Hypervirulent Clostridium Difficile Strains: Adherence, Toxin Production and Sporulation

Michelle Marie Merrigan

Loyola University Chicago

Recommended Citation

http://ecommons.luc.edu/luc_diss/178

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 2010 Michelle Marie Merrigan
LOYOLA UNIVERSITY CHICAGO

HYPERVIRULENT CLOSTRIDIUM DIFFICILE STRAINS:
ADHERENCE, TOXIN PRODUCTION AND SPORULATION

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

PROGRAM IN MICROBIOLOGY AND IMMUNOLOGY

BY

MICHELLE MARIE MERRIGAN

CHICAGO, IL

December 2010
ACKNOWLEDGEMENTS

I would first like to thank my mentor, Gayatri Vedantam. Her optimism, her work ethic and encouragement provide an excellent model of how to be a good scientist. More importantly, her selflessness, and the kindness and mutual respect with which she treats everyone serve as a model for how to be a better human being.

Second, I would like to thank my committee members, Adam Driks, David Keating, Chris Wiethoff, and Nancy Freitag for their active support, commitment and mentorship. They have each gone over and above the usual responsibilities. I hope to have the opportunity to someday give back as much as they gave me.

I would like to thank Mai Nguyen and Sarah Kralicek, for their friendship and endless help during long days in the lab under challenging conditions.

I would also like to thank Dr. Knight and the entire Microbiology and Immunology department, for making my graduate experience extra-ordinary.

I would also like to thank my many collaborators, without whom much of this work would not have been possible. I would like to thank Dale Gerding and Stuart Johnson and the members of their laboratories. When I worked for them as a technician, they encouraged my professional development, and gave me opportunities to learn and write. There are many helpful people who have passed though this research group whom I would like to thank, but particularly Farida Siddiqui and Jennifer O’Connor for their advice, inspiration and humor. My collaborators at UIC were so helpful and generous,
and I cannot thank them enough: Vish Viswanathan, Sia Koutsouris, Jenny and Bryan Roxas. Thanks to Anil Venugopal, for his invaluable hard work on the toxin project. Fellow student and current collaborator Mike Mallozzi also provided great advice and support.

I would like to thank my family and friends for their love and support. I would like to thank Carl Ebeling for his unique talent for going out of town whenever I needed to work especially hard, but really, I am grateful for his love and encouragement.

Last but not least, I would like to thank Susan Sambol, my friend and the manager of the Gerding laboratory. In 1996, she took a chance on hiring me, and had faith that someone with a degree in Animal Behavior Psychology could contribute something to a Microbiology lab. If it were not for her investment, not to mention the friendship, love and support of her and her whole family over the years, I would not have been able to reach this goal.
For Sue, and all of my mentors
The most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka!' but 'That's funny...'

