for the cellular organization of WT colonies grown on solid surfaces observed here. CwpV may interact with motility factors such as type IV pili to alter the cellular organisation of colonies, but as the factors responsible for this motility are yet to be definitively characterised it was not possible for us to investigate direct interactions between CwpV expression and these factors.

Identification of the CwpV ligand(s) responsible for its aggregative function is a future research priority. CwpV may interact with itself or with other *C. difficile* surface components. Whether or not different CwpV types recognize the same ligand(s) is another interesting question to be investigated. Fluorescent labeling of *C. difficile* expressing CwpV would be a useful tool, as this would allow cells with

and without CwpV expression to be mixed, and aggregates analysed for levels of CwpV expression. If CwpV interacts with itself you would expect aggregates to contain a majority of cells expressing CwpV, which could be identified by their fluorescence. As discussed in Chapter 4, there is a paucity of fluorescent reporters available for use under anaerobic conditions, and the FbFP protein (Drepper et al., 2007a) does not exhibit high sensitivity compared to background fluorescence of *C. difficile*. However, one possibility for the future is the use of FlAsH tags in proteins of interest, which can be irreversibly fluorescently labelled under aerobic or anaerobic conditions (Machleidt et al., 2007), and would allow identification of sub-populations of *C. difficile* that did, or did not, express the FlAsH tag, which could be engineered to correlate with CwpV expression. Alternatively, his-tagged purified CwpV repeats could be incubated with *C. difficile* cells that do, or do not, express CwpV to determine if they associate with the *C. difficile* cells, and if so whether this is dependent on CwpV expression. A similar experiment was used to demonstrate that the *Staphylococcus aureus* autoaggregative protein SasG binds to itself rather than any other surface component, mediating its biofilm-promoting function (Geoghegan et al., 2010). The effect of varying numbers of CwpV repeats on the affinity of the interaction can be investigated once the ligand has been identified.

Aggregation-promoting proteins have been described in a number of other bacterial species. In species where biofilm assays are available, autoaggregative proteins have often been shown to play roles in biofilm formation (Kuboniwa *et al.*, 2009, Corrigan *et al.*, 2007, Huang *et al.*, 2009). There have been no studies on *C. difficile* biofilms reported, and in our experience *C. difficile* does not exhibit significant biofilm formation on abiotic surfaces (data not shown). However, in *C.*

difficile-infected mice large aggregates of *C. difficile* cells, described as exaggerated mats, were reported to be associated with regions of severe inflammation (Lawley et al., 2009a). This description is indicative of a biofilm-like growth of *C. difficile* during infection. Multispecies biofilms in the human GI tract have been observed (Macfarlane & Dillon, 2007) and *C. difficile* is likely to require factors that promote appropriate intra- and inter-species interactions to colonise the gut. CwpV may play important roles in this process. An *in vitro* biofilm model would be useful for assessing the function of CwpV and other cell surface proteins in *C. difficile* growth on surfaces.

Aside from its aggregation-promoting function CwpV may have other functions. The fibronectin binding activity observed was of low affinity. Also, an increase from 2 to 9 repeats in purified protein did not lead to a significant increase in the binding affinity detected. If the repeat domains mediated a specific interaction with fibronectin, a more marked increase in affinity upon such an increase in the number of domains would be expected. It is therefore unclear as to the relevance of this interaction detected. No significant difference in adhesion to Caco-2 cells caused by CwpV expression was observed, which does not support an adhesive role for CwpV. This adhesion assay does rely upon low levels of *C. difficile* adhesion, and as yet has not been directly compared to adhesion levels of *C. difficile* to primary tissue or cells. This would give an indication as to how physiologically relevant such adhesion assays are, and therefore how well-suited they are to assessment of adhesive functions of surface proteins. CwpV was annotated as a putative hemagglutinin/adhesin in the 630 genome, reflecting its similarity to proteins known to have such functions. Purified CwpV type I repeats and WT, Δ*cwpV,* Δ*cwpV*(pOEI) *C. difficile* strains were tested for hemagglutination

activity but no such activity was observed, suggesting that type I CwpV does not act as a hemagglutinin (data not shown). An obvious important question for future work is whether CwpV expression affects colonization level *in vivo*, and if so whether this is due to its aggregation-promoting function or whether it does confer any other properties to *C. difficile*.

A variety of *in vitro* conditions have been investigated for their effect on CwpV expression, including addition of salts, taurocholic acid, glucose and serum (data not shown), as well as sporulation-inducing conditions and germination of spores. However, none yet have significantly affected CwpV expression level, with reproducible levels of ~5% of cells being CwpV-positive under all conditions. Therefore it is not known whether there are any environmental signals that affect CwpV expression levels. Given the relatively constant CwpV expression level observed in strain 630, it appears that under all conditions tested *cwpV* DNA switch inversion occurs at high frequency leading to an equilibrium of switch orientation, conferring the stable average CwpV expression level seen in large populations of cells such as colonies or cultures. A future research priority is the investigation of CwpV expression levels during infection *in vivo*. After infection of mice or hamsters, sections of gut tissue could be analysed for the presence of CwpV-positive *C. difficile* cells present using antibody staining and appropriate microscopic techniques.

Understanding the roles of CwpV during *in vivo* infection is a future research priority and will be addressed using the isogenic mutants described in this study. The Δ*recVON* and Δ*recVOFF* strains constructed and characterised in this study are ideal for assessing the effects of CwpV expression level on colonization, as

they are genetically identical aside from the orientation of the *cwpV* switch, determining fixed high or absence of CwpV expression respectively. It is of course possible that both these mutants are less well suited to host colonization than the WT strain with its ability to switch CwpV expression ON or OFF. Established animal models of infection focus more on the acute stage of *C. difficile* infection, rather than colonization (Douce & Goulding, 2010). However, recent developments of animal models better suited to studying the roles of *C. difficile* genes likely to be involved in colonization, such as CwpV, have been described (Lawley et al., 2009a, Steele et al., 2010) that will facilitate future investigations.

7 Discussion

C. difficile is a bacterial pathogen, which causes significant morbidity and mortality worldwide. There are shortcomings in the current therapeutic options available for treatment of *C. difficile* and there is no vaccine available for prevention of CDAD. Rational design of novel therapeutics and vaccines is limited by our lack of understanding of the molecular mechanisms underpinning colonisation of the host by *C. difficile.* This study has focused on the cell surface protein CwpV. A number of interesting properties and characteristics of CwpV have been determined, which shed light on the roles this protein may play during infection. These findings have generated new questions, both with regard to CwpV itself and to the molecular basis of *C. difficile* pathogenesis in general. A summary of the most important findings of this study, the questions that these findings raise and some suggestions as to how to address them are discussed in this final chapter.

7.1 CwpV phase variation

Discovery of the phase variable expression of CwpV is the first reported discovery of phase variable regulation of protein expression in *C. difficile.* It is also the first report of phase variation mediated by site-specific recombination in any gram-positive bacterial species (Wisniewski-Dye & Vial, 2008). It therefore represents an important finding, indicating that mechanisms of phase variation previously characterised in other bacterial species are also utilized by *C. difficile*. It is possible that other phase variable proteins will be discovered in *C. difficile,* and

the tools for studying phase variation mechanisms developed in this study would be useful for future studies.

The formation of a transcription-terminating stem loop at the boundary of the *cwpV* switch region in the OFF orientation, which is disrupted upon *cwpV* switch inversion to ON, confers the ON and OFF CwpV expression phenotypes. This represents a novel and elegant mechanism of phase variable regulation of protein expression. There are, however, still many aspects of the *cwpV* switch that merit further study. For example, determination of the affinity of the recombinase RecV for the two orientations of the *cwpV* switch would inform us how the relative proportions of the ON and OFF orientations of the *cwpV* switch are established and maintained. This has relevance to how RecV expression level could control the proportion of cells expressing CwpV. No conditions have yet been identified in which CwpV expression is specifically switched towards ON or OFF. A research priority is to study the proportion of CwpV-positive cells present during infection *in vivo*. This could be done by fluorescent microscopy analysis of gut tissue from patients or animals colonized by *C. difficile* stained with the relevant antibodies. If it does appear that CwpV is switched ON or OFF *in vivo*, the environmental signals responsible for this regulation can be investigated *in vitro*. Such work may lead to an improvement in the understanding of *C. difficile* regulatory networks, which are key to the *C. difficile* life-cycle.

7.2 *C. difficile* strain diversity

The diversity of CwpV C-terminal repeat sequences reflects the considerable sequence diversity found across *C. difficile* strains, as assessed by microarray

studies (Marsden et al., 2010, Stabler et al., 2006b, Janvilisri et al., 2009) and genomic sequencing (He et al., 2010). The five distinct repeat types identified so far have been discovered from a relatively limited survey of strains, and it is therefore likely that further repeat types, and/or mosaics of types exist. It will be interesting to determine the *cwpV* sequence of more strains and, if further CwpV types exist, their antigenicity, post-translational processing and aggregation-promoting activity should be tested to see if they fit the same patterns of conservation and variation as has been observed so far.

Recently, there has been much emphasis on the characterisation of so-called hypervirulent strains of *C. difficile,* such as the ribotype 027 strains, that have caused a number of nosocomial outbreaks of infection worldwide and appeared to be associated with increased incidence of infection (Stabler et al., 2009, Cartman *et al.*, 2010). However, the prevalence of different *C. difficile* strain types causing disease is dynamic, both temporally and geographically (Cheknis et al., 2009, Brazier et al., 2008). It is clear that common virulent strains have arisen from multiple lineages, and therefore research should not focus too closely on any one strain type or lineage, as virulent strains emerge independently from across the diversity of *C. difficile* strains (He et al., 2010). This is of particular importance for vaccine and antimicrobial development, where understanding of the total species genome or the pan-genome, is highly desirable in order to ensure that any vaccine developed has suitable strain coverage (Muzzi et al., 2007). Future genome sequencing projects will dramatically increase our understanding of the full diversity of *C. difficile* and of the phylogenetic relationships between strains. As already discussed this will be of general use to the selection of well-conserved vaccine targets, but will also allow us to understand better how CwpV has

evolved, ascertaining when and how many times different CwpV repeat types have been acquired. This may shed light onto the mechanisms involved in the acquisition and the functional importance of this variability.

7.3 CwpV and the *C. difficile* S-layer

This study represents the first detailed molecular study of a CWP family member via expression of mutated proteins in *C. difficile*. The cell wall anchoring function of the PF04122 domains has been demonstrated experimentally for the first time. The tools developed here to express CwpV from a plasmid copy amenable to mutagenesis, could be used in future to further characterize the mechanism of cell wall anchoring employed by this family of proteins. As this mechanism is likely to be conserved across the whole CWP family, including SlpA, it could represent a useful target for therapeutic intervention.

It is clear from this study that CwpV is post-translationally processed in a manner analogous to SlpA. High levels of CwpV can be tolerated in the *C. difficile* S-layer, with CwpV constituting a major component of the total S-layer protein in Δ*cwpVON* cells. It seems likely that this family of CWPs has evolved to allow *C. difficile* to have an S-layer of dynamic composition, allowing its functionalities to vary in response to external conditions. In order for the S-layer to maintain its integrity it seems likely that all members of this family contribute interactions to the two-dimensional S-layer packing. Interactions between these CWP proteins therefore merit further investigation. Plasmids encoding tagged CWPs for expression in *C. difficile*, as developed in this study, would be valuable tools for such investigations.

7.4 CwpV and *C. difficile* pathogenicity

Given the aggregation-promoting function of CwpV it is possible that its function may be linked to motility and/or biofilm production. These two general areas of *C. difficile* biology, particularly during *in vivo* infection, are currently poorly understood. It is clear that *C. difficile* does exhibit motility in liquid (Twine et al., 2009) and on solid surfaces *in vitro*. Exaggerated mats of *C. difficile* cells found on tissue from infected mice have been reported (Lawley et al., 2009a), and are suggestive of biofilm-like growth. Therefore these areas should be considered for future research prioritization, as they may be key to the *C. difficile* colonisation process.

In order to determine the role of CwpV in *C. difficile* pathogenesis, experiments using animal models of infection are likely to provide the most definitive results. The mutants created and characterised in this study (Δ*cwpV,* Δ*recVON* and Δ*recVOFF)* are ideal for carrying out animal experiments to determine the role of CwpV during infection. The importance of such experiments using isogenic *C. difficile* mutants has been demonstrated by recent studies on toxins A and B (Lyras et al., 2009, Marsden et al., 2010), which have called into question the findings of earlier experiments addressing the contribution of the two toxins to virulence. Recent developments of animal models (Lawley et al., 2009a, Steele et al., 2010), more suitable for studying colonisation than established models of acute toxin-mediated disease (Goulding et al., 2009), are likely to be most useful for assessing the function of surface molecules such as CwpV.

The tools used to genetically modify *C. difficile* for this study (Heap et al., 2010) demonstrate that the *C. difficile* research field is now well-positioned to create

isogenic mutants in putative colonisation factors and test these in animal models of infection. Such research is necessary to significantly improve our understanding of the factors key to colonisation of the host.

7.5 Overall conclusions

Understanding the common features of *C. difficile* as an ancient species is important, as in order for intervention strategies to be successful in the long-term they must target the full diversity of *C. difficile* strains, rather than specific lineages. This study of CwpV has provided a lesson in bacterial variability. CwpV appears to be a core component of *C. difficile*, including its phase variable regulation controlled by RecV, post-translational processing mechanism, and aggregation-promoting function. This suggests the importance of CwpV to *C. difficile* survival throughout the evolution of this diverse species. Despite this conservation CwpV appears to have been under positive selection for antigenic variability, hinting at a complex evolutionary history. Further investigation into the roles of CwpV in the *C. difficile* life cycle will teach us about what makes *C. difficile* a successful pathogen, providing opportunities for new rational intervention strategies.

8 References

Aas, J., C. E. Gessert & J. S. Bakken, (2003) Recurrent *Clostridium difficile* colitis: case series involving 18 patients treated with donor stool administered via a nasogastric tube. *Clin.Infect.Dis.* **36**: 580-585.

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

Abraham, J. M., C. S. Freitag, J. R. Clements & B. I. Eisenstein, (1985) An invertible element of DNA controls phase variation of type 1 fimbriae of *Escherichia coli*. *Pro*C.*Natl.Acad.Sci.USA* **82**: 5724-5727.

al Saif, N. & J. S. Brazier, (1996) The distribution of *Clostridium difficile* in the environment of South Wales. *J.Med.Microbiol.* **45**: 133-137.

Al-Eidan, F. A., J. C. McElnay, M. G. Scott & M. P. Kearney, (2000) *Clostridium difficile*-associated diarrhoea in hospitalised patients. *J Clin Pharm Ther* **25**: 101-109.

Ananthakrishnan, A. N., E. L. McGinley & D. G. Binion, (2008) Excess hospitalisation burden associated with *Clostridium difficile* in patients with inflammatory bowel disease. *Gut* **57**: 205-210.

Anton, P. M., M. O'Brien, E. Kokkotou, B. Eisenstein, A. Michaelis, D. Rothstein, S. Paraschos, C. P. Kelly & C. Pothoulakis, (2004) Rifalazil treats and prevents relapse of *Clostridium difficile*-associated diarrhea in hamsters. *Antimicrob.Agents Chemother.* **48**: 3975-3979.

Antonopoulos, D. A., S. M. Huse, H. G. Morrison, T. M. Schmidt, M. L. Sogin & V. B. Young, (2009) Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. *Infect.Immun.* **77**: 2367-2375.

Ausiello, C. M., M. Cerquetti, G. Fedele, F. Spensieri, R. Palazzo, M. Nasso, S. Frezza & P. Mastrantonio, (2006) Surface layer proteins from *Clostridium difficile* induce inflammatory and regulatory cytokines in human monocytes and dendritic cells. *Microbes Infect* **8**: 2640-2646.

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

Baines, S. D., J. Freeman & M. H. Wilcox, (2005) Effects of piperacillin/tazobactam on *Clostridium difficile* growth and toxin production in a human gut model. *J.Antimicrob.Chemother.* **55**: 974-982.

Baines, S. D., R. O'Connor, J. Freeman, W. N. Fawley, C. Harmanus, P. Mastrantonio, E. J. Kuijper & M. H. Wilcox, (2008) Emergence of reduced susceptibility to metronidazole in *Clostridium difficile*. *J.Antimicrob.Chemother.* **62**: 1046-1052.

Baines, S. D., K. Saxton, J. Freeman & M. H. Wilcox, (2006) Tigecycline does not induce proliferation or cytotoxin production by epidemic *Clostridium difficile* strains in a human gut model. *J.Antimicrob.Chemother.* **58**: 1062-1065.

Bakri, M. M., D. J. Brown, J. P. Butcher & A. D. Sutherland, (2009) *Clostridium difficile* in ready-to-eat salads, Scotland. *Emerg.Infect.Dis.* **15**: 817-818.

Barbut, F., A. Richard, K. Hamadi, V. Chomette, B. Burghoffer & J. C. Petit, (2000) Epidemiology of recurrences or reinfections of *Clostridium difficile*-associated diarrhea. *J.Clin.Microbiol.* **38**: 2386-2388.

Barker, R. H., Jr., R. Dagher, D. M. Davidson & J. K. Marquis, (2006) Review article: tolevamer, a novel toxin-binding polymer: overview of preclinical pharmacology and physicochemical properties. *Aliment.Pharmacol.Ther.* **24**: 1525-1534.

Bartlett, J. G., (1992) Antibiotic-associated diarrhea. *Clin.Infect.Dis.* **15**: 573-581.

Bartlett, J. G., (2009a) *Clostridium difficile* infection: historic review. *Anaerobe* **15**: 227-229.

Bartlett, J. G., (2009b) New antimicrobial agents for patients with *Clostridium difficile* infections. *Curr.Infect.Dis.Rep.* **11**: 21-28.

Bartlett, J. G., T. W. Chang, M. Gurwith, S. L. Gorbach & A. B. Onderdonk, (1978) Antibiotic-associated pseudomembranous colitis due to toxin-producing *Clostridia*. *N.Engl.J.Med.* **298**: 531-534.

Bartlett, J. G. & D. N. Gerding, (2008) Clinical recognition and diagnosis of *Clostridium difficile* infection. *Clin.Infect.Dis.* **46 Suppl 1**: S12-18.

Bartlett, J. G., A. B. Onderdonk & R. L. Cisneros, (1977a) Clindamycin-associated colitis in hamsters: protection with vancomycin. *Gastroenterology* **73**: 772-776.

Bartlett, J. G., A. B. Onderdonk, R. L. Cisneros & D. L. Kasper, (1977b) Clindamycin-associated colitis due to a toxin-producing species of *Clostridium* in hamsters. *J.Infect.Dis.* **136**: 701-705.

Bendtsen, J. D., H. Nielsen, G. von Heijne & S. Brunak, (2004) Improved prediction of signal peptides: SignalP 3.0. *J.Mol.Biol.* **340**: 783-795.

Berry, A. P. & P. N. Levett, (1986) Chronic diarrhoea in dogs associated with *Clostridium difficile* infection. *Vet.Re*C. **118**: 102-103.

Bidet, P., F. Barbut, V. Lalande, B. Burghoffer & J. C. Petit, (1999) Development of a new PCR-ribotyping method for *Clostridium difficile* based on ribosomal RNA gene sequencing. *FEMS Microbiol.Lett.* **175**: 261-266.

Bojesen, A. M., K. E. Olsen & M. F. Bertelsen, (2006) Fatal enterocolitis in Asian elephants (Elephas maximus) caused by *Clostridium difficile*. *Vet.Microbiol.* **116**: 329-335.

Bolton, R. P. & M. A. Culshaw, (1986) Faecal metronidazole concentrations during oral and intravenous therapy for antibiotic associated colitis due to *Clostridium difficile*. *Gut* **27**: 1169-1172.

Bolton, R. P., S. K. Tait, P. R. Dear & M. S. Losowsky, (1984) Asymptomatic neonatal colonisation by *Clostridium difficile*. *Arch.Dis.Child.* **59**: 466-472.

Borriello, S. P. & F. E. Barclay, (1985) Protection of hamsters against *Clostridium difficile* ileocaecitis by prior colonisation with non-pathogenic strains. *J.Med.Microbiol.* **19**: 339-350.

Borriello, S. P., P. Honour, T. Turner & F. Barclay, (1983) Household pets as a potential reservoir for *Clostridium difficile* infection. *J.Clin.Pathol.* **36**: 84-87.

Bouza, E., T. Pelaez, R. Alonso, P. Catalan, P. Munoz & M. R. Creixems, (2001) "Second-look" cytotoxicity: an evaluation of culture plus cytotoxin assay of *Clostridium difficile* isolates in the laboratory diagnosis of CDAD. *J.Hosp.Infect.* **48**: 233-237.

Brannigan, J. A., G. Dodson, H. J. Duggleby, P. C. Moody, J. L. Smith, D. R. Tomchick & A. G. Murzin, (1995) A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* **378**: 416-419.

Braun, V., T. Hundsberger, P. Leukel, M. Sauerborn & C. von Eichel-Streiber, (1996) Definition of the single integration site of the pathogenicity locus in *Clostridium difficile*. *Gene* **181**: 29-38.

Brazier, J. S., R. Raybould, B. Patel, G. Duckworth, A. Pearson, A. Charlett & B. I. Duerden, (2008) Distribution and antimicrobial susceptibility patterns of *Clostridium difficile* PCR ribotypes in English hospitals, 2007-08. *Euro Surveill* **13**.

Brooks, S., A. Khan, D. Stoica, J. Griffith, L. Friedeman, R. Mukherji, R. Hameed & N. Schupf, (1998) Reduction in vancomycin-resistant *Enterococcus* and *Clostridium difficile* infections following change to tympanic thermometers. *Infect.Control Hosp.Epidemiol.* **19**: 333-336.

Burbige, E. J. & F. D. Milligan, (1975) Pseudomembranous colitis. Association with antibiotics and therapy with cholestyramine. *JAMA* **231**: 1157-1158.

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

Calabi, E., F. Calabi, A. D. Phillips & N. F. Fairweather, (2002) Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. *Infect Immun* **70**: 5770-5778.

Calabi, E. & N. Fairweather, (2002a) Patterns of sequence conservation in the S-layer proteins and related sequences in *Clostridium difficile*. *J.Bacteriol.* **184**: 3886-3897.

Calabi, E. & N. Fairweather, (2002b) Patterns of sequence conservation in the S-layer proteins and related sequences in *Clostridium difficile*. *J. Bacteriol.* **184**: 3886-3897.

Calabi, E., S. Ward, B. Wren, T. Paxton, M. Panico, H. Morris, A. Dell, G. Dougan & N. Fairweather, (2001a) Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Molecular Microbiology* **40**: 1187-1199.

Calabi, E., S. Ward, B. Wren, T. Paxton, M. Panico, H. Morris, A. Dell, G. Dougan & N. Fairweather, (2001b) Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Mol.Microbiol.* **40**: 1187-1199.

Carter, G. P., D. Lyras, D. L. Allen, K. E. Mackin, P. M. Howarth, J. R. O'Connor & J. I. Rood, (2007) Binary toxin production in *Clostridium difficile* is regulated by CdtR, a LytTR family response regulator. *J.Bacteriol.* **189**: 7290-7301.

Carter, G. P., J. I. Rood & D. Lyras, (2010) The role of toxin A and toxin B in *Clostridium difficile*-associated disease: Past and present perspectives. *Gut Microbes* **1**: 58-64.

Carter, J., (2009) Deaths involving *Clostridium difficile*: England and Wales, 2008. *Health Stat.Q.*: 43-47.

Cartman, S. T., J. T. Heap, S. A. Kuehne, A. Cockayne & N. P. Minton, (2010) The emergence of 'hypervirulence' in *Clostridium difficile*. *Int.J.Med.Microbiol.*

Cartman, S. T. & N. P. Minton, (2010) A mariner-based transposon system for in vivo random mutagenesis of *Clostridium difficile*. *Appl.Environ.Microbiol.* **76**: 1103-1109.

Cerquetti, M., A. Serafino, A. Sebastianelli & P. Mastrantonio, (2002) Binding of *Clostridium difficile* to Caco-2 epithelial cell line and to extracellular matrix proteins. *FEMS Immunol.Med.Microbiol.* **32**: 211-218.

Chang, T. W., J. G. Bartlett, S. L. Gorbach & A. B. Onderdonk, (1978) Clindamycin-induced enterocolitis in hamsters as a model of pseudomembranous colitis in patients. *Infect.Immun.* **20**: 526-529.

Cheknis, A. K., S. P. Sambol, D. M. Davidson, K. J. Nagaro, M. C. Mancini, G. A. Hidalgo-Arroyo, J. S. Brazier, S. Johnson & D. N. Gerding, (2009) Distribution of *Clostridium difficile* strains from a North American, European and Australian trial of treatment for *C. difficile* infections: 2005-2007. *Anaerobe* **15**: 230-233.

Chen, X., K. Katchar, J. D. Goldsmith, N. Nanthakumar, A. Cheknis, D. N. Gerding & C. P. Kelly, (2008) A Mouse Model of *Clostridium difficile*-Associated Disease. *Gastroenterology*.

Cheng, V. C., W. C. Yam, J. F. Chan, K. K. To, P. L. Ho & K. Y. Yuen, (2009) *Clostridium difficile* ribotype 027 arrives in Hong Kong. *Int.J.Antimicrob.Agents* **34**: 492-493.

Clabots, C. R., S. Johnson, K. M. Bettin, P. A. Mathie, M. E. Mulligan, D. R. Schaberg, L. R. Peterson & D. N. Gerding, (1993) Development of a rapid and efficient restriction endonuclease analysis typing system for *Clostridium difficile* and correlation with other typing systems. *J.Clin.Microbiol.* **31**: 1870-1875.

Clabots, C. R., S. Johnson, M. M. Olson, L. R. Peterson & D. N. Gerding, (1992) Acquisition of *Clostridium difficile* by hospitalized patients: evidence for colonized new admissions as a source of infection. *J.Infect.Dis.* **166**: 561-567.

Climo, M. W., D. S. Israel, E. S. Wong, D. Williams, P. Coudron & S. M. Markowitz, (1998) Hospital-wide restriction of clindamycin: effect on the incidence of *Clostridium difficile*-associated diarrhea and cost. *Ann.Intern.Med.* **128**: 989-995.

Cohen, S. H., D. N. Gerding, S. Johnson, C. P. Kelly, V. G. Loo, L. C. McDonald, J. Pepin & M. H. Wilcox, (2010) Clinical practice guidelines for *Clostridium difficile* infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). *Infect.Control Hosp.Epidemiol.* **31**: 431-455.

Corrigan, R. M., D. Rigby, P. Handley & T. J. Foster, (2007) The role of *Staphylococcus aureus* surface protein SasG in adherence and biofilm formation. *Microbiology* **153**: 2435-2446.

Crawhall, J. C. & D. F. Elliott, (1955) Fission of peptide bonds by lithium borohydride. *Nature* **175**: 299-300.

Curry, S. R., J. W. Marsh, K. A. Shutt, C. A. Muto, M. M. O'Leary, M. I. Saul, A. W. Pasculle & L. H. Harrison, (2009) High frequency of rifampin resistance identified in an epidemic *Clostridium difficile* clone from a large teaching hospital. *Clin.Infect.Dis.* **48**: 425-429.