Isaac Asimov
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACKNOWLEDGEMENTS</strong></td>
<td>iii</td>
</tr>
<tr>
<td><strong>LIST OF TABLES</strong></td>
<td>x</td>
</tr>
<tr>
<td><strong>LIST OF FIGURES</strong></td>
<td>xi</td>
</tr>
<tr>
<td><strong>LIST OF ABBREVIATIONS</strong></td>
<td>xiii</td>
</tr>
<tr>
<td><strong>CHAPTER ONE: LITERATURE REVIEW</strong></td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Epidemiology of CDI: Causes, Incidence and Control</td>
<td>3</td>
</tr>
<tr>
<td>Disease Characteristics and History</td>
<td>3</td>
</tr>
<tr>
<td>Epidemiology and Antibiotics</td>
<td>4</td>
</tr>
<tr>
<td>Risk Factors for CDI</td>
<td>6</td>
</tr>
<tr>
<td>Immunity and Vaccines</td>
<td>7</td>
</tr>
<tr>
<td>Costs and Prevention</td>
<td>8</td>
</tr>
<tr>
<td>Pathogenesis</td>
<td>10</td>
</tr>
<tr>
<td>Toxins</td>
<td>10</td>
</tr>
<tr>
<td>The Structure and Regulation of the Pathogenicity Locus</td>
<td>13</td>
</tr>
<tr>
<td>Mechanism of Toxin Action and Host Responses</td>
<td>20</td>
</tr>
<tr>
<td>Binary Toxin and Other Virulence Factors</td>
<td>25</td>
</tr>
<tr>
<td>Sporulation</td>
<td>26</td>
</tr>
<tr>
<td>The Emergence of Hypervirulent Strains</td>
<td>31</td>
</tr>
<tr>
<td>Colonization and Adherence</td>
<td>36</td>
</tr>
<tr>
<td>Colonization and Animal Models</td>
<td>36</td>
</tr>
<tr>
<td>Adherence</td>
<td>38</td>
</tr>
<tr>
<td><strong>CHAPTER TWO: MATERIALS AND EXPERIMENTAL METHODS</strong></td>
<td>45</td>
</tr>
<tr>
<td><em>C. difficile</em> Strains and Media</td>
<td>45</td>
</tr>
<tr>
<td><em>E. coli</em> Strains and Media</td>
<td>47</td>
</tr>
<tr>
<td>Anaerobic Bacterial Adherence Assay</td>
<td>48</td>
</tr>
<tr>
<td>Immunofluorescence Microscopy</td>
<td>48</td>
</tr>
<tr>
<td>Total Surface Protein Interference Assay</td>
<td>50</td>
</tr>
<tr>
<td>Antibody Interference Assay</td>
<td>50</td>
</tr>
<tr>
<td>Total Soluble Protein Isolation</td>
<td>50</td>
</tr>
<tr>
<td>S-layer Protein Extraction</td>
<td>51</td>
</tr>
<tr>
<td>Protein Quantitation</td>
<td>51</td>
</tr>
<tr>
<td>DNA Sequencing of <em>slpA</em></td>
<td>52</td>
</tr>
<tr>
<td>Immunodetection</td>
<td>53</td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER THREE: EXPERIMENTAL RESULTS</strong></td>
</tr>
<tr>
<td></td>
<td>The Adherence of CD to Host Cells and the Role of SlpA in this</td>
</tr>
<tr>
<td></td>
<td><strong>Introduction</strong></td>
</tr>
<tr>
<td></td>
<td><em>C. difficile</em> strains differ in adherence to host epithelia in culture</td>
</tr>
<tr>
<td></td>
<td>Adherence of CD strains over the growth cycle</td>
</tr>
<tr>
<td></td>
<td>Protein profiles of HV CD and the identification of SlpA</td>
</tr>
<tr>
<td></td>
<td>Surface protein extracts block adherence of CD to host cells</td>
</tr>
<tr>
<td></td>
<td>Surface protein interference is not strain specific</td>
</tr>
<tr>
<td></td>
<td>Antibodies to SlpA interfere with adherence</td>
</tr>
<tr>
<td></td>
<td>Survey of Multiple CD strains reveals wide range of adherence values</td>
</tr>
<tr>
<td></td>
<td>Sequencing of <em>slpA</em> from multiple <em>Clostridium difficile</em> strains.</td>
</tr>
<tr>
<td></td>
<td>SlpA binding to host cells of <em>C. difficile</em> SlpA subunits</td>
</tr>
<tr>
<td></td>
<td>Independently</td>
</tr>
<tr>
<td></td>
<td><strong>Summary</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Toxin Production and Gene Expression in Hypervirulent C. difficile</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Introduction</strong></td>
</tr>
<tr>
<td></td>
<td>HV CD strains do not secrete toxins during exponential growth</td>
</tr>
<tr>
<td></td>
<td>PaLoc Gene Expression Over the Growth Cycle</td>
</tr>
<tr>
<td></td>
<td>HV strains exhibit high <em>tcdA</em> and <em>tcdB</em> expression</td>
</tr>
<tr>
<td></td>
<td><strong>Summary</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Sporulation</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Introduction</strong></td>
</tr>
<tr>
<td></td>
<td>HV strains sporulated more efficiently than non-HV strains</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>CHAPTER FOUR: DISCUSSION</strong></td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Commensals and Pathogens</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Adaptation of CD to Animal Hosts</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Adherence Varies Between Strains and Is Multifactorial</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>Contribution of SlpA to Adherence and Colonization</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td>SlpA Subunit Sequence and Function</td>
<td>117</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>C. difficile</em> strains used in this study</td>
<td>45</td>
</tr>
<tr>
<td>2. <em>E. coli</em> strains used in this study</td>
<td>47</td>
</tr>
<tr>
<td>3. Plasmids used in this study</td>
<td>56</td>
</tr>
<tr>
<td>4. Primers used for RT-PCR</td>
<td>60</td>
</tr>
<tr>
<td>5. IC50 concentration of recombinant proteins</td>
<td>97</td>
</tr>
<tr>
<td>6. Positively selected core genes in <em>C. difficile</em></td>
<td>119</td>
</tr>
<tr>
<td>7. Gene deletions specific to the Hypervirulent (HY) clade</td>
<td>122</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pathogenicity Locus</td>
<td>15</td>
</tr>
<tr>
<td>2. Model of PaLoc regulation</td>
<td>19</td>
</tr>
<tr>
<td>3. Structure of Toxins</td>
<td>23</td>
</tr>
<tr>
<td>4. Stages of Sporulation</td>
<td>28</td>
</tr>
<tr>
<td>5. Phylogenetic relationship of <em>Clostridium difficile</em> strains</td>
<td>33</td>
</tr>
<tr>
<td>6. Arrangement of SlpA gene cluster</td>
<td>41</td>
</tr>
<tr>
<td>7. Crystal structure of SlpA</td>
<td>43</td>
</tr>
<tr>
<td>8. Schematic representation of <em>slpA</em> constructs cloned</td>
<td>55</td>
</tr>
<tr>
<td>9. Immunofluorescence microscopy of <em>C. difficile</em> adherence</td>
<td>65</td>
</tr>
<tr>
<td>10. Adherence of HV and non-HV <em>C. difficile</em> strains to C2BBE monolayers</td>
<td>68</td>
</tr>
<tr>
<td>11. Adherence of <em>C. difficile</em> during exponential phase and early stationary phase</td>
<td>70</td>
</tr>
<tr>
<td>12. <em>C. difficile</em> protein studies</td>
<td>72</td>
</tr>
<tr>
<td>13. <em>C. difficile</em> adherence interference assay using total SLP protein</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Section Title</td>
</tr>
<tr>
<td>---</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>14.</td>
<td>CD Supernatants and Cross-interference assays</td>
</tr>
<tr>
<td>15.</td>
<td>Adherence interference assay using anti-SlpA antisera</td>
</tr>
<tr>
<td>16.</td>
<td>Adherence of multiple <em>C. difficile</em> strains to C2BBE monolayers</td>
</tr>
<tr>
<td>17.</td>
<td>Low Molecular Weight Subunit Alignment</td>
</tr>
<tr>
<td>18.</td>
<td>High Molecular Weight Subunit Alignment</td>
</tr>
<tr>
<td>19.</td>
<td>Sequence Analysis of SlpA subunits of 6 CD strains</td>
</tr>
<tr>
<td>20.</td>
<td>SlpA crystal structure and recombinant proteins</td>
</tr>
<tr>
<td>21.</td>
<td>Adherence in the presence of recombinant LMW or HMW SlpA</td>
</tr>
<tr>
<td>22.</td>
<td>Growth curves and total toxin levels by ELISA</td>
</tr>
<tr>
<td>23.</td>
<td>Transcription of <em>tcdA</em> during exponential and stationary phase</td>
</tr>
<tr>
<td>24.</td>
<td>Transcription of <em>tcdB</em> during exponential and stationary phase</td>
</tr>
<tr>
<td>25.</td>
<td>Transcription of <em>tcdR</em> during exponential and stationary phase</td>
</tr>
<tr>
<td>26.</td>
<td>Transcription of <em>tcdC</em> during exponential and stationary phase</td>
</tr>
<tr>
<td>27.</td>
<td>Sporulation efficiency</td>
</tr>
<tr>
<td>28.</td>
<td>Schematic of the <em>C. difficile</em> PaLoc region differences</td>
</tr>
</tbody>
</table>
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amp</td>
<td>Ampicillin</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair(s)</td>
</tr>
<tr>
<td>BHI</td>
<td>3.7% brain heart infusion medium</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>C2BBE</td>
<td>Caco2-BBE epithelial cells</td>
</tr>
<tr>
<td>CD</td>
<td>Clostridium difficile</td>
</tr>
<tr>
<td>CDI</td>
<td>Clostridium difficile Infection</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s Modified Eagle Medium</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein isothiocyanate</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-thiolgalactopyranoid</td>
</tr>
<tr>
<td>L</td>
<td>Liter</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani media</td>
</tr>
</tbody>
</table>
kb  kilobase
kD  Kilodalton
M   Molar
ug  microgram
mg  milligram
ul  microliter
mM  millimolar
uM  micrometer
PBS Phosphate buffered saline
PCR Polymerase chain reaction
Q-PCR Quantitative PCR
RNA Ribonucleic acid
RT-PCR Real-Time PCR
SDS Sodium dodecyl sulfate
TAE Tris acetate ethylenediaminetetraacetic
TBE Tris borate EDTA
TBS 20mM Tris-Cl, 150mM NaCl, pH 7.5
TBST 20mM Tris-Cl, 150mM NaCl, 1% Tween 20, pH 7.5
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG</td>
<td>2.5mM Tris-Cl, 19.2% Glycine</td>
</tr>
<tr>
<td>TGS</td>
<td>2.5mM Tris-Cl, 19.2% Glycine, 0.01% SDS</td>
</tr>
<tr>
<td>Tm</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>v/v</td>
<td>Volume/ volume</td>
</tr>
<tr>
<td>w/v</td>
<td>Weight/ volume</td>
</tr>
</tbody>
</table>
CHAPTER ONE
LITERATURE REVIEW

Introduction

Infection with the bacterium *Clostridium difficile* is a pointed example of the costs and benefits of one of the most important tools of modern medicine: antimicrobial agents. The first antimicrobial class developed, the penicillins, has a broad range of activity against multiple kinds of bacteria, which led to their wide usage. But the cost of broad-range activity is collateral damage to the beneficial commensal microbiota of the human body, the value of which has only recently been appreciated. *C. difficile* has exploited this feature of modern medicine to become a significant pathogen of the human digestive system only in recent times.

*C. difficile* is a Gram positive, obligately anaerobic, rod-shaped bacterium of the family Clostridiaceae, of the class Clostridia in the phylum Firmicutes. While *C. difficile* (CD) is not part of the normal gut flora of adults, upon antibiotic treatment and concomitant reduction of commensal flora, *C. difficile* efficiently colonizes the gut [1, 2]. *C. difficile* can produce two toxins which damage the colon, causing disease ranging from mild diarrhea, to colitis, to fatal multi-organ failure [1, 2].

Even though the toxins were the earliest identified virulence factor, it has only been in recent years that advances have been made in understanding how the toxins are regulated, produced, and the mechanism by which they damage the host. Even
less is known about other non-toxin virulence factors relating to interactions with the host.

Treatment of CD infection involves further usage of antibiotics, which is effective, but also can predispose the gut to the same vulnerable state that precipitated the initial *C. difficile* infection. Re-infection and recurrences are unfortunately common [3]. Further compounding this problem is the ability of CD to form a spore, a dormant form of the bacterium that is resistant to antibiotics and cleaning agents. The ability to persist in the environment has made CD endemic in hospitals and long-term care facilities.

Since 2000, rates of CDI have increased, and epidemics of CDI characterized by greater incidence, severity and fatality, have been shown to be caused by “hypervirulent” (HV) variants of *C. difficile* [4-6]. HV epidemics are characterized by both increased morbidity and increased environmental predominance and spread.

Antibiotic usage, virulence factors, and spores have always made CD a challenge to control, which has only increased with the emergence of HV strains, especially since the virulence mechanisms of HV strains are not well understood. These challenges highlight the need to acquire basic information about this pathogen.

This work seeks to contribute to the greater knowledge about this pathogen, by elucidating the phenomena of bacterial adherence, toxin regulation, and spore production. Specifically, I have sought to characterize the interaction between hypervirulent *C. difficile* and the epithelia, and one bacterial surface protein that may mediate this interaction. I have also sought to clarify other factors that may lead to
hypervirulence, by determining the course of toxin production and sporulation over the growth cycle of *C. difficile*.