Dang, T. H., L. Riva Lde, R. P. Fagan, E. M. Storck, W. P. Heal, C. Janoir, N. F. Fairweather & E. W. Tate, (2010) Chemical probes of surface layer biogenesis in *Clostridium difficile*. *ACS Chem Biol* **5**: 279-285.

Dautin, N., T. J. Barnard, D. E. Anderson & H. D. Bernstein, (2007) Cleavage of a bacterial autotransporter by an evolutionarily convergent autocatalytic mechanism. *EMBO J.* **26**: 1942-1952.

Davey, P., E. Brown, L. Fenelon, R. Finch, I. Gould, G. Hartman, A. Holmes, C. Ramsay, E. Taylor, M. Wilcox & P. Wiffen, (2005) Interventions to improve antibiotic prescribing practices for hospital inpatients. *Cochrane Database Syst.Rev.*: CD003543.

Davey, P., E. Brown, L. Fenelon, R. Finch, I. Gould, A. Holmes, C. Ramsay, E. Taylor, P. Wiffen & M. Wilcox, (2006) Systematic review of antimicrobial drug prescribing in hospitals. *Emerg.Infect.Dis.* **12**: 211-216.

Dawson, L. F., R. A. Stabler & B. W. Wren, (2008) Assessing the role of p-cresol tolerance in *Clostridium difficile*. *J.Med.Microbiol.* **57**: 745-749.
Delaney, J. A., (2007) Antimicrobial drugs and community acquired *Clostridium difficile*-associated disease, UK. *Emerg.Infect.Dis.* **13**: 761-763.
Delmee, M. & V. Avesani, (1990) Virulence of ten serogroups of *Clostridium difficile* in hamsters. *J.Med.Microbiol.* **33**: 85-90.
Dendukuri, N., V. Costa, M. McGregor & J. M. Brophy, (2005) Probiotic therapy for the prevention and treatment of *Clostridium difficile*-associated diarrhea: a systematic review. *CMAJ* **173**: 167-170.
Dingle, T., G. L. Mulvey, R. M. Humphries & G. D. Armstrong, (2010) A real-time quantitative PCR assay for evaluating *Clostridium difficile* adherence to differentiated intestinal Caco-2 cells. *J.Med.Microbiol.*
Douce, G. & D. Goulding, (2010) Refinement of the Hamster Model of *Clostridium difficile* Disease. *Methods Mol.Biol.* **646**: 215-227.
Drepper, T., T. Eggert, F. Circolone, A. Heck, U. Krauss, J. K. Guterl, M. Wendorff, A. Losi, W. Gartner & K. E. Jaeger, (2007a) Reporter proteins for in vivo fluorescence without oxygen. *Nat.Biotechnol.* **25**: 443-445.
Drepper, T., T. Eggert, F. Circolone, A. Heck, U. Krausz, J.-K. Guterl, M. Wendorff, A. Losi, W. Gartner & K.-E. Jaeger, (2007b) Reporter proteins for *in vivo* fluorescence without oxygen. *Nat Biotech* **25**: 443-445.
Drudy, D., E. Calabi, L. Kyne, S. Sougioultzis, E. Kelly, N. Fairweather & C. P. Kelly, (2004) Human antibody response to surface layer proteins in *Clostridium difficile* infection. *FEMS Immunol.Med.Microbiol.* **41**: 237-242.
Drudy, D., S. Fanning & L. Kyne, (2006a) Toxin A-negative, toxin B-positive *Clostridium difficile*. *Int J Infect Dis*.
Drudy, D., S. Fanning & L. Kyne, (2007a) Toxin A-negative, toxin B-positive *Clostridium difficile*. *Int J Infect Dis* **11**: 5-10.
Drudy, D., N. Harnedy, S. Fanning, M. Hannan & L. Kyne, (2007b) Emergence and control of fluoroquinolone-resistant, toxin A-negative, toxin B-positive *Clostridium difficile*. *Infect.Control Hosp.Epidemiol.* **28**: 932-940.
Drudy, D., D. P. O'Donoghue, A. Baird, L. Fenelon & C. O'Farrelly, (2001) Flow cytometric analysis of *Clostridium difficile* adherence to human intestinal epithelial cells. *J.Med.Microbiol.* **50**: 526-534.
Drudy, D., T. Quinn, R. O'Mahony, L. Kyne, P. O'Gaora & S. Fanning, (2006b) High-level resistance to moxifloxacin and gatifloxacin associated with a novel mutation in gyrB in toxin-A-negative, toxin-B-positive *Clostridium difficile*. *J.Antimicrob.Chemother.* **58**: 1264-1267.
Dubberke, E. R. & A. I. Wertheimer, (2009) Review of current literature on the economic burden of *Clostridium difficile* infection. *Infect.Control Hosp.Epidemiol.* **30**: 57-66.
Dupuy, B., R. Govind, A. Antunes & S. Matamouros, (2008) *Clostridium difficile* toxin synthesis is negatively regulated by TcdC. *J.Med.Microbiol.* **57**: 685-689.
Dupuy, B. & S. Matamouros, (2006) Regulation of toxin and bacteriocin synthesis in *Clostridium* species by a new subgroup of RNA polymerase sigma-factors. *Res.Microbiol.* **157**: 201-205.
Dupuy, B., S. Raffestin, S. Matamouros, N. Mani, M. R. Popoff & A. L. Sonenshein, (2006) Regulation of toxin and bacteriocin gene expression in *Clostridium* by interchangeable RNA polymerase sigma factors. *Mol.Microbiol.* **60**: 1044-1057.
Dupuy, B. & A. L. Sonenshein, (1998) Regulated transcription of *Clostridium difficile* toxin genes. *Mol.Microbiol.* **27**: 107-120.

Eastwood, K., P. Else, A. Charlett & M. Wilcox, (2009) Comparison of nine commercially available *Clostridium difficile* toxin detection assays, a real-time PCR assay for *C. difficile* tcdB, and a glutamate dehydrogenase detection assay to cytotoxin testing and cytotoxigenic culture methods. *J.Clin.Microbiol.* **47**: 3211-3217.

Eglow, R., C. Pothoulakis, S. Itzkowitz, E. J. Israel, C. J. O'Keane, D. Gong, N. Gao, Y. L. Xu, W. A. Walker & J. T. LaMont, (1992) Diminished *Clostridium difficile* toxin A sensitivity in newborn rabbit ileum is associated with decreased toxin A receptor. *J.Clin.Invest.* **90**: 822-829.

Eidhin, D., A. Ryan, R. Doyle, J. B. Walsh & D. Kelleher, (2006) Sequence and phylogenetic analysis of the gene for surface layer protein, *slpA*, from 14 PCR ribotypes of *Clostridium difficile*. *J Med Microbiol* **55**: 69-83.

Emerson, J. E., C. B. Reynolds, R. P. Fagan, H. A. Shaw, D. Goulding & N. F. Fairweather, (2009) A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. *Mol.Microbiol.* **74**: 541-556.

Emerson, J. E., R. A. Stabler, B. W. Wren & N. F. Fairweather, (2008) Microarray analysis of the transcriptional responses of *Clostridium difficile* to environmental and antibiotic stress. In., pp. 757-764.

Eveillard, M., V. Fourel, M. C. Barc, S. Kerneis, M. H. Coconnier, T. Karjalainen, P. Bourlioux & A. L. Servin, (1993) Identification and characterization of adhesive factors of *Clostridium difficile* involved in adhesion to human colonic enterocyte-like Caco-2 and mucus-secreting HT29 cells in culture. *Mol.Microbiol.* **7**: 371-381.

Fagan, R. P., D. Albesa-Jove, O. Qazi, D. I. Svergun, K. A. Brown & N. F. Fairweather, (2009) Structural insights into the molecular organization of the S-layer from *Clostridium difficile*. *Mol.Microbiol.* **71**: 1308-1322.

Fekety, R., K. H. Kim, D. Brown, D. H. Batts, M. Cudmore & J. Silva, Jr., (1981) Epidemiology of antibiotic-associated colitis; isolation of *Clostridium difficile* from the hospital environment. *Am.J.Med.* **70**: 906-908.

Fekety, R., J. Silva, R. Toshniwal, M. Allo, J. Armstrong, R. Browne, J. Ebright & G. Rifkin, (1979) Antibiotic-associated colitis: effects of antibiotics on *Clostridium difficile* and the disease in hamsters. *Rev Infect Dis* **1**: 386-397.

Fernie, D. S., R. O. Thomson, I. Batty & P. D. Walker, (1983) Active and passive immunization to protect against antibiotic associated caecitis in hamsters. *Dev.Biol.Stand.* **53**: 325-332.

Freeman, J., S. D. Baines, D. Jabes & M. H. Wilcox, (2005) Comparison of the efficacy of ramoplanin and vancomycin in both *in vitro* and *in vivo* models of clindamycin-induced *Clostridium difficile* infection. *J.Antimicrob.Chemother.* **56**: 717-725.

Freeman, J., S. D. Baines, K. Saxton & M. H. Wilcox, (2007) Effect of metronidazole on growth and toxin production by epidemic *Clostridium difficile* PCR ribotypes 001 and 027 in a human gut model. *J.Antimicrob.Chemother.* **60**: 83-91.

Freeman, J., M. P. Bauer, S. D. Baines, J. Corver, W. N. Fawley, B. Goorhuis, E. J. Kuijper & M. H. Wilcox, (2010) The changing epidemiology of *Clostridium difficile* infections. *Clin.Microbiol.Rev.* **23**: 529-549.

Freeman, J., F. J. O'Neill & M. H. Wilcox, (2003) Effects of cefotaxime and desacetylcefotaxime upon *Clostridium difficile* proliferation and toxin production in a triple-stage chemostat model of the human gut. *J.Antimicrob.Chemother.* **52**: 96-102.

French, E., A. Rodriguez-Palacios & J. T. Lejeune, (2010) Enteric Bacterial Pathogens with Zoonotic Potential Isolated from Farm-Raised Deer. *Foodborne Pathog Dis*.

Freter, R., E. Stauffer, D. Cleven, L. V. Holdeman & W. E. Moore, (1983) Continuous-flow cultures as *in vitro* models of the ecology of large intestinal flora. *Infect.Immun.* **39**: 666-675.

Gally, D. L., J. Leathart & I. C. Blomfield, (1996) Interaction of FimB and FimE with the *fim* switch that controls the phase variation of type 1 fimbriae in *Escherichia coli* K-12. *Mol.Microbiol.* **21**: 725-738.

Gao, X. W., M. Mubasher, C. Y. Fang, C. Reifer & L. E. Miller, (2010) Dose-Response Efficacy of a Proprietary Probiotic Formula of Lacto*Bacillus* acidophilus CL1285 and Lacto*Bacillus* casei LBC80R for Antibiotic-Associated Diarrhea and *Clostridium difficile*-Associated Diarrhea Prophylaxis in Adult Patients. *Am.J.Gastroenterol.*

Gardiner, D. F., T. Rosenberg, J. Zaharatos, D. Franco & D. D. Ho, (2009) A DNA vaccine targeting the receptor-binding domain of *Clostridium difficile* toxin A. *Vaccine* **27**: 3598-3604.

Garey, K. W., M. Salazar, D. Shah, R. Rodrigue & H. L. DuPont, (2008) Rifamycin antibiotics for treatment of *Clostridium difficile*-associated diarrhea. *Ann.Pharmacother.* **42**: 827-835.

Geoghegan, J. A., R. M. Corrigan, D. T. Gruszka, P. Speziale, J. P. O'Gara, J. R. Potts & T. J. Foster, (2010) The Role of Surface Protein SasG in Biofilm Formation by *Staphylococcus aureus*. *J.Bacteriol.*

Ghantoji, S. S., K. Sail, D. R. Lairson, H. L. DuPont & K. W. Garey, (2010) Economic healthcare costs of *Clostridium difficile* infection: a systematic review. *J.Hosp.Infect.* **74**: 309-318.

Giannasca, P. J., Z. X. Zhang, W. D. Lei, J. A. Boden, M. A. Giel, T. P. Monath & W. D. Thomas, Jr., (1999) Serum antitoxin antibodies mediate systemic and mucosal protection from *Clostridium difficile* disease in hamsters. *Infect.Immun.* **67**: 527-538.

Giel, J. L., J. A. Sorg, A. L. Sonenshein & J. Zhu, (2010) Metabolism of bile salts in mice influences spore germination in *Clostridium difficile*. *PLoS One* **5**: e8740.

Goorhuis, A., D. Bakker, J. Corver, S. B. Debast, C. Harmanus, D. W. Notermans, A. A. Bergwerff, F. W. Dekker & E. J. Kuijper, (2008a) Emergence of *Clostridium difficile* infection due to a new hypervirulent strain, polymerase chain reaction ribotype 078. *Clin.Infect.Dis.* **47**: 1162-1170.

Goorhuis, A., S. B. Debast, L. A. van Leengoed, C. Harmanus, D. W. Notermans, A. A. Bergwerff & E. J. Kuijper, (2008b) *Clostridium difficile* PCR ribotype 078: an emerging strain in humans and in pigs? *J.Clin.Microbiol.* **46**: 1157; author reply 1158.

Goulding, D., H. Thompson, J. Emerson, N. F. Fairweather, G. Dougan & G. R. Douce, (2009) Distinctive profiles of infection and pathology in hamsters infected with *Clostridium difficile* strains 630 and B1. *Infect.Immun.* **77**: 5478-5485.

Graul, T., A. M. Cain & K. D. Karpa, (2009) Lacto*Bacillus* and bifidobacteria combinations: a strategy to reduce hospital-acquired *Clostridium difficile* diarrhea incidence and mortality. *Med Hypotheses* **73**: 194-198.

Greco, A., J. G. Ho, S. J. Lin, M. M. Palcic, M. Rupnik & K. K. Ng, (2006) Carbohydrate recognition by *Clostridium difficile* toxin A. *Nat.Struct.Mol.Biol.* **13**: 460-461.