**Epidemiology of CDI: Causes, Incidence and Control**

**Disease Characteristics and History**

The first indication of *C. difficile* infection (CDI) is usually diarrhea, accompanied by fever and bloating if the infection progresses. As the toxins produced by CD accumulate in the colon, they continue to damage the colonic mucosa, and produce an inflammatory response. A sloughed-off layer of mucous, fibrin, dead epithelia and leukocytes, known as a pseudomembrane, forms over the interior surface of the colon [7]. More serious sequelae include toxic megacolon, in which the colon becomes paralyzed and distended with gas, and carries the possibility of rupture [8]. If dehydration is severe, kidney failure can occur. Death can result from any of these sequelae if left untreated, and additionally if the toxins become systemic and damage the heart [9].

The disease and the bacterium were identified separately many years before they were linked. Symptoms consistent with pseudomembranous colitis (PMC) were first described in a surgical patient in 1893 [10]. In the early years of penicillin usage, researchers noticed that treated animals frequently died of colitis [11, 12]. It was only later, with the advent of increased usage of antibiotics in the 1950’s and 60’s, that the syndrome of antibiotic-associated diarrhea was observed in humans, although it was initially misattributed to *Staphlococcus aureus* [13, 14].
Historically, before its pathogenic nature was discovered, the bacterium was identified in 1935 (known at that time as “Bacillis difficilis”) in one of its more benign appearances, in the normal microbiota of asymptomatic infants [15]. *C. difficile* is also frequently found in soil [16] and the gut microbiota of animals [17, 18].

Research to discover the organism responsible for “clindamycin colitis” began in the 1970’s. In 1974, Tedesco and colleagues associated the usage of the broad-spectrum antibiotic clindamycin with pseudomembraneous colitis, one of the more serious consequences of CDI, but the agent was till unknown. Toxin of some kind was detected in the stools of PMC patients [19], and the infection was thought to be bacterial because it responded to vancomycin [20]. Because the toxin was neutralized by *Clostridium sordellii* antitoxin antisera [21], but no *C. sordellii* could be isolated from patients, the search began for other *Clostridia* in the stools of patients and animals with antibiotic-associated disease. In 1978, Bartlett discovered that *Clostridium difficile* was responsible for the disease in humans and animals [22].

**Epidemiology and Antibiotics**

With the widespread usage of myriad antibiotics, *C. difficile* is now endemic in hospitals worldwide. Susceptibility to CDI begins as soon as the microbiota are compromised by nearly any antibiotic, and continues until the biota recover and return in quantity sufficient to exclude CD, which can be a significant time after the course of treatment ends. Some antibiotics evidence higher relative risk for CDI than others, such as ampicillin (or amoxycillin) and cephalosporins [23-25]. Third-
generation cephalosporins remain widely prescribed and show high relative risk for CDI [26, 27], particularly because CD is naturally resistant to this class.

If CD is resistant to the antibiotic, then it will be able to colonize during the course of treatment, in addition to after treatment. This longer window of susceptibility favors resistant strains, and as would be expected with selection pressures, the most prevalent epidemic CD strains are often resistant to the antibiotics in common use [28]. For example, as mentioned earlier, in the 1970’s rates were highest for clindamycin, and many epidemics throughout the 1980’s and 1990’s were caused by clindamycin resistant strains [29]. These strains contained the \textit{erm} gene, which confers resistance to macrolides, lincosamides and streptogramins [29, 30]. Recent epidemics have been characterized by recently evolved resistance to the fluoroquinolone class of antibiotics, which are currently in wide use [6, 28].

Fortunately, resistance to the two antibiotics primarily used to treat CDI, vancomycin and metronidazole, has not emerged, although increases in the amount of metronidazole required to treat CDI have been observed. [31]. However, in the search for additional treatment options, other drugs that have been used, such as rifaxamin, and have become resistant rapidly [32]. For a review of the clinical treatment of CDI, including recent challenges, the reader is referred to the review by Kelly and Lamont in the NEJM [33].

CD is one of the most commonly isolated cause of nosocomial diarrhea [34, 35], but estimates of how many patients become colonized with CD, and the proportion that become ill has varied widely from the time the organism was
described. Before the disease was recognized, and control measures implemented, rates of diarrhea were as high as 20% for those treated with clindamycin, with a 10% rate of PMC. [36]. Colonization rates in the general population are estimated to be around 2-3% [37, 38]. Current data estimates that 15-25% of nosocomial diarrhea cases are due to CD [8]. Several carefully controlled studies determined that about 20% of patients acquired CD during their hospital stay, and that most of those remain asymptomatic. [39-41].

One general trend is that all rates are unfortunately increasing. Epidemiological data from the UK indicated that the number of deaths reported with CD as a cause or contributor has grown approximately 8-fold from 1999 to 2007, and that CD was responsible for more deaths than Methicillin-Resistant *Staphylococcus aureus* (MRSA) [42]. In the USA, in the late 1990’s, one measure of yearly CDI cases numbered 82,000, but rates in the mid 2000’s had doubled[43]. Mortality rates in the US between 1999 and 2004 have increased 35% per year [44]. These increases are primarily attributed to the emergence of more virulent strains of CD, which will be discussed in detail later in this work.

**Risk Factors for CDI**

Factors that increase a person’s risk of contracting CDI include both intrinsic characteristics and medical manipulations. Antibiotic use evidences the highest risk, with a relative risk of 5.9 for diarrhea and 4.2 for asymptomatic carriage [25]. In addition to exposure to precipitating antibiotics, hospitalization places a vulnerable patient in an environment frequently contaminated with CD spores. Factors related to
hospitalization, such as length of stay, severity of co-morbid conditions, and proximity to a contaminated environment also increase relative risk [25, 45]. Advanced age is correlated with increasing propensity for CDI [46-48], mainly due to these patients having more severe co-morbidities [5, 49] and possible immune senescence, but not necessarily greater exposure to antimicrobials [50].

Exposure to antacid medication has been a risk factor of recent debate. The data have been contradictory, with some studies indicating increased risk CDI [51, 52], while other studies finding no association [28, 53, 54], more recent data have accumulated indicating that antacid medication does increase risk, and that the risk is proportionate to the degree of acid suppression [55]. Because CD spores, the assumed infectious particles, are naturally resistant to acid, the mechanism of this effect remains unknown, although it is proposed that vegetative cells may be the infectious agent in these cases [56].

**Immunity and Vaccines**

The patient’s immune response is one characteristic that is predictive of a positive or negative outcome. People who have an existing response in the form of IgG to Toxin A were significantly less likely to develop symptomatic disease [57]. Because of this association of Toxin A antibodies and disease in human populations, most vaccine efforts have focused on the toxins.

Early work with Toxin A in animals showed that toxoid vaccines elicited immunity [58, 59]. In humans, A toxoid vaccine was shown to elicit IgG against toxin A levels that exceeded those shown protective in the survey of symptomatic and
asymptomatic patients [60], and is also effective in recurrent cases [61]. A DNA vaccine, in which a gene for the receptor binding region of Toxin A was injected into mice, also produced high titer antibodies [62].

Adhesins and other surface proteins are also attractive vaccine candidates, such as the surface proteins in the MeNZB meningococcal vaccine [63]. CDI patients develop antibodies to surface proteins following infection [64]. As with the toxins, it has been noted that patients with lower levels of antibodies to CD surface proteins were more likely to have relapses of CDI [65]. In animal models, while passive immunization with antibodies to the surface layer protein SlpA prolonged life, a vaccine against SlpA was not been able to elicit protective immunity [66]. However, vaccines with total cell wall extracts have shown some effect [67].

Economic models of a *C. difficile* vaccine have shown that it would be cost effective to treat both patients at risk for CDI and those who already have contracted the disease to prevent recurrence [68].

**Costs and Prevention**

As mentioned earlier, rates are increasing across the board, as evidenced by increasing rates of infection previously low-risk populations, such as children and pregnant and peripartum women [69-71].