Griffiths, D., W. Fawley, M. Kachrimanidou, R. Bowden, D. W. Crook, R. Fung, T. Golubchik, R. M. Harding, K. J. Jeffery, K. A. Jolley, R. Kirton, T. E. Peto, G. Rees, N. Stoesser, A. Vaughan, A. S. Walker, B. C. Young, M. Wilcox & K. E. Dingle, (2010) Multilocus sequence typing of *Clostridium difficile*. *J.Clin.Microbiol.* **48**: 770-778.

Gustafsson, A., S. Lund-Tonnesen, A. Berstad, T. Midtvedt & E. Norin, (1998) Faecal short-chain fatty acids in patients with antibiotic-associated diarrhoea, before and after faecal enema treatment. *Scand.J.Gastroenterol.* **33**: 721-727.

Hafiz, S. & C. L. Oakley, (1976) *Clostridium difficile*: isolation and characteristics. *J.Med.Microbiol.* **9**: 129-136.

Hall, I. C. & E. O'Toole, (1935) Intestinal flora in newborn infants with a description of a new pathogenic anaerobe, *Bacillus difficilis*. *Am.J.Dis.Child.* **49**: 390-402.

He, M., M. Sebaihia, T. D. Lawley, R. A. Stabler, L. F. Dawson, M. J. Martin, K. E. Holt, H. M. Seth-Smith, M. A. Quail, R. Rance, K. Brooks, C. Churcher, D. Harris, S. D. Bentley, C. Burrows, L. Clark, C. Corton, V. Murray, G. Rose, S. Thurston, A. van Tonder, D. Walker, B. W. Wren, G. Dougan & J. Parkhill, (2010) Evolutionary dynamics of *Clostridium difficile* over short and long time scales. *Pro*C.*Natl.Acad.Sci.USA* **107**: 7527-7532.

Heap, J. T., S. A. Kuehne, M. Ehsaan, S. T. Cartman, C. M. Cooksley, J. C. Scott & N. P. Minton, (2010) The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. *J.Microbiol.Methods* **80**: 49-55.

Heap, J. T., O. J. Pennington, S. T. Cartman, G. P. Carter & N. P. Minton, (2007) The ClosTron: a universal gene knock-out system for the genus *Clostridium*. *J.Microbiol.Methods* **70**: 452-464.

Heap, J. T., O. J. Pennington, S. T. Cartman & N. P. Minton, (2009) A modular system for *Clostridium* shuttle plasmids. *J.Microbiol.Methods* **78**: 79-85.

Heger, A. & L. Holm, (2000) Rapid automatic detection and alignment of repeats in protein sequences. *Proteins* **41**: 224-237.

Heichman, K. A. & R. C. Johnson, (1990) The Hin invertasome: protein-mediated joining of distant recombination sites at the enhancer. *Science* **249**: 511-517.

Hennequin, C., C. Janoir, M.-C. Barc, A. Collignon & T. Karjalainen, (2003) Identification and characterization of a fibronectin-binding protein from *Clostridium difficile*. *Microbiology* **149**: 2779-2787.

Herpers, B. L., B. Vlaminckx, O. Burkhardt, H. Blom, H. S. Biemond-Moeniralam, M. Hornef, T. Welte & E. J. Kuijper, (2009) Intravenous tigecycline as adjunctive or alternative therapy for severe refractory *Clostridium difficile* infection. *Clin.Infect.Dis.* **48**: 1732-1735.

HICPAC, (1995) Recommendations for preventing the spread of vancomycin resistance. Recommendations of the Hospital Infection Control Practices Advisory Committee (HICPAC). *MMWR Recomm Rep* **44**: 1-13.

Hidalgo, I. J., T. J. Raub & R. T. Borchardt, (1989) Characterization of the human colon carcinoma cell line (Caco-2) as a model system for intestinal epithelial permeability. *Gastroenterology* **96**: 736-749.

Hinde, P., P. Deighan & C. J. Dorman, (2005) Characterization of the detachable Rho-dependent transcription terminator of the fimE gene in *Escherichia coli* K-12. *J.Bacteriol.* **187**: 8256-8266.

Hollingshead, S. K., V. A. Fischetti & J. R. Scott, (1986) Complete nucleotide sequence of type 6 M protein of the group A Streptococcus. Repetitive structure and membrane anchor. *J.Biol.Chem.* **261**: 1677-1686.

Hopkins, M. J. & G. T. Macfarlane, (2003) Nondigestible oligosaccharides enhance bacterial colonization resistance against *Clostridium difficile in vitro*. *Appl.Environ.Microbiol.* **69**: 1920-1927.

Huang, Y. J., H. W. Liao, C. C. Wu & H. L. Peng, (2009) MrkF is a component of type 3 fimbriae in Klebsiella pneumoniae. *Res.Microbiol.* **160**: 71-79.

Hussain, H. A., A. P. Roberts & P. Mullany, (2005) Generation of an erythromycin-sensitive derivative of *Clostridium difficile* strain 630 (630Deltaerm) and demonstration that the conjugative transposon Tn916DeltaE enters the genome of this strain at multiple sites. *J.Med.Microbiol.* **54**: 137-141.

Ingham, K. C., S. Brew, D. Vaz, D. N. Sauder & M. J. McGavin, (2004) Interaction of *Staphylococcus aureus* fibronectin-binding protein with fibronectin: affinity, stoichiometry, and modular requirements. *J.Biol.Chem.* **279**: 42945-42953.

Itoh, K., W. K. Lee, H. Kawamura, T. Mitsuoka & T. Magaribuchi, (1986) Isolation of *Clostridium difficile* from various colonies of laboratory mice. *Lab.Anim.* **20**: 266-270.

Jabbar, U., J. Leischner, D. Kasper, R. Gerber, S. P. Sambol, J. P. Parada, S. Johnson & D. N. Gerding, (2010) Effectiveness of alcohol-based hand rubs for removal of *Clostridium difficile* spores from hands. *Infect.Control Hosp.Epidemiol.* **31**: 565-570.

Jangi, S. & J. T. Lamont, (2010) Asymptomatic colonization by *Clostridium difficile* in infants: implications for disease in later life. *J.Pediatr.Gastroenterol.Nutr.* **51**: 2-7.

Jank, T. & K. Aktories, (2008) Structure and mode of action of clostridial glucosylating toxins: the ABCD model. *Trends Microbiol.* **16**: 222-229.

Janoir, C., S. Pechine, C. Grosdidier & A. Collignon, (2007) Cwp84, a surface-associated protein of *Clostridium difficile*, is a cysteine protease with degrading activity on extracellular matrix proteins. *J Bacteriol* **189**: 7174-7180.

Janvilisri, T., J. Scaria, A. D. Thompson, A. Nicholson, B. M. Limbago, L. G. Arroyo, J. G. Songer, Y. T. Grohn & Y. F. Chang, (2009) Microarray identification of *Clostridium difficile* core components and divergent regions associated with host origin. *J.Bacteriol.* **191**: 3881-3891.

Jernigan, J. A., Y. Siegman-Igra, R. C. Guerrant & B. M. Farr, (1998) A randomized crossover study of disposable thermometers for prevention of *Clostridium difficile* and other nosocomial infections. *Infect.Control Hosp.Epidemiol.* **19**: 494-499.

Johal, S. S., C. P. Lambert, J. Hammond, P. D. James, S. P. Borriello & Y. R. Mahida, (2004) Colonic IgA producing cells and macrophages are reduced in recurrent and non-recurrent *Clostridium difficile* associated diarrhoea. *J.Clin.Pathol.* **57**: 973-979.

Johansson, D. G., B. Macao, A. Sandberg & T. Hard, (2008) SEA domain autoproteolysis accelerated by conformational strain: mechanistic aspects. *J.Mol.Biol.* **377**: 1130-1143.

Johnson, S., A. Adelmann, C. R. Clabots, L. R. Peterson & D. N. Gerding, (1989) Recurrences of *Clostridium difficile* diarrhea not caused by the original infecting organism. *J.Infect.Dis.* **159**: 340-343.

Johnson, S., D. N. Gerding, M. M. Olson, M. D. Weiler, R. A. Hughes, C. R. Clabots & L. R. Peterson, (1990) Prospective, controlled study of vinyl glove use to interrupt *Clostridium difficile* nosocomial transmission. *Am.J.Med.* **88**: 137-140.

Johnson, S., S. R. Homann, K. M. Bettin, J. N. Quick, C. R. Clabots, L. R. Peterson & D. N. Gerding, (1992) Treatment of asymptomatic *Clostridium difficile* carriers (fecal excretors) with vancomycin or metronidazole. A randomized, placebo-controlled trial. *Ann.Intern.Med.* **117**: 297-302.

Johnson, S., C. Schriever, M. Galang, C. P. Kelly & D. N. Gerding, (2007) Interruption of recurrent *Clostridium difficile*-associated diarrhea episodes by serial therapy with vancomycin and rifaximin. *Clin.Infect.Dis.* **44**: 846-848.

Jones, M. A. & D. Hunter, (1983) Isolation of *Clostridium difficile* from pigs. *Vet.Re*C. **112**: 253.

Jonsson, K., C. Signas, H. P. Muller & M. Lindberg, (1991) Two different genes encode fibronectin binding proteins in *Staphylococcus aureus*. The complete nucleotide sequence and characterization of the second gene. *Eur.J.Biochem.* **202**: 1041-1048.

Joyce, S. A. & C. J. Dorman, (2002) A Rho-dependent phase-variable transcription terminator controls expression of the FimE recombinase in *Escherichia coli*. *Mol.Microbiol.* **45**: 1107-1117.

Kaier, K., D. Luft, M. Dettenkofer, M. Kist & U. Frank, (2010) Correlations between bed occupancy rates and *Clostridium difficile* infections: a time-series analysis. *Epidemiol.Infect.*: 1-4.

Kapoor, K., M. Chandra, D. Nag, J. K. Paliwal, R. C. Gupta & R. C. Saxena, (1999) Evaluation of metronidazole toxicity: a prospective study. *Int J Clin Pharmacol Res* **19**: 83-88.

Karas, J. A., D. A. Enoch & S. H. Aliyu, (2010) A review of mortality due to *Clostridium difficile* infection. *J.Infect.*

Karberg, M., H. Guo, J. Zhong, R. Coon, J. Perutka & A. M. Lambowitz, (2001) Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nat.Biotechnol.* **19**: 1162-1167.

Karjalainen, T., N. Saumier, M.-C. Barc, M. Delmee & A. Collignon, (2002a) *Clostridium difficile* genotyping based on *slpA* variable region in S-layer gene sequence: an alternative to serotyping. *J. Clin. Microbiol.* **40**: 2452-2458.

Karjalainen, T., N. Saumier, M. C. Barc, M. Delmee & A. Collignon, (2002b) *Clostridium difficile* genotyping based on *slpA* variable region in S-layer gene sequence: an alternative to serotyping. *J.Clin.Microbiol.* **40**: 2452-2458.

Karjalainen, T., A. J. Waligora-Dupriet, M. Cerquetti, P. Spigaglia, A. Maggioni, P. Mauri & P. Mastrantonio, (2001) Molecular and genomic analysis of genes encoding surface-anchored proteins from *Clostridium difficile*. *Infect.Immun.* **69**: 3442-3446.

Karlsson, S., L. G. Burman & T. Akerlund, (2008) Induction of toxins in *Clostridium difficile* is associated with dramatic changes of its metabolism. *Microbiology* **154**: 3430-3436.

Karlsson, S., B. Dupuy, K. Mukherjee, E. Norin, L. G. Burman & T. Akerlund, (2003) Expression of *Clostridium difficile* toxins A and B and their sigma factor TcdD is controlled by temperature. *Infect.Immun.* **71**: 1784-1793.

Kato, H., Y. Ito, R. J. van den Berg, E. J. Kuijper & Y. Arakawa, (2007) First isolation of *Clostridium difficile* 027 in Japan. *Euro Surveill* **12**: E070111 070113.

Kato, H., N. Kato, K. Watanabe, T. Yamamoto, K. Suzuki, S. Ishigo, S. Kunihiro, I. Nakamura, G. E. Killgore & S. Nakamura, (2001) Analysis of *Clostridium*

difficile isolates from nosocomial outbreaks at three hospitals in diverse areas of Japan. *J.Clin.Microbiol.* **39**: 1391-1395.

Kato, H., T. Yokoyama & Y. Arakawa, (2005) Typing by sequencing the slpA gene of *Clostridium difficile* strains causing multiple outbreaks in Japan. *J.Med.Microbiol.* **54**: 167-171.

Keighley, M. R., D. W. Burdon, Y. Arabi, J. A. Williams, H. Thompson, D. Youngs, M. Johnson, S. Bentley, R. H. George & G. A. Mogg, (1978) Randomised controlled trial of vancomycin for pseudomembranous colitis and postoperative diarrhoea. *Br.Med.J.* **2**: 1667-1669.

Khoruts, A., J. Dicksved, J. K. Jansson & M. J. Sadowsky, (2010) Changes in the composition of the human fecal microbiome after bacteriotherapy for recurrent *Clostridium difficile*-associated diarrhea. *J.Clin.Gastroenterol.* **44**: 354-360.

Killgore, G., A. Thompson, S. Johnson, J. Brazier, E. Kuijper, J. Pepin, E. H. Frost, P. Savelkoul, B. Nicholson, R. J. van den Berg, H. Kato, S. P. Sambol, W. Zukowski, C. Woods, B. Limbago, D. N. Gerding & L. C. McDonald, (2008) Comparison of seven techniques for typing international epidemic strains of *Clostridium difficile*: restriction endonuclease analysis, pulsed-field gel electrophoresis, PCR-ribotyping, multilocus sequence typing, multilocus variable-number tandem-repeat analysis, amplified fragment length polymorphism, and surface layer protein A gene sequence typing. *J.Clin.Microbiol.* **46**: 431-437.

Kim, H., T. V. Riley, M. Kim, C. K. Kim, D. Yong, K. Lee, Y. Chong & J. W. Park, (2008) Increasing prevalence of toxin A-negative, toxin B-positive isolates of *Clostridium difficile* in Korea: impact on laboratory diagnosis. *J.Clin.Microbiol.* **46**: 1116-1117.

Kim, K., L. K. Pickering, H. L. DuPont, N. Sullivan & T. Wilkins, (1984) In vitro and in vivo neutralizing activity of human colostrum and milk against purified toxins A and B of *Clostridium difficile*. *J.Infect.Dis.* **150**: 57-62.

Kim, K. H., R. Fekety, D. H. Batts, D. Brown, M. Cudmore, J. Silva, Jr. & D. Waters, (1981) Isolation of *Clostridium difficile* from the environment and contacts of patients with antibiotic-associated colitis. *J.Infect.Dis.* **143**: 42-50.

Kim, P. H., J. P. Iaconis & R. D. Rolfe, (1987) Immunization of adult hamsters against *Clostridium difficile*-associated ileocecitis and transfer of protection to infant hamsters. *Infect.Immun.* **55**: 2984-2992.

Kink, J. A. & J. A. Williams, (1998) Antibodies to recombinant *Clostridium difficile* toxins A and B are an effective treatment and prevent relapse of *C. difficile*-associated disease in a hamster model of infection. *Infect.Immun.* **66**: 2018-2025.

Kirby, J. M., H. Ahern, A. K. Roberts, V. Kumar, Z. Freeman, K. R. Acharya & C. C. Shone, (2009) Cwp84, a surface-associated cysteine protease, plays a role in the maturation of the surface layer of *Clostridium difficile*. *J.Biol.Chem.* **284**: 34666-34673.

Klemm, P., (1986) Two regulatory fim genes, fimB and fimE, control the phase variation of type 1 fimbriae in *Escherichia coli*. *EMBO J.* **5**: 1389-1393.

Kornblihtt, A. R., K. Umezawa, K. Vibe-Pedersen & F. E. Baralle, (1985) Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. *EMBO J.* **4**: 1755-1759.

Kotloff, K. L., J. C. Wade & J. G. Morris, Jr., (1988) Lack of association between *Clostridium difficile* toxin and diarrhea in infants. *Pediatr.Infect.Dis.J.* **7**: 662-663.

Krivan, H. C., G. F. Clark, D. F. Smith & T. D. Wilkins, (1986) Cell surface binding site for *Clostridium difficile* enterotoxin: evidence for a glycoconjugate containing the sequence Gal alpha 1-3Gal beta 1-4GlcNAC. *Infect.Immun.* **53**: 573-581.

Kuboniwa, M., A. Amano, E. Hashino, Y. Yamamoto, H. Inaba, N. Hamada, K. Nakayama, G. D. Tribble, R. J. Lamont & S. Shizukuishi, (2009) Distinct roles of long/short fimbriae and gingipains in homotypic biofilm development by *Porphyromonas gingivalis*. *BMC Microbiol* **9**: 105.

Kuehne, S. A., S. T. Cartman, J. T. Heap, M. L. Kelly, A. Cockayne & N. P. Minton, (2010) The role of toxin A and toxin B in *Clostridium difficile* infection. *Nature*.

Kuijper, E. J., B. Coignard, J. S. Brazier, C. Suetens, D. Drudy, C. Wiuff, H. Pituch, P. Reichert, F. Schneider, A. F. Widmer, K. E. Olsen, F. Allerberger, D. W. Notermans, F. Barbut, M. Delmee, M. Wilcox, A. Pearson, B. C. Patel, D. J. Brown, R. Frei, T. Akerlund, I. R. Poxton & P. Tull, (2007) Update of *Clostridium difficile*-associated disease due to PCR ribotype 027 in Europe. *Euro Surveill* **12**: E1-2.

Kutty, P. K., C. W. Woods, A. C. Sena, S. R. Benoit, S. Naggie, J. Frederick, S. Evans, J. Engel & L. C. McDonald, (2010) Risk factors for and estimated incidence of community-associated *Clostridium difficile* infection, North Carolina, USA. *Emerg.Infect.Dis.* **16**: 197-204.

Kyne, L., M. B. Hamel, R. Polavaram & C. P. Kelly, (2002) Health care costs and mortality associated with nosocomial diarrhea due to *Clostridium difficile*. *Clin.Infect.Dis.* **34**: 346-353.

Kyne, L., M. Warny, A. Qamar & C. P. Kelly, (2000) Asymptomatic carriage of *Clostridium difficile* and serum levels of IgG antibody against toxin A. *N.Engl.J.Med.* **342**: 390-397.

Kyne, L., M. Warny, A. Qamar & C. P. Kelly, (2001) Association between antibody response to toxin A and protection against recurrent *Clostridium difficile* diarrhoea. *Lancet* **357**: 189-193.

Lamontagne, F., A. C. Labbe, O. Haeck, O. Lesur, M. Lalancette, C. Patino, M. Leblanc, M. Laverdiere & J. Pepin, (2007) Impact of emergency colectomy on survival of patients with fulminant *Clostridium difficile* colitis during an epidemic caused by a hypervirulent strain. *Ann.Surg.* **245**: 267-272.

Larson, H. E. & S. P. Borriello, (1990) Quantitative study of antibiotic-induced susceptibility to *Clostridium difficile* enterocecitis in hamsters. *Antimicrob.Agents Chemother.* **34**: 1348-1353.

Lawley, T. D., S. Clare, A. W. Walker, D. Goulding, R. A. Stabler, N. Croucher, P. Mastroeni, P. Scott, C. Raisen, L. Mottram, N. F. Fairweather, B. W. Wren, J. Parkhill & G. Dougan, (2009a) Antibiotic treatment of *Clostridium difficile* carrier mice triggers a supershedder state, spore-mediated transmission, and severe disease in immunocompromised hosts. *Infect.Immun.* **77**: 3661-3669.

Lawley, T. D., N. J. Croucher, L. Yu, S. Clare, M. Sebaihia, D. Goulding, D. J. Pickard, J. Parkhill, J. Choudhary & G. Dougan, (2009b) Proteomic and genomic characterization of highly infectious *Clostridium difficile* 630 spores. *J.Bacteriol.* **191**: 5377-5386.

Lawrence, S. J., L. A. Puzniak, B. N. Shadel, K. N. Gillespie, M. H. Kollef & L. M. Mundy, (2007) *Clostridium difficile* in the intensive care unit: epidemiology, costs, and colonization pressure. *Infect.Control Hosp.Epidemiol.* **28**: 123-130.

Leav, B. A., B. Blair, M. Leney, M. Knauber, C. Reilly, I. Lowy, D. N. Gerding, C. P. Kelly, K. Katchar, R. Baxter, D. Ambrosino & D. Molrine, (2010) Serum anti-toxin B antibody correlates with protection from recurrent *Clostridium difficile* infection (CDI). *Vaccine* **28**: 965-969.

Lee, A. S. & K. P. Song, (2005) LuxS/autoinducer-2 quorum sensing molecule regulates transcriptional virulence gene expression in *Clostridium difficile*. *Biochem.Biophys.Res.Commun.* **335**: 659-666.

Lee, B. Y., M. J. Popovich, Y. Tian, R. R. Bailey, P. J. Ufberg, A. E. Wiringa & R. R. Muder, (2010) The potential value of *Clostridium difficile* vaccine: An economic computer simulation model. *Vaccine*.

Lefebvre, S. L., L. G. Arroyo & J. S. Weese, (2006a) Epidemic *Clostridium difficile* strain in hospital visitation dog. *Emerg.Infect.Dis.* **12**: 1036-1037.

Lefebvre, S. L., D. Waltner-Toews, A. S. Peregrine, R. Reid-Smith, L. Hodge, L. G. Arroyo & J. S. Weese, (2006b) Prevalence of zoonotic agents in dogs visiting hospitalized people in Ontario: implications for infection control. *J.Hosp.Infect.* **62**: 458-466.

Lefebvre, S. L. & J. S. Weese, (2009) Contamination of pet therapy dogs with MRSA and *Clostridium difficile*. *J.Hosp.Infect.* **72**: 268-269.

Lemee, L., I. Bourgeois, E. Ruffin, A. Collignon, J. F. Lemeland & J. L. Pons, (2005) Multilocus sequence analysis and comparative evolution of virulence-associated genes and housekeeping genes of *Clostridium difficile*. *Microbiology* **151**: 3171-3180.

Lemee, L., A. Dhalluin, M. Pestel-Caron, J. F. Lemeland & J. L. Pons, (2004) Multilocus sequence typing analysis of human and animal *Clostridium difficile* isolates of various toxigenic types. *J.Clin.Microbiol.* **42**: 2609-2617.

Lesnik, E. A., R. Sampath, H. B. Levene, T. J. Henderson, J. A. McNeil & D. J. Ecker, (2001) Prediction of rho-independent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Res.* **29**: 3583-3594.

Levy, P., O. Loreal, A. Munier, Y. Yamada, J. Picard, G. Cherqui, B. Clement & J. Capeau, (1994) Enterocytic differentiation of the human Caco-2 cell line is correlated with down-regulation of fibronectin and laminin. *FEBS Lett.* **338**: 272-276.

Libby, J. M., B. S. Jortner & T. D. Wilkins, (1982) Effects of the two toxins of *Clostridium difficile* in antibiotic-associated cecitis in hamsters. *Infect.Immun.* **36**: 822-829.

Loo, V. G., L. Poirier, M. A. Miller, M. Oughton, M. D. Libman, S. Michaud, A. M. Bourgault, T. Nguyen, C. Frenette, M. Kelly, A. Vibien, P. Brassard, S. Fenn, K. Dewar, T. J. Hudson, R. Horn, P. Rene, Y. Monczak & A. Dascal, (2005) A predominantly clonal multi-institutional outbreak of *Clostridium difficile*-associated diarrhea with high morbidity and mortality. *N.Engl.J.Med.* **353**: 2442-2449.

Louie, T. J., J. Emery, W. Krulicki, B. Byrne & M. Mah, (2009) OPT-80 eliminates *Clostridium difficile* and is sparing of bacteroides species during treatment of *C. difficile* infection. *Antimicrob.Agents Chemother.* **53**: 261-263.

Lowy, I., D. C. Molrine, B. A. Leav, B. M. Blair, R. Baxter, D. N. Gerding, G. Nichol, W. D. Thomas, Jr., M. Leney, S. Sloan, C. A. Hay & D. M. Ambrosino, (2010) Treatment with monoclonal antibodies against *Clostridium difficile* toxins. *N.Engl.J.Med.* **362**: 197-205.

Lu, C. L., C. Y. Liu, C. H. Liao, Y. T. Huang, H. P. Wang & P. R. Hsueh, (2010) Severe and refractory *Clostridium difficile* infection successfully treated with tigecycline and metronidazole. *Int.J.Antimicrob.Agents* **35**: 311-312.

Lyerly, D. M., K. E. Saum, D. K. MacDonald & T. D. Wilkins, (1985) Effects of *Clostridium difficile* toxins given intragastrically to animals. *Infect.Immun.* **47**: 349-352.

Lyras, D., V. Adams, I. Lucet & J. I. Rood, (2004) The large resolvase TnpX is the only transposon-encoded protein required for transposition of the Tn4451/3 family of integrative mobilizable elements. *Mol.Microbiol.* **51**: 1787-1800.

Lyras, D., J. R. O'Connor, P. M. Howarth, S. P. Sambol, G. P. Carter, T. Phumoonna, R. Poon, V. Adams, G. Vedantam, S. Johnson, D. N. Gerding & J. I. Rood, (2009) Toxin B is essential for virulence of *Clostridium difficile*. *Nature* **458**: 1176-1179.

Macfarlane, G. T., S. Macfarlane & G. R. Gibson, (1998) Validation of a three-stage compound continuous culture system for investigating the effect of retention time on the ecology and metabolism of bacteria in the human colon. *Microb.Ecol.* **35**: 180-187.

Macfarlane, S. & J. F. Dillon, (2007) Microbial biofilms in the human gastrointestinal tract. *J.Appl.Microbiol.* **102**: 1187-1196.

Machleidt, T., M. Robers & G. T. Hanson, (2007) Protein labeling with FlAsH and ReAsH. *Methods Mol.Biol.* **356**: 209-220.

Mani, N. & B. Dupuy, (2001) Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. *Pro*C.*Natl.Acad.Sci.USA* **98**: 5844-5849.

Mani, N., D. Lyras, L. Barroso, P. Howarth, T. Wilkins, J. I. Rood, A. L. Sonenshein & B. Dupuy, (2002) Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. *J.Bacteriol.* **184**: 5971-5978.

Marcotte, E. M., M. Pellegrini, T. O. Yeates & D. Eisenberg, (1999) A census of protein repeats. *J.Mol.Biol.* **293**: 151-160.

Marsden, G. L., I. J. Davis, V. J. Wright, M. Sebaihia, E. J. Kuijper & N. P. Minton, (2010) Array comparative hybridisation reveals a high degree of similarity between UK and European clinical isolates of hypervirulent *Clostridium difficile*. *BMC Genomics* **11**: 389.

Matamouros, S., P. England & B. Dupuy, (2007) *Clostridium difficile* toxin expression is inhibited by the novel regulator TcdC. *Mol.Microbiol.* **64**: 1274-1288.

Mattick, J. S., (2002) Type IV pili and twitching motility. *Annu.Rev.Microbiol.* **56**: 289-314.

Mattila, E., V. J. Anttila, M. Broas, H. Marttila, P. Poukka, K. Kuusisto, L. Pusa, K. Sammalkorpi, J. Dabek, O. P. Koivurova, M. Vahatalo, V. Moilanen & T. Widenius, (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. *Scand.J.Infect.Dis.* **40**: 702-708.