Clinically, the cost of *C. difficile* infection (CDI) is great. Nosocomial *C. difficile* infections prolong hospital stays and increase costs. One recent meta analysis estimated that each case of CDI adds approximately US$3000 to $4000 to patient’s hospital bill in the USA ($8500 outside the USA), and recurrent CDI can add up to
$18,000 [72]. In 2002, Kyne et al used costs per patient described in prospective study to project the total cost of CDI to the US healthcare system, which was $1.1 billion [73]. With the increase in rates due to recent epidemics, costs are now estimated to be $3.2 billion per year for the United States [74].

To acquire *C. difficile*, a susceptible patient must ingest *C. difficile* spores from the environment. In the hospital, spores are found on surfaces, medical devices and can be transferred to patients by transient hand carriage by healthcare personnel [2]. CD spores are not killed by alcohol-based hand gels [75], and the recent rise in CDI rates has been partially attributed to the increased usage of these gels in lieu of traditional soap and water hand washing [76]. Unfortunately, exposure to some detergents only increases sporulation rates of vegetative CD [77]. Sodium hypochlorite is one agent consistently effective against spores [77, 78]. Rates of transmission have been reduced by encouraging proper hand washing and glove use, employing disposable medical equipment, and using bleach to clean surfaces [79-81].

An effective strategy for CDI prevention has been to limit the usage of precipitating antibiotics. For example, usage of the antibiotic clindamycin was restricted due to its association with CDI, and outbreaks of resistant strains have been controlled by limiting its usage [82]. Cephalosporin usage has also followed a similar pattern [83]. During the outbreak of hypervirulent CD in Québec in the mid-2000’s, limiting the use of cephalosporins, ciprofloxacin, clindamycin and macrolides reduced CDI cases by 60% [84]. The most successful approach for reducing epidemics has been to implement infection control at multiple levels, including
increased cleaning measures and hygiene, antibiotic stewardship, and increased patient monitoring. This “bundle” approach has been effective at reducing CDI rates by as much as 78% [85].

**Pathogenesis**

**Toxins**

Even when CD was first described in healthy infants, it was noted that culture filtrate of the organism could kill rabbits when injected [15]. We now know that disease-causing CD produce at least two toxins. When CDI was first being described, the two toxins were described independently, and initially confused [86]. At the time, there was controversy about the contribution of each toxin in the disease process, which interestingly, has been revisited recently.

Toxin in the stools of PMC patients was cytotoxic to mammalian cells in culture [19], and supernatants of CD were able to produce PMC lesions in animals [87]. In 1981, two independent research groups used ion exchange chromatography to separate secreted proteins, and showed that there was another large toxin present, which was called Toxin A [88, 89]. Toxin A produced less rounding of fibroblasts in culture, but was able to produce fluid accumulation in ligated ileal loops of rabbits, while Toxin B could not. Also, the work of Lyerly and Wilkins in 1985 showed that when given intragastrically, Toxin A produced a similar profile as total supernatant, which Toxin B did not, unless mixed with small amounts of Toxin A, or given to animals with bruised ceca [90]. It was concluded that Toxin A was the chief
virulence factor, and Toxin B, while a potent cytotoxin in culture, required Toxin A to be present, to begin the cascade of damage to the colonic mucosa and produce disease at the organismal level [86].

Before the development of genetic tools to manipulate CD, natural mutations and variations in the toxin genes were sought out to shed some light on their function and role in pathogenesis, especially in the debate as to which toxin was most important. Clinical reports surfaced of variant CD strains that were causing disease, but produced no detectable Toxin A. This challenged current thinking, and was not only academically interesting, but it also had ramifications for the clinical practice. At that time, clinical immunoassays only tested for the presence of Toxin A. The first Toxin A-negative Toxin B positive (A-/B+) strain described was strain 8864, described by Lyerly et al in 1992 [91], which demonstrated that in produce no Toxin A in vitro by immunoassay, and that Toxin B from 8864 was more lethal and weakly enterotoxic to animals. Strain 8864 was later shown to have a deletion encompassing the 3’ends of tcdA and tcdC and that the Toxin A protein lacks the receptor binding repeats and translocation domain [92] While strain 8864 had an obscure clinical origin, other disease-causing A-B+ strains were being reported more frequently in the early 1990’s. Strain 1470 and related isolates [93, 94] also have deletions in tcdA and were negative in clinical tests for Toxin A, were initially thought to be non-pathogenic in mice [93], but further work in the more sensitive hamster model [95] and fatal cases in humans [96] and human epidemics soon dismissed this idea [97-100]. However, Toxin B from these strains shared homology
with a toxin from a related *Clostridium* species, TecSL from *C. sordelii* [101], in its enzymatic domain, which allows it a larger range of host molecule specificity and increased toxicity. This variation in Toxin B’s enzymatic domain left the debate concerning Toxin A vs. B open.

Recent advances in the genetic manipulation of CD have finally allowed a molecular scientific answer to this question. Using an unstable plasmid as a recombination vector [102], Julian Rood’s research group were able to construct isogenic mutants in the genes for Toxin A and Toxin B in strain 630 [103]. When these mutant strains were introduced into the hamster model, it was shown that Toxin B was the required virulence factor. Hamsters infected with the Toxin A mutants CD showed no difference in mortality as compared to wild-type. In contrast, the group of hamsters infected with the Toxin B mutant CD showed significantly less mortality, and analysis of the CD obtained from the fatal cases indicated that they were revertant strains in which recombination vector was excised, also supporting the idea that Toxin B production is favored for pathogenesis and survival in the gut [103].

However, a more recent study has found conflicting results. Using the ClosTron system of a re-targetable group II intron [104], null mutations were made in *tcdA* and *tcdB* [105]. In contrast to the previous study, both toxins contributed to virulence in the hamster model. However, hamsters infected with CD lacking Toxin B survived for several days longer than those infected with CD lacking Toxin A [105]. While both may participate in the disease process, it appears that Toxin B makes the larger contribution to virulence.
The Structure and Regulation of the Pathogenicity Locus

Early research also recognized that there were strains of CD that did not cause disease or produce toxins in culture. The mechanism behind these non-toxigenic strains was not known, but it was suspected that they lacked the genes for the toxins, which was confirmed when the toxin genes were sequenced, and later when the whole genetic island on which they reside was described. The individual genes for Toxin A and Toxin B were sequenced in 1990 [106, 107], and the neighboring genes were later sequenced and designated the Pathogenicity Locus or PaLoc by Eichel Streiber’s group in 1996 [108]. Non-toxigenic strains of *C. difficile* that lack this genetic locus entirely, having only a 115 bp “scar” in the same region [109].

The genes for toxins A and B, *tcdA* and *tcdB* (formerly *toxA* and *toxB*), are located in the PaLoc with three other genes, *tcdR*, *tcdE* and *tcdC* (Figure 1A) [108]. The toxin proteins are among the largest bacterial toxins described, and are encoded by single genes. In strain VPI 10463, *tcdA* is an 8133 bp gene which codes for a 2710 amino acid, 308 kD protein, while *tcdB* is 7098 nucleotides, which codes for a TcdB molecule of 2366 amino acids and 279kD [106, 107].

Toxin B and Toxin A are 44.8% identical and 63.1% similar at the amino acid level, and may have originated from a gene duplication event [110]. Both share similar overall topology, with an N-terminal enzymatic domain, a central putative translocation domain, and a C-terminal receptor-binding domain consisting of multiple repeats (Figure 1B). Briefly, the enzymatic domain modifies host signaling proteins, the translocation domain may be involved in the transit of the toxin through
host membranes, and the receptor binding domain is responsible for the binding of toxin to sugar moieties on the host cell surface. TcdA and TcdB are most similar (74% homologous) in the N-terminal enzymatic domain [111], which confers substrate specificity. The two CD toxins belong to a superfamily of similar molecules known as the Large Clostridial Toxins, which also includes the lethal and hemorrhagic toxins of *C. sordellii*, and *C. novyii* alpha toxin [112], all with similar 3-domain structures and mechanisms of action.
Figure 1. A. Schematic of the Pathogenicity Locus. Arrows indicate the direction of transcription. B. Structure of Toxins A and B. Adapted from Jank et al, 2007 Glycobiology 14(4) 15R-22R.
In addition to the toxins, the locus encodes for three other genes, \textit{tcdR}, \textit{tcdE} and \textit{tcdC} (Figure 1A). \textit{tcdR} (formerly \textit{tcdD} or \textit{txeR}) codes for an alternate sigma factor required for the expression of \textit{tcdA} and \textit{tcdB} [113]. TcdR is a 22kD protein that features a helix-turn-helix DNA binding motif [114], and has defined a new group of the sigma 70 family of sigma factors, Group 5 or Extra-Cytoplasmic Function (ECF) [115, 116]. As a sigma factor, it binds to RNA polymerase, and allows transcription at specific promoters. Mani and Dupuy first demonstrated that TcdR was required for \textit{tcdA} and \textit{tcdB} expression \textit{in vivo} and could activate toxin promoter-specific transcription \textit{in vitro} [113]. TcdR is also homologous to toxin-inducing regulators for other Clostridia, such as BotR of \textit{C. botulinum} and UviA of \textit{C. perfringens}, and can substitute for UviA in \textit{in vivo} transcriptional studies [115, 117-119].

\textit{tcdE} encodes a protein of unknown function with homology to phage holins. Because the toxins do not have classical secretion signals, TcdE has been implicated in toxin release during stationary phase [120]. When expressed in \textit{E. coli}, \textit{tcdE} produces membrane lysis as visualized by microscopy [120]. Along these same lines, infection of CD with lysogenic phages has been shown to increase release of toxins, although not transcription of toxin genes, presumably by providing additional holins [121].

\textit{tcdC} is transcribed on the opposite strand from the other PaLoc genes, and is postulated to act as negative regulator [122]. This unique 26 kDa protein has no homology to any described regulatory protein [122], and it is membrane-associated [123]. The function of TcdC in the regulation of toxin production is the center of
much recent debate. The earliest data regarding the function of TcdC was a transcriptional study by Hundsberger et al [119]. Using semi-quantitative RT-PCR, they noted that transcription of tcdA, tcdB, tcdR and tcdE was not detectable during exponential growth, but increased as bacteria entered stationary phase. However, tcdC transcripts evidenced the opposite pattern, being transcribed in exponential phase, and declining in stationary phase.

This pattern of transcription led to the hypothesis that TcdC was acting as a negative regulator in exponential phase, the mechanism was unknown. The first functional studies of TcdC demonstrated that, in a heterologous system, TcdC strongly decreased the TcdR-dependent transcription of tcdA promoter linked to a reporter gene. Gel shift and surface plasmon resonance assays indicated that TcdC appears to act as a novel kind of anti-sigma factor by inhibiting TcdR from complexing with RNA polymerase core enzyme, or preventing the holoenzyme from forming an open complex [122].

These data, and data concerning the function of TcdR as a positive regulator have led to the current model of PaLoc regulation (Figure 2) [124]. In general, Toxins A and B are produced in stationary phase in response to starvation. When nutrients are abundant, expression of the toxin genes is repressed by the global stationary-phase regulator CodY, which binds to a region upstream of tcdR and prevents its transcription. Without TcdR to act as a required sigma factor, transcription of tcdB and tcdA is not possible (Figure 2A) [119, 122, 125]. In a C. difficile CodY mutant, toxin production was de-repressed and toxin production occurred during exponential
phase [125]. During stationary phase, lower levels of particular nutrients decrease the binding affinity of CodY for its target sequence, allowing for read-through transcription of \(tcdR\). \(tcdR\) is transcribed, and then goes on in a positive feedback loop to promote its own transcription as well as that of \(tcdB\), \(tcdE\) and \(tcdA\) (Figure 2B).

Transcription is primarily monocistronic, initiated from individual promoters preceding \(tcdB\) and \(tcdA\), but polycistronic read-through transcripts do originate from upstream promoters as well [119, 126, 127].
Figure 2 Model for the regulation of the Pathogenicity Locus of *Clostridium difficile*. A. Exponential Phase. B. Stationary Phase. Adapted from O’Connor et al, 2008, Gastroenterology 136(6), 1913-24.
Nutritional effects on toxin production are myriad and complex. It had long been reported that toxin production was catabolite repressed [126, 128]. Catabolite control appears to be exerted through CodY, although not directly. The presence of branched-chain amino acids (isoleucine and valine) and GTP increase the binding affinity of CodY to the DNA [125, 129]. Glucose exerts its effect indirectly through the TCA cycle by influencing the fermentation of amino acids [128]. Biotin is essential for growth, but may repress toxin production [128, 130], as may certain amino acids such as lysine and cysteine [131].

**Mechanism of Toxin Action and Host Responses**

In defining the mechanism of the CD toxins, some information is relevant to only one toxin or another, but chiefly evidence has accumulated that demonstrates that they function similarly, with particular limitations noted. Following secretion, the first step in intoxication is binding to the host cell surface. In this case, more is known about Toxin A than Toxin B. The C-terminal CROP region has homology to a glucan binding protein from *Streptococcus mutans* [110, 132]. Toxin A can bind to several types of carbohydrate moieties, which vary by cell, tissue and species. For example, Toxin A binds to and agglutinates rabbit erythrocytes, but the carbohydrates on the surface are not found in other mammalian erythrocytes [133, 134]. In 1991, Tucker showed that toxins bound to human carbohydrate antigens I, X and Y- all of which contain the core Galbeta1-4GlcNAc, although some doubt has been raised about this result [135]. Recent crystal structures have indicated that Toxin A binds to two
carbohydrate molecules (in this case, alpha-Gal-(1,3)-beta-Gal-(1,4)-beta-GlcNAc O(CH(2))(8)CO(2)CH(3)) in each of seven conserved binding pockets in the CROP region (Figure 3A) [136]. On the surface of epithelia, this moiety could be attached to a glycoproteins or glycolipids. Recent data suggest that a glycoprotein on human cells, gp96, binds to Toxin A [137]. The varieties of carbohydrate modifications on human gut epithelia are one proposed explanation for why animal and human newborns do not become ill when colonized with CD, as these structures change over development [138].

The receptor for Toxin B is unknown, but in vitro tests indicate that it is distributed differently than the receptor for Toxin A. Toxin A can bind to both the apical and basal side of polarized T84 cells in culture, and decrease trans-epithelial resistance, but Toxin B could only do so from the basal side, or with small amounts of A from the apical side, once again pointing to some synergy between the two toxins [139].

Once bound, the toxins are endocytosed via a clathrin-mediated endocytosis in a dynamin-dependent manner [140, 141]. Early work by [142] indicated that acidification was required for Toxin A to become active in vivo. The cytotoxicity of Toxin B was also shown be abrogated in the presence of bafilomycin, which blocks the acidification of endosomes [143, 144]. Fractionation studies indicated that the enzymatic fragment is delivered into the cytoplasm, while the translocation domains and CROP regions remain in the endosome [145]. The acidification allows the membrane translocation region to change conformation and insert into the membrane
of the early endosome [144]. If the hydrophobic translocation domain is deleted, the cytotoxicity of Toxin B is reduced by more than 90% [146]. The pore-forming process for both Toxins requires cholesterol-containing membranes [147].

Recent work has defined how the enzymatic domain enters the cytosol. Early work suggested that the toxins required proteolytic cleavage to become active, as protease inhibitors could blunt the action of Toxin A on cells [148]. Recently, it has been discovered that a host co-factor, inositol hexakisphosphate, induces autocatalytic cleavage and activation of *C. difficile* toxin B in the cytoplasm [149]. The autocatalytic cleavage was initially thought to be a serine protease located C-terminal to the hydrophobic domain [149], but later work determined that the autocatalytic domain was a cysteine protease domain located closer to the enzymatic domain (Figure 3B) [150].
Once the enzymatic domain enters the cytosol, it exerts its effect on the small GTP binding proteins of the Rho and Ras superfamilies, particularly Rho, Rac, Ras, and Cdc42 [151]. Small GTPases are signaling proteins that cycle between an active, GTP-bound state and an inactive, GDP-bound form. Guanidine exchange factors, or GEF’s receive activation signals and exchange GDP for GTP. In this form, they can interact with a variety of effectors. The active state ceases with GTP hydrolysis to GDP, aided by GTPase activating proteins (GAP’s) [152].