Mayfield, J. L., T. Leet, J. Miller & L. M. Mundy, (2000) Environmental control to reduce transmission of *Clostridium difficile*. *Clin.Infect.Dis.* **31**: 995-1000.

McDevitt, D. & T. J. Foster, (1995) Variation in the size of the repeat region of the fibrinogen receptor (clumping factor) of *Staphylococcus aureus* strains. *Microbiology* **141 (Pt 4)**: 937-943.

McDonald, L. C., G. E. Killgore, A. Thompson, R. C. Owens, Jr., S. V. Kazakova, S. P. Sambol, S. Johnson & D. N. Gerding, (2005) An epidemic, toxin gene-variant strain of *Clostridium difficile*. *N.Engl.J.Med.* **353**: 2433-2441.

McFarland, L. V., (2008) Antibiotic-associated diarrhea: epidemiology, trends and treatment. *Future Microbiol* **3**: 563-578.

McFarland, L. V., M. E. Mulligan, R. Y. Kwok & W. E. Stamm, (1989) Nosocomial acquisition of *Clostridium difficile* infection. *N.Engl.J.Med.* **320**: 204-210.
McFarland, L. V., C. M. Surawicz, M. Rubin, R. Fekety, G. W. Elmer & R. N. Greenberg, (1999) Recurrent *Clostridium difficile* disease: epidemiology and clinical characteristics. *Infect.Control Hosp.Epidemiol.* **20**: 43-50.
McPherson, S., C. J. Rees, R. Ellis, S. Soo & S. J. Panter, (2006) Intravenous immunoglobulin for the treatment of severe, refractory, and recurrent *Clostridium difficile* diarrhea. *Dis.Colon Rectum* **49**: 640-645.
Merrigan, M., S. Sambol, S. Johnson & D. N. Gerding, (2003a) Susceptibility of hamsters to human pathogenic *Clostridium difficile* strain B1 following clindamycin, ampicillin or ceftriaxone administration. *Anaerobe* **9**: 91-95.
Merrigan, M., S. Sambol, S. Johnson & D. N. Gerding, (2003b) Susceptibility of hamsters to human pathogenic *Clostridium difficile* strain B1 following clindamycin, ampicillin or ceftriaxone administration. *Anaerobe* **9**: 91-95.
Merrigan, M. M., S. P. Sambol, S. Johnson & D. N. Gerding, (2009) New approach to the management of *Clostridium difficile* infection: colonisation with non-toxigenic *C. difficile* during daily ampicillin or ceftriaxone administration. *Int.J.Antimicrob.Agents* **33 Suppl 1**: S46-50.
Miller, M., (2010) Fidaxomicin (OPT-80) for the treatment of *Clostridium difficile* infection. *Expert Opin Pharmacother* **11**: 1569-1578.
Mogg, G. A., R. H. George, D. Youngs, M. Johnson, H. Thompson, D. W. Burdon & M. R. Keighley, (1982) Randomized controlled trial of colestipol in antibiotic-associated colitis. *Br.J.Surg.* **69**: 137-139.
Mohr, G., D. Smith, M. Belfort & A. M. Lambowitz, (2000) Rules for DNA target-site recognition by a lactococcal group II intron enable retargeting of the intron to specific DNA sequences. *Genes Dev.* **14**: 559-573.
Mulligan, M. E., W. L. George, R. D. Rolfe & S. M. Finegold, (1980) Epidemiological aspects of *Clostridium difficile*-induced diarrhea and colitis. *Am.J.Clin.Nutr.* **33**: 2533-2538.
Musher, D. M., N. Logan, R. J. Hamill, H. L. Dupont, A. Lentnek, A. Gupta & J. F. Rossignol, (2006) Nitazoxanide for the treatment of *Clostridium difficile* colitis. *Clin.Infect.Dis.* **43**: 421-427.
Musher, D. M., N. Logan, V. Mehendiratta, N. A. Melgarejo, S. Garud & R. J. Hamill, (2007) *Clostridium difficile* colitis that fails conventional metronidazole therapy: response to nitazoxanide. *J.Antimicrob.Chemother.* **59**: 705-710.
Muto, C. A., M. K. Blank, J. W. Marsh, E. N. Vergis, M. M. O'Leary, K. A. Shutt, A. W. Pasculle, M. Pokrywka, J. G. Garcia, K. Posey, T. L. Roberts, B. A. Potoski, G. E. Blank, R. L. Simmons, P. Veldkamp, L. H. Harrison & D. L. Paterson, (2007) Control of an outbreak of infection with the hypervirulent *Clostridium difficile* BI strain in a university hospital using a comprehensive "bundle" approach. *Clin.Infect.Dis.* **45**: 1266-1273.
Muzzi, A., V. Masignani & R. Rappuoli, (2007) The pan-genome: towards a knowledge-based discovery of novel targets for vaccines and antibacterials. *Drug Discov Today* **12**: 429-439.
Naggie, S., J. Frederick, B. C. Pien, B. A. Miller, D. T. Provenzale, K. C. Goldberg & C. W. Woods, (2010) Community-associated *Clostridium difficile* infection: experience of a veteran affairs medical center in southeastern USA. *Infection*.
Nagy, E. & M. J. Dowzicky, (2010) In vitro activity of tigecycline and comparators against a European compilation of anaerobes collected as part of the

Tigecycline Evaluation and Surveillance Trial (TEST). *Scand.J.Infect.Dis.* **42**: 33-38.

Nakamura, S., K. Yamakawa, J. Izumi, S. Nakashio & S. Nishida, (1985) Germinability and heat resistance of spores of *Clostridium difficile* strains. *Microbiol.Immunol.* **29**: 113-118.

NHS, (2010) *Clostridium difficile*.

O'Brien, J., B., M. S. McCabe, V. Athie-Morales, G. S. McDonald, D. B. Ni Eidhin & D. P. Kelleher, (2005) Passive immunisation of hamsters against *Clostridium difficile* infection using antibodies to surface layer proteins. *FEMS Microbiol.Lett.* **246**: 199-205.

O'Brien, J. A., B. J. Lahue, J. J. Caro & D. M. Davidson, (2007) The emerging infectious challenge of *Clostridium difficile*-associated disease in Massachusetts hospitals: clinical and economic consequences. *Infect.Control Hosp.Epidemiol.* **28**: 1219-1227.

O'Connor, J. R., S. Johnson & D. N. Gerding, (2009) *Clostridium difficile* infection caused by the epidemic BI/NAP1/027 strain. *Gastroenterology* **136**: 1913-1924.

O'Connor, J. R., D. Lyras, K. A. Farrow, V. Adams, D. R. Powell, J. Hinds, J. K. Cheung & J. I. Rood, (2006) Construction and analysis of chromosomal *Clostridium difficile* mutants. *Mol.Microbiol.* **61**: 1335-1351.

Ochsner, U. A., S. J. Bell, A. L. O'Leary, T. Hoang, K. C. Stone, C. L. Young, I. A. Critchley & N. Janjic, (2009) Inhibitory effect of REP3123 on toxin and spore formation in *Clostridium difficile*, and in vivo efficacy in a hamster gastrointestinal infection model. *J.Antimicrob.Chemother.* **63**: 964-971.

Olson, M. M., C. J. Shanholtzer, J. T. Lee, Jr. & D. N. Gerding, (1994) Ten years of prospective *Clostridium difficile*-associated disease surveillance and treatment at the Minneapolis VA Medical Center, 1982-1991. *Infect.Control Hosp.Epidemiol.* **15**: 371-381.

Ossiprandi, M. C., M. Buttrini, E. Bottarelli & L. Zerbini, (2009) Preliminary molecular analysis of *Clostridium difficile* isolates from healthy horses in northern Italy. *Comp.Immunol.Microbiol.Infect.Dis.*

Otten, A. M., R. J. Reid-Smith, A. Fazil & J. S. Weese, (2010) Disease transmission model for community-associated *Clostridium difficile* infection. *Epidemiol.Infect.* **138**: 907-914.

Ozaki, E., H. Kato, H. Kita, T. Karasawa, T. Maegawa, Y. Koino, K. Matsumoto, T. Takada, K. Nomoto, R. Tanaka & S. Nakamura, (2004) *Clostridium difficile* colonization in healthy adults: transient colonization and correlation with enterococcal colonization. *J.Med.Microbiol.* **53**: 167-172.

Palmer, C., E. M. Bik, D. B. Digiulio, D. A. Relman & P. O. Brown, (2007) Development of the human infant intestinal microbiota. *PLoS Biol* **5**: e177.

Paredes, C. J., K. V. Alsaker & E. T. Papoutsakis, (2005) A comparative genomic view of clostridial sporulation and physiology. *Nat.Rev.Microbiol.* **3**: 969-978.

Paredes, C. J., I. Rigoutsos & E. T. Papoutsakis, (2004) Transcriptional organization of the *Clostridium* acetobutylicum genome. *Nucleic Acids Res.* **32**: 1973-1981.

Paulus, H., (2000) Protein splicing and related forms of protein autoprocessing. *Annu Rev Biochem* **69**: 447-496.

Pechine, S., C. Janoir, H. Boureau, A. Gleizes, N. Tsapis, S. Hoys, E. Fattal & A. Collignon, (2007) Diminished intestinal colonization by *Clostridium difficile* and immune response in mice after mucosal immunization with surface proteins of *Clostridium difficile*. *Vaccine* **25**: 3946-3954.

Pechine, S., C. Janoir & A. Collignon, (2005) Variability of *Clostridium difficile* surface proteins and specific serum antibody response in patients with *Clostridium difficile*-associated disease. *J.Clin.Microbiol.* **43**: 5018-5025.

Pelaez, T., L. Alcala, R. Alonso, A. Martin-Lopez, V. Garcia-Arias, M. Marin & E. Bouza, (2005) In vitro activity of ramoplanin against *Clostridium difficile*, including strains with reduced susceptibility to vancomycin or with resistance to metronidazole. *Antimicrob.Agents Chemother.* **49**: 1157-1159.

Pepin, J., L. Valiquette, M. E. Alary, P. Villemure, A. Pelletier, K. Forget, K. Pepin & D. Chouinard, (2004) *Clostridium difficile*-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. *CMAJ* **171**: 466-472.

Pepin, J., L. Valiquette, S. Gagnon, S. Routhier & I. Brazeau, (2007) Outcomes of *Clostridium difficile*-associated disease treated with metronidazole or vancomycin before and after the emergence of NAP1/027. *Am.J.Gastroenterol.* **102**: 2781-2788.

Perego, M. & J. A. Hoch, (2008) Commingling regulatory systems following acquisition of virulence plasmids by *Bacillus* anthracis. *Trends Microbiol.* **16**: 215-221.

Perelle, S., M. Gibert, P. Bourlioux, G. Corthier & M. R. Popoff, (1997) Production of a complete binary toxin (actin-specific ADP-ribosyltransferase) by *Clostridium difficile* CD196. *Infect.Immun.* **65**: 1402-1407.

Perrin, J., C. Buogo, A. Gallusser, A. P. Burnens & J. Nicolet, (1993a) Intestinal carriage of *Clostridium difficile* in neonate dogs. *Zentralbl.Veterinarmed.B* **40**: 222-226.

Perrin, J., I. Cosmetatos, A. Gallusser, L. Lobsiger, R. Straub & J. Nicolet, (1993b) *Clostridium difficile* associated with typhlocolitis in an adult horse. *J.Vet.Diagn.Invest.* **5**: 99-101.

Perry, C., R. Marshall & E. Jones, (2001) Bacterial contamination of uniforms. *J.Hosp.Infect.* **48**: 238-241.

Perutka, J., W. Wang, D. Goerlitz & A. M. Lambowitz, (2004) Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. *J.Mol.Biol.* **336**: 421-439.

Peterson, J., S. Garges, M. Giovanni, P. McInnes, L. Wang, J. A. Schloss, V. Bonazzi, J. E. McEwen, K. A. Wetterstrand, C. Deal, C. C. Baker, V. Di Francesco, T. K. Howcroft, R. W. Karp, R. D. Lunsford, C. R. Wellington, T. Belachew, M. Wright, C. Giblin, H. David, M. Mills, R. Salomon, C. Mullins, B. Akolkar, L. Begg, C. Davis, L. Grandison, M. Humble, J. Khalsa, A. R. Little, H. Peavy, C. Pontzer, M. Portnoy, M. H. Sayre, P. Starke-Reed, S. Zakhari, J. Read, B. Watson & M. Guyer, (2009) The NIH Human Microbiome Project. *Genome Res.* **19**: 2317-2323.

Pillai, A. & R. Nelson, (2008) Probiotics for treatment of *Clostridium difficile*-associated colitis in adults. *Cochrane Database Syst.Rev.*: CD004611.

Pituch, H., (2009) *Clostridium difficile* is no longer just a nosocomial infection or an infection of adults. *Int.J.Antimicrob.Agents* **33 Suppl 1**: S42-45.

Popoff, M. R., E. J. Rubin, D. M. Gill & P. Boquet, (1988) Actin-specific ADP-ribosyltransferase produced by a *Clostridium difficile* strain. *Infect Immun* **56**: 2299-2306.

Purdy, D., T. A. O'Keeffe, M. Elmore, M. Herbert, A. McLeod, M. Bokori-Brown, A. Ostrowski & N. P. Minton, (2002) Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. *Mol Microbiol* **46**: 439-452.

Ramirez, N., M. Liggins & E. Abel-Santos, (2010) Kinetic evidence for the presence of putative germination receptors in *C. difficile* spores. *J.Bacteriol.*

Razaq, N., S. Sambol, K. Nagaro, W. Zukowski, A. Cheknis, S. Johnson & D. N. Gerding, (2007) Infection of hamsters with historical and epidemic BI types of *Clostridium difficile*. *J.Infect.Dis.* **196**: 1813-1819.