Toxins A and B perturb this process by glucosylating the small GTPases, which permanently fixes them in the inactive form by the transfer of the glucose moiety of UDP-glucose to a threonine residue in the switch region of the GTPases [153, 154]. In the enzymatic domain of the toxins, the DXD motif is essential; mutating it leads to a loss of enzyme activity [155]. This motif is thought to mediate nucleotide-sugar binding in the presence of a Mn\(^{2+}\) ion [155], [156]. This glucosylation blocks the cycling of the GTPases between the membrane and the cytosol, parking them inactive at the membrane [157]. It also blocks the action of GEFs and GAPs, so no new GTP can be exchanged, and most importantly, it blocks the interaction with effectors [158, 159], the chief effect of this irreversible glucosylation is significant effects on the host cell cytoskeleton. Rho family members are in particular involved in the regulation of the actin cytoskeleton, the disruption of which is responsible for the cell rounding and neurite formation seen in intoxicated cells in culture [151]. Another downstream effect of inactivated GTPases is the induction of apoptosis via Caspase 3, 6, 9 and Bid [160].
The toxins also have effects independent of GTPase activity, for example, they cause mitochondrial damage [161], which also induces apoptosis. In macrophages, Toxins A and B activate the inflammasome, induce the release of IL-1B, and start the cascade of an inflammatory immune response [162].

The consequences of this intoxication on gut epithelia are disruption of tight junctions [163, 164]. Toxins also have systemic effects, which may be responsible for death due to CDI. In a study using zebrafish embryos as a model system, Hamm et al demonstrated that Toxin B localizes to the heart, decreases contractility and induces apoptosis in heart tissue [9].

**Binary Toxin and other virulence factors**

In 1997, a new *C. difficile* toxin was discovered [165]. Similar to binary toxins found in other Clostridia, such as iota toxin of *C. perfringens*, it has a structure consisting of two separate components, which combine to make the fully functional toxin [165]. Each component is encoded by one of two two neighboring genes. Most described Clostridial binary toxins are believed to function by the same mechanism [166]. In CD, CdtA is the enzymatic component, and CdtB is the binding and translocation component. The two fragments are secreted separately, and combine in the external millieu for the fully functional toxin. The binary toxin is taken up by receptor mediated endocytosis, and the acid environment induces pore formation and the translocation of the enzymatic component [166]. Once inside, they ADP-ribosylate actin, which leads to the disintegration of the cytoskeleton. Expression of
cdtAB genes is regulated by a third member of the CDT locus, the orphan response regulator CdtR [167].

Surveys indicate that a minor percentage of C. difficile strains carry the binary toxin Locus, with estimates ranging from 5.8 to 15.5 % [168-170] It is distinct from the PaLoc, and has been found in strains which do not carry the PaLoc [169]. In animal models its function remains obscure. Strains that produce binary toxin but not the LCT’s are able to colonize animals, but produce no diarrhea or disease. However, purified toxin can cause fluid accumulation in the ligated ileal loop assay [171]. Its contribution to CDI, whether alone or in combination with the LCT’s in humans is still under investigation.

*C. difficile* produces other non-toxin virulence factors such as several hydrolytic enzymes, fimbriae and a capsule [172]. The hydrolytic enzymes includes hyaluronidase, collagenase, and a chondroitin-4-sulfatase, which are predicted to act to release nutrients in the gut [173]. While unrelated to the production of toxin, the presence of these enzymes was more frequent in highly toxigenic strains [174].

**Sporulation**

CD spores are metabolically inactive forms of the bacteria that terminally differentiate from vegetative cells. Spores are resistant to heat, desiccation, and chemical onslaughts. Spores can persist in aerobic environments, such as on surfaces in hospitals. Because of these features, and because they are resistant to destruction by stomach acid, spores are the infectious particle leading to CDI. Thus, their development and function are highly relevant to CD pathogenesis.
The production of spores begins when the vegetative cell undergoes a special kind of cell division in which the cell is divided into two unequal compartments. The larger compartment is called the mother cell and the smaller is the forespore. The forespore becomes enveloped by and develops inside the mother cell. The mother cell lyses to release the spore. Sporulation is a complex process of seven stages of development (Figure 4). First, the nuclear material is replicated and forms an axial filament. Instead of a central septum forming as in normal cell division, an asymmetrical invagination of the membrane pinches one copy of the chromosome with the developing forespore septum, in stage II. The remaining portion of the chromosome is actively transported into the forespore that is then sealed by the completion of the septum. In stage III, the membrane of the mother cell grows and engulfs the forespore forming a vesicle with two membranes. During stage IV, a cortex of peptidoglycan forms between the two cells, and calcium and diplocolinic acid accumulates in the forespore. Next, a thick proteinaceous coat comprised of multiple proteins is assembled on the outer forespore membrane in stage V. The spore matures in stage VI, with the completion of coat synthesis, accompanied by an increase in refractility upon the dehydration of the forespore. In the final step, the mother cell lyses by the action of lytic enzymes and the mature spore is released. [175].
Figure 4. Stages of Sporulation as described in *Bacillus subtilis*. Figure courtesy of Timothy Paustian, University of Wisconsin-Madison.
This complex developmental program is the result of stepwise gene expression, controlled by a cascade of sigma factors. In general, sporulation is induced by starvation, and additionally for \textit{C. difficile} by the presence of oxygen. In the better-described system of \textit{Bacillus subtilis}, nutrient deprivation stimulates a phospho-relay resulting in the phosphorylation of the master regulator, Spo0A. \textit{C. difficile} does not have a multicomponent phosphorelay [176], but one orphan sensor kinase has recently been identified, and was able to phosphorylate CD Spo0A [177]. A mutant in this gene was compromised for sporulation [178].

Once active, Spo0A activates downstream regulators involved in sporulation, and represses other regulators involved in vegetative functions. The chief controllers of sporulation are two cascades of sigma factors, one that is active in the spore, consisting of Sigma F and Sigma G, and one that is active in the mother cell, consisting of Sigma E and Sigma K [175]. These sigma factors are synthesized in an inactive form and activated by specific sequential signal. Spo0A, along with housekeeping and stationary phase regulators, initiates the formation of the septum. Once the compartments are divided, the inhibitors of Sigma F are deactivated in the forespore and Sigma F is activated. Sigma E in the mother cell is activated by proteolytic cleavage. Sigma E directs the production of coat proteins, and by communication with the forespore (by an unknown mechanism) activates Sigma G. Sigma G directs the production of a signaling and proteolytic cascade that results in the activation of sigma K, the final sigma factor. Sigma K directs the assembly of the outer layer of the spore [175].
Interestingly, in both *B. subtilis* and CD, the Sigma K gene is interrupted by a pro-phage element, although not the same one. While disposable in *B. subtilis*, the excision of this element during sporulation is required for proper timing of the activation of Sigma K in *C. difficile* [179]. CD lacks a homologue of the protease that activates Sigma K, and CD Sigma K lacks the signal sequence for cleavage, so the excision of this element is the main regulator of the activity of Sigma K. Premature activity of Sigma K results in a 100-fold decrease in sporulation efficiency in a CD strain without the prophage element [179]. Variations in sporulation efficiency in different strains or populations could conceivably affect the spread of CD in the nosocomial environment.

Germination is an equally complex process, and is key to the initiation of disease. Again, not much is known about CD germination. In *B. subtilis*, germination begins when germinants such as glucose, peptidoglycan, or calcium are sensed by receptors in the inner membrane of the spore. After receptor binding, an irreversible program begins in which divalent cations and calcium-dipicolinic acid are released, leading to an influx of water, which results in an expansion of the core [180]. Finally, the hydrolysis of the cortex peptidoglycan is followed by active metabolism and outgrowth [180].

While *B. subtilis* spores germinate in response to nutrients such as glucose, CD must sense several features characteristic of its appropriate niche in the human gut. No homologues exist in the CD genome to the *B. subtilis* germinant receptors GerA, B and K [181, 182]. However, it has long been noted that adding a primary
bile acid, taurocholate, to culture media improves spore recovery [183-185]. Taurocholate and glycine are abundant in bile, and have been demonstrated to activate germination [181]. The receptors for taurocholate and glycine are unknown, but kinetic data suggest a sequential progression of binding of taurocholate followed by glycine to specific receptors [186].

The interaction of taurocholate and the microbiota may underlie how CD is normally excluded from the human gut. While taurocholate is a CD germinant, it is hydrolyzed in the gut by other bacteria to secondary bile salts such as cholate and chenodeoxycholate. These derivatives, chenodeoxycholate and deoxycholate, inhibit germination and vegetative outgrowth [187]. In vivo, antibiotic-treated mice have a higher proportion of primary to secondary bile salts in their ceca, and cecal contents better promote the growth of CD ex vivo [188]. This may explain the mechanism by which antibiotic treatment predisposes patients to CDI.

**The Emergence of Hypervirulent Strains**

As mentioned in the Epidemiology section, in the past ten years epidemics of CDI characterized by greater incidence, severity and fatality have been shown to be caused by highly virulent variants of toxigenic *C. difficile* [4-6]. Past epidemics were usually confined to a particular institution, although occasionally there have been epidemics of a particular strain over a wider geographic area, such as the J9 strain outbreak in the eastern US in the 1990’s [189].

Three major outbreaks in the early 2000’s were indicators of an unprecedented larger problem. In 2000, a cluster of severe cases of CDI occurred at
the University of Pittsburgh Medical Center. Incidence of CDI roughly doubled in one year, with a sharp increase in mortality and morbidity, including nine patients requiring surgery to remove their permanently damaged colons [190]. The next major outbreak was in Montreal, Canada during 2003, where CDI incidence quadrupled as compared to previous years, and CD caused or contributed to 244 deaths [5]. Increased morbidity was also seen, with 33 patients requiring colectomies [5]. The highest rate of mortality (approximately 11% of CDI cases) occurred during two outbreaks between 2003 and 2005 at the Stoke-Mandeville hospital the UK in which 38 people died [191].

As outbreaks accumulated, genetic typing studies indicated that these outbreaks were caused by a single genetic cluster of *C. difficile* [6]. Hypervirulent strains cluster into a distinct genetic group by several different typing methods: Multi-Locus Sequence Typing [192], toxinotyping [193, 194] where they are designated type III, PFGE and PCR ribotyping [194] where they are designated type NAP1/027, and Restriction Endonuclease Analysis [6] where they are called group BI. After this identification enabled tracking, BI strains have been found responsible for CDI epidemics worldwide [6, 195-199]. A large study using comparative phylogenomics has shown that there are four distinct clades of *C. difficile*, and that the HV strains comprise their own clade, HY, as shown in Figure 5 [200]. Toxigenic strains comprise the HA1 clade. Non-toxigenic strains and Toxin-variant strains comprise a third clade, A-B+, and animal derived isolates form the HA2 clade.
Figure 5 Phylogenetic relationships of *Clostridium difficile* strains as determined by microarray analyses. Strains distributed into four major clades (HY, A-B+, HA1, and HA2). Strain names are shown at the end of the branches and are colored according to the animal source of isolation. Black, human; blue, mouse; green, bovine; red, swine; light blue, equine. Branches with ** have a P value of 1.0 and represent 100% of all phylogenies showing a given topology. * indicates a P value of ≥0.98. Figure from Stabler et al, J Bacteriol. 2006.
To understand the virulence mechanisms of HV strains, initial work focused on known virulence factors, namely antibiotic resistance and toxins. As mentioned earlier, strains of CD resistant to the antibiotics in common usage have a competitive advantage. Fluoroquinolone usage has been a significant risk factor in HV epidemics [28]. The HV strains are resistant to the fluoroquinolone class of antibiotics, and particularly to gatifloxacin and moxifloxacin, which have the strongest anti-anaerobe activity [6]. Resistance arises from a single amino acid substitution in the target of fluoroquinolones, DNA gyrase [201]. This resistance appears to be recently evolved, as historical CD clinical isolates of the BI genetic group are not fluoroquinolone resistant [6, 75].

Because of the severity of disease seen in the early epidemics, toxin production was a target of initial investigations. HV CD is toxinotype III and produces both toxins, and also Binary Toxin [5]. As mentioned earlier, the contribution to Binary Toxin to human disease is unknown. Sequencing the \( tcdB \) from HV strains has shown that the enzymatic domain is only about 90% identical as compared to the sequenced strain 630 [200, 202]. This altered enzymatic domain may have different activity against the GTPases. The C-terminal binding domain is also different, which may affect the cell type tropism of HVCD. In vitro, purified TcdB from an HV strain displayed increased cytotoxicity across a broader range of cell types than non-HV strains, however the mechanism for this difference is not yet known [202].

The majority of the research into HV toxins has focused on the relative
amounts of the two toxins produced. These values, and their potential impact on
disease severity are the subject of debate [203, 204]. In 2005, a study by Warny et al
found that the median amounts of toxins A and B in a group of HV strains were 16
and 23 fold higher, respectively, than the median toxin amounts for a group of non-
HV strains, and concluded that HV strains also expressed toxins during exponential
growth, although no quantitative toxin measurements were presented for the 0-24
hour time period [194]. On the other hand, in 2006, Akerlund et al found that one HV
strain produced 3-13 fold more toxin that a group of non-HV strains. In 2007, using a
human gut model, Freeman et al [205] reported that while mean toxin titers were not
significantly higher for an HV strain than for a non-HV strain, the HV strain had an
extended duration of toxin production. The fold differences espoused by Warny et al
have been most frequently cited, in scientific journals and popular media [206, 207]
,[208, 209].

While the toxin genes in HV strains were intact, initial analysis of the PaLoc
found that tcdC from HV strains contains an 18bp deletion as compared to published
sequences. This variation in this negative regulator was initially suspected to be
responsible for the HV toxin phenotype [5]. Further sequencing of tcdC identified a
single base pair deletion at position 117 that produces a frameshift mutation,
rendering the protein truncated and likely non-functional [122, 194, 208].
Biochemical studies demonstrated that the 18bp variation had no effect on TcdC
function [122]. Clinical studies also showed no correlation with disease severity with
the 18bp deletion, which has also been found in non-HV strains [210].
While the *tdcC* truncation has been well described in HV strains, its direct contribution hypervirulence is unknown. Whether the lack of this negative regulator is sufficient to account for toxin production during exponential growth, as proposed by Warny et al [194], has not been specifically tested.

Although an increase in toxin production may account in part for increased disease severity, it does not fully explain why hypervirulent strains also predominate in the hospital environment during and after an epidemic [197], or the rapid geographic spread of hypervirulent *C. difficile*. As mentioned earlier, the most likely form of *C. difficile* in the healthcare environment is the highly resistant bacterial spore, which is spread to susceptible patients either by environmental contact or carriage by healthcare personnel [211, 212]. A study by Akerlund et al showed that one strain of HV CD had increased sporulation [213]. It is not improbable to suggest that the rate of sporulation may contribute to the spread and persistence of HV CD.

The research described in this work more clearly defines the factors involved in hypervirulence, by determining the course of toxin production and sporulation in HV strains over the growth cycle.

**Colonization and Adherence**

**Colonization and animal models**

*Clostridium difficile* is unusual in that the animal model for the disease was delineated before the etiologic agent was discovered. In the early work of Bartlett in identifying *C. difficile*, it was noted that antibiotic-treated hamsters and humans shared the same symptoms and the presence of a Clostridial toxin in their gut contents
In the hamster, the cecum is the main site of CD proliferation, which can become dilated and hemorrhagic, and also exhibits a loss of epithelium consistent with pseudomembranes.

Colonization is key for CDI to occur. Unlike C. botulinum, which can cause disease from the ingestion of botulinum toxin in food, CD has to colonize the gut to produce appreciable toxins, and most frequently begins with the ingestion of spores. Early work also noted that animals treated with antibiotics frequently became colonized with CD strains from the environment. If the contaminating strain was non-toxigenic, further inoculations of toxigenic CD did not result in colonization of that strain or disease.

This observation of a competition effect has been further characterized and refined by the laboratory of Dale Gerding, who has patented the use of certain non-toxigenic strains as a preventative treatment for human patients. Colonization is also key in the protection effect. In the hamster, when detectable colonization of the non-toxigenic strain is achieved, protection from challenge by a toxigenic strain occurs in 80-100% of animals. This is true even for challenge with hypervirulent strains.