Rea, M. C., A. Dobson, O. O'Sullivan, F. Crispie, F. Fouhy, P. D. Cotter, F. Shanahan, B. Kiely, C. Hill & R. P. Ross, (2010a) Microbes and Health Sackler Colloquium: Effect of broad- and narrow-spectrum antimicrobials on *Clostridium difficile* and microbial diversity in a model of the distal colon. *Pro*C.*Natl.Acad.Sci.USA*.

Rea, M. C., C. S. Sit, E. Clayton, P. M. O'Connor, R. M. Whittal, J. Zheng, J. C. Vederas, R. P. Ross & C. Hill, (2010b) Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against *Clostridium difficile*. *Pro*C.*Natl.Acad.Sci.USA* **107**: 9352-9357.

Rehg, J. E. & Y. S. Lu, (1981) *Clostridium difficile* colitis in a rabbit following antibiotic therapy for pasteurellosis. *J.Am.Vet.Med.Asso*C. **179**: 1296-1297.

Rehg, J. E. & S. P. Pakes, (1981) *Clostridium difficile* antitoxin neutralization of cecal toxin(s) from guinea pigs with penicillin-associated colitis. *Lab.Anim.Sci.* **31**: 156-160.

Reller, M. E., R. C. Alcabasa, C. A. Lema & K. C. Carroll, (2010) Comparison of two rapid assays for *Clostridium difficile* Common antigen and a *C. difficile* toxin A/B assay with the cell culture neutralization assay. *Am J Clin Pathol* **133**: 107-109.

Riley, T. V., J. E. Adams, G. L. O'Neill & R. A. Bowman, (1991) Gastrointestinal carriage of *Clostridium difficile* in cats and dogs attending veterinary clinics. *Epidemiol.Infect.* **107**: 659-665.

Rivera, E. V. & S. Woods, (2003) Prevalence of asymptomatic *Clostridium difficile* colonization in a nursing home population: a cross-sectional study. *J Gend Specif Med* **6**: 27-30.

Rodriguez-Palacios, A., R. J. Reid-Smith, H. R. Staempfli & J. S. Weese, (2010) *Clostridium difficile* survives minimal temperature recommended for cooking ground meats. *Anaerobe*.

Rodriguez-Palacios, A., H. R. Stampfli, T. Duffield, A. S. Peregrine, L. A. Trotz-Williams, L. G. Arroyo, J. S. Brazier & J. S. Weese, (2006) *Clostridium difficile* PCR ribotypes in calves, Canada. *Emerg.Infect.Dis.* **12**: 1730-1736.

Rodriguez-Palacios, A., H. R. Stampfli, M. Stalker, T. Duffield & J. S. Weese, (2007) Natural and experimental infection of neonatal calves with *Clostridium difficile*. *Vet.Microbiol.* **124**: 166-172.

Rolfe, R. D. & J. P. Iaconis, (1983) Intestinal colonization of infant hamsters with *Clostridium difficile*. *Infect.Immun.* **42**: 480-486.

Ruby, R., K. G. Magdesian & P. H. Kass, (2009) Comparison of clinical, microbiologic, and clinicopathologic findings in horses positive and negative for *Clostridium difficile* infection. *J.Am.Vet.Med.Asso*C. **234**: 777-784.

Rupnik, M., (2010) *Clostridium difficile* Toxinotyping. *Methods Mol.Biol.* **646**: 67-76.

Rupnik, M., V. Avesani, M. Janc, C. von Eichel-Streiber & M. Delmee, (1998) A novel toxinotyping scheme and correlation of toxinotypes with serogroups of *Clostridium difficile* isolates. *J.Clin.Microbiol.* **36**: 2240-2247.

Rupnik, M., B. Dupuy, N. F. Fairweather, D. N. Gerding, S. Johnson, I. Just, D. M. Lyerly, M. R. Popoff, J. I. Rood, A. L. Sonenshein, M. Thelestam, B. W.

Wren, T. D. Wilkins & C. von Eichel-Streiber, (2005) Revised nomenclature of *Clostridium difficile* toxins and associated genes. *J.Med.Microbiol.* **54**: 113-117.

Ryan, J., C. Murphy, C. Twomey, R. Paul Ross, M. C. Rea, J. MacSharry, B. Sheil & F. Shanahan, (2010) Asymptomatic carriage of *Clostridium difficile* in an Irish continuing care institution for the elderly: prevalence and characteristics. *Ir J Med Sci* **179**: 245-250.

Salcedo, J., S. Keates, C. Pothoulakis, M. Warny, I. Castagliuolo, J. T. LaMont & C. P. Kelly, (1997) Intravenous immunoglobulin therapy for severe *Clostridium difficile* colitis. *Gut* **41**: 366-370.

Sambol, S. P., M. M. Merrigan, J. K. Tang, S. Johnson & D. N. Gerding, (2002) Colonization for the prevention of *Clostridium difficile* disease in hamsters. *J.Infect.Dis.* **186**: 1781-1789.

Sambol, S. P., J. K. Tang, M. M. Merrigan, S. Johnson & D. N. Gerding, (2001) Infection of hamsters with epidemiologically important strains of *Clostridium difficile*. *J.Infect.Dis.* **183**: 1760-1766.

Samore, M. H., L. Venkataraman, P. C. DeGirolami, R. D. Arbeit & A. W. Karchmer, (1996) Clinical and molecular epidemiology of sporadic and clustered cases of nosocomial *Clostridium difficile* diarrhea. *Am.J.Med.* **100**: 32-40.

Sara, M., (2001) Conserved anchoring mechanisms between crystalline cell surface S-layer proteins and secondary cell wall polymers in Gram-positive bacteria? *Trends Microbiol.* **9**: 47-49; discussion 49-50.

Schwan, C., B. Stecher, T. Tzivelekidis, M. van Ham, M. Rohde, W. D. Hardt, J. Wehland & K. Aktories, (2009) *Clostridium difficile* toxin CDT induces formation of microtubule-based protrusions and increases adherence of bacteria. *PLoS Pathog* **5**: e1000626.

Schwarz-Linek, U., M. Hook & J. R. Potts, (2004) The molecular basis of fibronectin-mediated bacterial adherence to host cells. *Mol.Microbiol.* **52**: 631-641.

Sebaihia, M., B. W. Wren, P. Mullany, N. F. Fairweather, N. Minton, R. Stabler, N. R. Thomson, A. P. Roberts, A. M. Cerdeno-Tarraga, H. Wang, M. T. Holden, A. Wright, C. Churcher, M. A. Quail, S. Baker, N. Bason, K. Brooks, T. Chillingworth, A. Cronin, P. Davis, L. Dowd, A. Fraser, T. Feltwell, Z. Hance, S. Holroyd, K. Jagels, S. Moule, K. Mungall, C. Price, E. Rabbinowitsch, S. Sharp, M. Simmonds, K. Stevens, L. Unwin, S. Whithead, B. Dupuy, G. Dougan, B. Barrell & J. Parkhill, (2006) The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nat.Genet.* **38**: 779-786.

Shah, D., M. D. Dang, R. Hasbun, H. L. Koo, Z. D. Jiang, H. L. DuPont & K. W. Garey, (2010) *Clostridium difficile* infection: update on emerging antibiotic treatment options and antibiotic resistance. *Expert Rev Anti Infect Ther* **8**: 555-564.

Shapiro, J. A., (1995) The significances of bacterial colony patterns. *Bioessays* **17**: 597-607.

Sharp, S. E., W. M. Ivie, M. R. Buckles, D. M. Coover, J. C. Pohl & P. A. Hatcher, (2009) A simple 3-step algorithm for improved laboratory detection of *Clostridium difficile* toxin without the need for tissue culture cytotoxicity neutralization assays. *Diagn.Microbiol.Infect.Dis.* **64**: 344-346.

Shim, J. K., S. Johnson, M. H. Samore, D. Z. Bliss & D. N. Gerding, (1998) Primary symptomless colonisation by *Clostridium difficile* and decreased risk of subsequent diarrhoea. *Lancet* **351**: 633-636.

Silverman, M. S., I. Davis & D. R. Pillai, (2010) Success of self-administered home fecal transplantation for chronic *Clostridium difficile* infection. *Clin.Gastroenterol.Hepatol.* **8**: 471-473.

Sjodahl, J., (1977) Repetitive sequences in protein A from *Staphylococcus aureus*. Arrangement of five regions within the protein, four being highly homologous and Fc-binding. *Eur.J.Biochem.* **73**: 343-351.

Smith, A., (2005) Outbreak of *Clostridium difficile* infection in an English hospital linked to hypertoxin-producing strains in Canada and the US. *Euro Surveill* **10**: E050630 050632.

Smith, L. D. & E. O. King, (1962) Occurrence of *Clostridium difficile* in infections of man. *J.Bacteriol.* **84**: 65-67.

Snyder, (1937) Further studies on *Bacillus difficilis*. *J.Infect.Dis.* **60**.

Songer, J. G. & M. A. Anderson, (2006) *Clostridium difficile*: an important pathogen of food animals. *Anaerobe* **12**: 1-4.

Songer, J. G., H. T. Trinh, G. E. Killgore, A. D. Thompson, L. C. McDonald & B. M. Limbago, (2009) *Clostridium difficile* in retail meat products, USA, 2007. *Emerg.Infect.Dis.* **15**: 819-821.

Sorg, J. A. & S. S. Dineen, (2009) Laboratory maintenance of *Clostridium difficile*. *Curr Protoc Microbiol* **Chapter 9**: Unit9A 1.

Sorg, J. A. & A. L. Sonenshein, (2008) Bile salts and glycine as cogerminants for *Clostridium difficile* spores. *J.Bacteriol.* **190**: 2505-2512.

Sougioultzis, S., L. Kyne, D. Drudy, S. Keates, S. Maroo, C. Pothoulakis, P. J. Giannasca, C. K. Lee, M. Warny, T. P. Monath & C. P. Kelly, (2005) *Clostridium difficile* toxoid vaccine in recurrent *C. difficile*-associated diarrhea. *Gastroenterology* **128**: 764-770.

Spigaglia, P., F. Barbanti, P. Mastrantonio, J. S. Brazier, F. Barbut, M. Delmee, E. Kuijper & I. R. Poxton, (2008) Fluoroquinolone resistance in *Clostridium difficile* isolates from a prospective study of *C. difficile* infections in Europe. *J.Med.Microbiol.* **57**: 784-789.

Stabler, R. A., D. N. Gerding, J. G. Songer, D. Drudy, J. S. Brazier, H. T. Trinh, A. A. Witney, J. Hinds & B. W. Wren, (2006a) Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. *Journal of Bacteriology* **188**: 7297-7305.

Stabler, R. A., D. N. Gerding, J. G. Songer, D. Drudy, J. S. Brazier, H. T. Trinh, A. A. Witney, J. Hinds & B. W. Wren, (2006b) Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. *J.Bacteriol.* **188**: 7297-7305.

Stabler, R. A., M. He, L. Dawson, M. Martin, E. Valiente, C. Corton, T. D. Lawley, M. Sebaihia, M. A. Quail, G. Rose, D. N. Gerding, M. Gibert, M. R. Popoff, J. Parkhill, G. Dougan & B. W. Wren, (2009) Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biol.* **10**: R102.

Steele, J., H. Feng, N. Parry & S. Tzipori, (2010) Piglet models of acute or chronic *Clostridium difficile* illness. *J.Infect.Dis.* **201**: 428-434.

Stubbs, S., M. Rupnik, M. Gibert, J. Brazier, B. Duerden & M. Popoff, (2000) Production of actin-specific ADP-ribosyltransferase (binary toxin) by strains of *Clostridium difficile*. *FEMS Microbiol.Lett.* **186**: 307-312.

Stubbs, S. L., J. S. Brazier, G. L. O'Neill & B. I. Duerden, (1999) PCR targeted to the 16S-23S rRNA gene intergenic spacer region of *Clostridium difficile* and construction of a library consisting of 116 different PCR ribotypes. *J.Clin.Microbiol.* **37**: 461-463.

Swindells, J., N. Brenwald, N. Reading & B. Oppenheim, (2010) Evaluation of diagnostic tests for *Clostridium difficile* infection. *J.Clin.Microbiol.* **48**: 606-608.

Tajima, N., F. Kawai, S. Y. Park & J. R. Tame, (2010) A novel intein-like autoproteolytic mechanism in autotransporter proteins. *J.Mol.Biol.* **402**: 645-656.

Tasteyre, A., M. C. Barc, A. Collignon, H. Boureau & T. Karjalainen, (2001a) Role of FliC and FliD flagellar proteins of *Clostridium difficile* in adherence and gut colonization. *Infect.Immun.* **69**: 7937-7940.

Tasteyre, A., M. C. Barc, T. Karjalainen, P. Dodson, S. Hyde, P. Bourlioux & P. Borriello, (2000a) A *Clostridium difficile* gene encoding flagellin. *Microbiology* **146 (Pt 4)**: 957-966.

Tasteyre, A., T. Karjalainen, V. Avesani, M. Delmee, A. Collignon, P. Bourlioux & M. C. Barc, (2000b) Phenotypic and genotypic diversity of the flagellin gene (*fliC*) among *Clostridium difficile* isolates from different serogroups. *J.Clin.Microbiol.* **38**: 3179-3186.

Tasteyre, A., T. Karjalainen, V. Avesani, M. Delmee, A. Collignon, P. Bourlioux & M. C. Barc, (2001b) Molecular characterization of *fliD* gene encoding flagellar cap and its expression among *Clostridium difficile* isolates from different serogroups. *J.Clin.Microbiol.* **39**: 1178-1183.

Tedesco, F. J., R. W. Barton & D. H. Alpers, (1974) Clindamycin-associated colitis. A prospective study. *Ann.Intern.Med.* **81**: 429-433.

Thomas, C. M. & K. M. Nielsen, (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat.Rev.Microbiol.* **3**: 711-721.

Torres, J. F., D. M. Lyerly, J. E. Hill & T. P. Monath, (1995) Evaluation of formalin-inactivated *Clostridium difficile* vaccines administered by parenteral and mucosal routes of immunization in hamsters. *Infect.Immun.* **63**: 4619-4627.

Twine, S. M., C. W. Reid, A. Aubry, D. R. McMullin, K. M. Fulton, J. Austin & S. M. Logan, (2009) Motility and flagellar glycosylation in *Clostridium difficile*. *J.Bacteriol.* **191**: 7050-7062.

Ulett, G. C., R. I. Webb & M. A. Schembri, (2006) Antigen-43-mediated autoaggregation impairs motility in *Escherichia coli*. *Microbiology* **152**: 2101-2110.

Underwood, S., S. Guan, V. Vijayasubhash, S. D. Baines, L. Graham, R. J. Lewis, M. H. Wilcox & K. Stephenson, (2009) Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. *J.Bacteriol.* **191**: 7296-7305.

Valiquette, L., B. Cossette, M. P. Garant, H. Diab & J. Pepin, (2007) Impact of a reduction in the use of high-risk antibiotics on the course of an epidemic of *Clostridium difficile*-associated disease caused by the hypervirulent NAP1/027 strain. *Clin.Infect.Dis.* **45 Suppl 2**: S112-121.

van den Berg, R. J., I. Schaap, K. E. Templeton, C. H. Klaassen & E. J. Kuijper, (2007) Typing and subtyping of *Clostridium difficile* isolates by using multiple-locus variable-number tandem-repeat analysis. *J.Clin.Microbiol.* **45**: 1024-1028.

van der Woude, M. W., (2006) Re-examining the role and random nature of phase variation. *FEMS Microbiol.Lett.* **254**: 190-197.

van der Woude, M. W. & A. J. Baumler, (2004) Phase and antigenic variation in bacteria. *Clin.Microbiol.Rev.* **17**: 581-611, table of contents.

van Dissel, J. T., N. de Groot, C. M. Hensgens, S. Numan, E. J. Kuijper, P. Veldkamp & J. van 't Wout, (2005) Bovine antibody-enriched whey to aid in

the prevention of a relapse of *Clostridium difficile*-associated diarrhoea: preclinical and preliminary clinical data. *J.Med.Microbiol.* **54**: 197-205.
van Nood, E., P. Speelman, E. J. Kuijper & J. J. Keller, (2009) Struggling with recurrent *Clostridium difficile* infections: is donor faeces the solution? *Euro Surveill* **14**.
Varga, J. J., V. Nguyen, D. K. O'Brien, K. Rodgers, R. A. Walker & S. B. Melville, (2006) Type IV pili-dependent gliding motility in the Gram-positive pathogen *Clostridium perfringens* and other clostridia. *Mol.Microbiol.* **62**: 680-694.
Viscidi, R., B. E. Laughon, R. Yolken, P. Bo-Linn, T. Moench, R. W. Ryder & J. G. Bartlett, (1983) Serum antibody response to toxins A and B of *Clostridium difficile*. *J.Infect.Dis.* **148**: 93-100.
Von Abercron, S. M., F. Karlsson, G. T. Wigh, M. Wierup & K. Krovacek, (2009) Low occurrence of *Clostridium difficile* in retail ground meat in Sweden. *J.Food Prot.* **72**: 1732-1734.
von Eichel-Streiber, C., P. Boquet, M. Sauerborn & M. Thelestam, (1996) Large clostridial cytotoxins - a family of glycosyltransferases modifying small GTP-binding proteins. *Trends Microbiol* **4**: 375-382.
von Eichel-Streiber, C., R. Laufenberg-Feldmann, S. Sartingen, J. Schulze & M. Sauerborn, (1992) Comparative sequence analysis of the *Clostridium difficile* toxins A and B. *Mol.Gen.Genet.* **233**: 260-268.
Voth, D. E. & J. D. Ballard, (2005) *Clostridium difficile* toxins: mechanism of action and role in disease. *Clin.Microbiol.Rev.* **18**: 247-263.
Waligora, A. J., M. C. Barc, P. Bourlioux, A. Collignon & T. Karjalainen, (1999) *Clostridium difficile* cell attachment is modified by environmental factors. *Appl.Environ.Microbiol.* **65**: 4234-4238.
Waligora, A. J., C. Hennequin, P. Mullany, P. Bourlioux, A. Collignon & T. Karjalainen, (2001a) Characterization of a cell surface protein of *Clostridium difficile* with adhesive properties. *Infect.Immun.* **69**: 2144-2153.
Waligora, A. J., C. Hennequin, P. Mullany, P. Bourlioux, A. Collignon & T. Karjalainen, (2001b) Characterization of a cell surface protein of *Clostridium difficile* with adhesive properties. *Infect Immun* **69**: 2144-2153.
Walker, K. J., S. S. Gilliland, K. Vance-Bryan, J. A. Moody, A. J. Larsson, J. C. Rotschafer & D. R. Guay, (1993) *Clostridium difficile* colonization in residents of long-term care facilities: prevalence and risk factors. *J Am Geriatr Soc* **41**: 940-946.
Washio, T., J. Sasayama & M. Tomita, (1998) Analysis of complete genomes suggests that many prokaryotes do not rely on hairpin formation in transcription termination. *Nucleic Acids Res.* **26**: 5456-5463.
Weber, D. J., W. A. Rutala, M. B. Miller, K. Huslage & E. Sickbert-Bennett, (2010) Role of hospital surfaces in the transmission of emerging health care-associated pathogens: norovirus, *Clostridium difficile*, and Acinetobacter species. *Am.J.Infect.Control* **38**: S25-33.
Weese, J. S., B. P. Avery, J. Rousseau & R. J. Reid-Smith, (2009) Detection and enumeration of *Clostridium difficile* spores in retail beef and pork. *Appl.Environ.Microbiol.* **75**: 5009-5011.
Weese, J. S., R. J. Reid-Smith, B. P. Avery & J. Rousseau, (2010) Detection and characterization of *Clostridium difficile* in retail chicken. *Lett.Appl.Microbiol.* **50**: 362-365.
Weiss, K., (2009) Toxin-binding treatment for *Clostridium difficile*: a review including reports of studies with tolevamer. *Int.J.Antimicrob.Agents* **33**: 4-7.

Wheeldon, L. J., T. Worthington, A. C. Hilton, T. S. Elliott & P. A. Lambert, (2008a) Physical and chemical factors influencing the germination of *Clostridium difficile* spores. *J.Appl.Microbiol.* **105**: 2223-2230.

Wheeldon, L. J., T. Worthington, P. A. Lambert, A. C. Hilton, C. J. Lowden & T. S. Elliott, (2008b) Antimicrobial efficacy of copper surfaces against spores and vegetative cells of *Clostridium difficile*: the germination theory. *J.Antimicrob.Chemother.* **62**: 522-525.

Wilcox, M. H., (2004) Descriptive study of intravenous immunoglobulin for the treatment of recurrent *Clostridium difficile* diarrhoea. *J.Antimicrob.Chemother.* **53**: 882-884.

Wilcox, M. H., (2007a) Diagnosis of *Clostridium difficile*-associated diarrhea and odor. *Clin.Infect.Dis.* **45**: 1110.

Wilcox, M. H., (2007b) Evidence for low risk of *Clostridium difficile* infection associated with tigecycline. *Clin.Microbiol.Infect.* **13**: 949-952.

Wilcox, M. H., J. G. Cunniffe, C. Trundle & C. Redpath, (1996) Financial burden of hospital-acquired *Clostridium difficile* infection. *J.Hosp.Infect.* **34**: 23-30.

Wilcox, M. H., W. N. Fawley, N. Wigglesworth, P. Parnell, P. Verity & J. Freeman, (2003) Comparison of the effect of detergent versus hypochlorite cleaning on environmental contamination and incidence of *Clostridium difficile* infection. *J.Hosp.Infect.* **54**: 109-114.

Wilcox, M. H., J. Freeman, W. Fawley, S. MacKinlay, A. Brown, K. Donaldson & O. Corrado, (2004) Long-term surveillance of cefotaxime and piperacillin-tazobactam prescribing and incidence of *Clostridium difficile* diarrhoea. *J.Antimicrob.Chemother.* **54**: 168-172.

Wilcox, M. H., L. Mooney, R. Bendall, C. D. Settle & W. N. Fawley, (2008) A case-control study of community-associated *Clostridium difficile* infection. *J.Antimicrob.Chemother.* **62**: 388-396.

Wilson, K. H., (1983) Efficiency of various bile salt preparations for stimulation of *Clostridium difficile* spore germination. *J.Clin.Microbiol.* **18**: 1017-1019.

Wilson, K. H. & R. Freter, (1986) Interaction of *Clostridium difficile* and *Escherichia coli* with microfloras in continuous-flow cultures and gnotobiotic mice. *Infect.Immun.* **54**: 354-358.

Wilson, K. H., M. J. Kennedy & F. R. Fekety, (1982) Use of sodium taurocholate to enhance spore recovery on a medium selective for *Clostridium difficile*. *J.Clin.Microbiol.* **15**: 443-446.

Wilson, K. H. & J. N. Sheagren, (1983) Antagonism of toxigenic *Clostridium difficile* by nontoxigenic *C. difficile*. *J.Infect.Dis.* **147**: 733-736.

Wilson, K. H., J. N. Sheagren, R. Freter, L. Weatherbee & D. Lyerly, (1986) Gnotobiotic models for study of the microbial ecology of *Clostridium difficile* and *Escherichia coli*. *J.Infect.Dis.* **153**: 547-551.

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

Wistrom, J., S. R. Norrby, E. B. Myhre, S. Eriksson, G. Granstrom, L. Lagergren, G. Englund, C. E. Nord & B. Svenungsson, (2001) Frequency of antibiotic-associated diarrhoea in 2462 antibiotic-treated hospitalized patients: a prospective study. *J.Antimicrob.Chemother.* **47**: 43-50.

Wolfe, K. H., (2001) Yesterday's polyploids and the mystery of diploidization. *Nat Rev Genet* **2**: 333-341.

Wren, M. W., R. Kinson, M. Sivapalan, M. Shemko & N. R. Shetty, (2009a) Detection of *Clostridium difficile* infection: a suggested laboratory diagnostic algorithm. *Br.J.Biomed.Sci.* **66**: 175-179.

Wren, M. W., M. Sivapalan, R. Kinson & N. R. Shetty, (2009b) Laboratory diagnosis of *Clostridium difficile* infection. An evaluation of tests for faecal toxin, glutamate dehydrogenase, lactoferrin and toxigenic culture in the diagnostic laboratory. *Br.J.Biomed.Sci.* **66**: 1-5.

Wright, A., D. Drudy, L. Kyne, K. Brown & N. F. Fairweather, (2008a) Immunoreactive cell wall proteins of *Clostridium difficile* identified by human sera. *J.Med.Microbiol.* **57**: 750-756.

Wright, A., D. Drudy, L. Kyne, B. K. & N. F. Fairweather, (2008b) Immunoreactive cell wall proteins of *Clostridium difficile* identified by human sera. *J. Med. Microbiol* **57**: 750-756.

Wright, A., R. Wait, S. Begum, B. Crossett, J. Nagy, K. Brown & N. Fairweather, (2005) Proteomic analysis of cell surface proteins from *Clostridium difficile*. *Proteomics* **5**: 2443-2452.

Xu, Q., D. Buckley, C. Guan & H. C. Guo, (1999) Structural insights into the mechanism of intramolecular proteolysis. *Cell* **98**: 651-661.

Yaeger, M., N. Funk & L. Hoffman, (2002) A survey of agents associated with neonatal diarrhea in Iowa swine including *Clostridium difficile* and porcine reproductive and respiratory syndrome virus. *J.Vet.Diagn.Invest.* **14**: 281-287.

Yager, T. D. & P. H. von Hippel, (1991) A thermodynamic analysis of RNA transcript elongation and termination in *Escherichia coli*. *Biochemistry* **30**: 1097-1118.

Yassin, S. F., (2009) Pseudomembranous Colitis, Surgical Treatment. *eMedicine General Surgery*.

Yoon, S. S. & L. J. Brandt, (2010) Treatment of Refractory/Recurrent *C. difficile*-associated Disease by Donated Stool Transplanted Via Colonoscopy: A Case Series of 12 Patients. *J.Clin.Gastroenterol.*

Young, K. W., I. C. Munro, S. L. Taylor, P. Veldkamp & J. T. van Dissel, (2007) The safety of whey protein concentrate derived from the milk of cows immunized against *Clostridium difficile*. *Regul Toxicol Pharmacol* **47**: 317-326.

Zar, F. A., S. R. Bakkanagari, K. M. Moorthi & M. B. Davis, (2007) A comparison of vancomycin and metronidazole for the treatment of *Clostridium difficile*-associated diarrhea, stratified by disease severity. *Clin.Infect.Dis.* **45**: 302-307.

Zerey, M., B. L. Paton, A. E. Lincourt, K. S. Gersin, K. W. Kercher & B. T. Heniford, (2007) The burden of *Clostridium difficile* in surgical patients in the United States. *Surg Infect (Larchmt)* **8**: 557-566.

Zhong, J., M. Karberg & A. M. Lambowitz, (2003) Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Res.* **31**: 1656-1664.

Zhou, H. X., (2003) Quantitative account of the enhanced affinity of two linked scFvs specific for different epitopes on the same antigen. *J.Mol.Biol.* **329**: 1-8.

Zidaric, V., M. Zemljic, S. Janezic, A. Kocuvan & M. Rupnik, (2008) High diversity of *Clostridium difficile* genotypes isolated from a single poultry farm producing replacement laying hens. *Anaerobe* **14**: 325-327.

Zieg, J., M. Silverman, M. Hilmen & M. Simon, (1977) Recombinational switch for gene expression. *Science* **196**: 170-172.

Zuker, M., (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**: 3406-3415.