Competition between mucosal bacterial species is common. As mentioned in the sporulation section, the mechanism by which the gut biota normally excludes CD is becoming better defined. Intra-species competition, such as seen in the protection effect, also occurs. For example, resident Staphlococcus aureus in the respiratory tract is able to exclude other S. aureus from colonizing, even when there is no
difference in fitness between strains [223]. Hypothesized mechanisms for intra-species exclusion include monopolization of binding sites or essential nutrients [221, 223]. However, the exact mechanism by which one CD strain excludes another is not defined.

**Adherence**

Gut pathogens must associate with the mucosal epithelium to begin the process of pathogenesis. Compared to the wealth of information that has accumulated about the toxins, relatively little is known about the interaction of *C. difficile* with the gut. While the toxins are secreted and purifiable, investigations of molecules on the bacterial surface have been much more difficult, and have been limited by the lack of genetic tools available until recently.

Once the hamster model of CDI was established, early work focused on strain differences in colonization factors. In 1988, Borriello examined strain differences in proliferation in various gut locations, and attempted to relate adherence to gross features such as flagella or fimbriae, but the conclusions that could be drawn were limited [224]. In 1991, Gonzalez-Valencia [225] attempted to correlate CD obtained from symptomatic or asymptomatic patients with the adherence to HEP2 cells in culture. Results were inconclusive, which is not surprising considering the multiple intrinsic and extrinsic factors that determine a patient’s symptomatic status.

Multiple *C. difficile* adhesins have been described by the research group of Tuomo Karjalainen and Anne Collignon. Their early work found that heat-shocked CD grown in the presence of blood were more adherent than non-heat-shocked [226],
but the blood effect later turned out to be an error [227]. The heat labile protein identified by antigenic screening of a CD genomic library was a 27 kD adhesin that was not characterized further [226]. Other putative adhesins identified by this group have been the flagellin FliC, the flagellar cap FliD [228], and a fibronectin-binding protein Fbp68 [229], a heat-shock protein, GroEL [230]. Adherence of CD is reduced in the presence of anti-GroEL antibodies [230], although heat shock proteins are not usually displayed on the bacterial surface, so the mechanism for this effect is unknown.

The best characterized adhesins so far are two; the surface associated adhesin Cwp66 [231], and the surface layer protein SlpA [232-234]. Cwp66 was first described as a heat-shock induced adhesin [231] that is located on the surface [182, 235]. Initial work indicated that it was transcribed in early log phase [236], although later work reported that it was still highly transcribed in overnight cultures, and could be up-regulated by osmotic stress or the presence of antibiotics [237].

The S-layer protein SlpA was first described as one of a number of putative-surface associated genes located in a cluster of 17 ORF’s along with cwp66 (Figure 6) [238]. These genes share homology to the cell-wall anchoring domain of CwlB/LytC of B. subtilis, indicating that these genes most likely code for surface-associated proteins [238].

S-layer proteins are secreted by many bacterial species and self-assemble into an ordered lattice on the cell surface. In addition to providing structural integrity to the cells, and acting as molecular sieves, S-layer proteins have been implicated in
adherence and immune evasion [239]. S-layer proteins have been demonstrated to be responsible for the adherence to the epithelia in several species of Lactobacilli [240, 241]. S-layer proteins were the first prokaryotic proteins demonstrated to be true glycoproteins [242]. On the bacterial surface, glycosylation can have multiple functions, including stability, and evasion of the immune system of the host by blocking complement mediated lysis [243-245] [246].
Figure 6. Arrangement of genes containing \( slpA \) and its paralogs (ORFs 1–12) from CD strain 630. Black shading indicates regions of homology to the cell wall binding domain, and the non-homologous regions are shown in white. Dark grey indicates non-surface associated genes. Figure adapted from A. Wright et al. 2005.
The *C. difficile* S-layer protein SlpA is unique in that it contains two subunits [247], the high-molecular weight (HMW) and the low molecular weight (LMW) proteins, which are cleaved from a common precursor by the surface-associated protease Cwp84 [248] [234], and assemble on the bacterial surface into the paracrystalline lattice [247]. Because it covers the whole surface, SlpA is the most abundantly produced protein in the cell. The interaction of the two SlpA subunits is so robust that they assemble into lattices *in vitro* after denaturation and renaturation [249]. The crystal structure of the LMW subunit from strain 630 has recently been solved, along with a solution structure model of the two subunits interacting with each other [250] (Figure 7). The two subunits associate with each other with high affinity through the N-terminal of HMW subunit and the C-terminal of LMW subunit.

Characterized as an adhesin by Neil Fairweather’s group, studies have shown that recombinant SlpA proteins bind to host tissues and extracellular matrix (ECM) proteins, and that CD adherence is decreased in the presence of anti-SLP antisera [232]. Concerning the relative function of the two subunits, Takeoka [247] has shown that both subunits are displayed on the exterior surface using immunogold electron microscopy. Takumi [249] found that antibodies to the LMW subunit reduced adherence more than antibodies to the HMW subunit. In contrast, Calabi [232] found that recombinant HMW subunit protein bound to gastrointestinal tissue while the LMW subunit did not, and that antibodies to the HMW subunit inhibited binding of the protein to Hep2 cells.
Figure 7. Small-Angle-X-ray-Scattering (SAXS) structure of *C. difficile* SlpA from strain 630. A and B. Two orientations, differing by a 90° rotation, are shown of the complex of the HMW (red) and LMW (white) subunits. B. Model of the orientation of the HMW and LMW SLPs on the surface of *C. difficile*. The SAXS structure of the complex is shown as above with a ribbon representation of the LMW crystal structure overlaid in blue. The HMW SLP is also shown interacting with the cell wall; however, the extent and exact nature of this interaction is currently unknown. Figure and description adapted from Fagan et al, Mol. Micro. 2009.
While it was initially appeared that SlpA was glycosylated, and this modification was hypothesized to influence host interactions, [234], later work using mass spectrometry definitively demonstrated that SlpA was not glycosylated, and that previous results were most likely due to SlpA preparations contaminated with peptidoglycan [251].

This work investigated the interaction between hypervirulent C. difficile and the epithelia, and characterized one of the surface proteins, SlpA, that mediates this interaction.
CHAPTER TWO
MATERIALS AND METHODS

*C. difficile* Strains and Media

*C. difficile* human clinical isolates were obtained from the culture collection of Dr. Dale Gerding (Table 1). Four HV *C. difficile* strains isolated from geographically distinct regions were chosen, as was BI-1, a historical isolate from the BI genetic group that predates the epidemics. The non-HV toxigenic strains included strain 630, the first fully sequenced strain, and strain VPI 10463, a known high-toxin producer [213], which are both very rarely found in clinical settings. Strains J9 and K14 (representative of the REA type J and K groups respectively) have caused hospital outbreaks, are frequently isolated from hospital settings in the USA, but have never been reported as a cause of increased CDI severity and are not referred to as ‘hypervirulent’ [95, 189, 192, 200]. Strains M3, M23 and T7 are non-toxigenic.

**Table 1 C. difficile strains used in this study.**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Source</th>
<th>Year</th>
<th>Source</th>
<th>Toxigenic</th>
<th>Hypervirulent</th>
</tr>
</thead>
<tbody>
<tr>
<td>630</td>
<td>Patient</td>
<td>1982</td>
<td>Switzerland</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>VPI 10463</td>
<td>Patient</td>
<td>1980</td>
<td>Eastern USA</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>J9</td>
<td>Patient</td>
<td>1987</td>
<td>Illinois</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Patient</td>
<td>unknown</td>
<td>unknown</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>---</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>J32</td>
<td>Patient</td>
<td>unknown</td>
<td>unknown</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>K14</td>
<td>Patient</td>
<td>1994</td>
<td>Illinois</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>K29</td>
<td>Patient</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BI-1</td>
<td>Patient</td>
<td>1988</td>
<td>Minnesota</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>BI-6</td>
<td>Patient</td>
<td>2003</td>
<td>Oregon</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>BI-8</td>
<td>Patient</td>
<td>2004</td>
<td>Maine</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>BI-17</td>
<td>Patient</td>
<td>2004</td>
<td>Montreal</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>BI-23</td>
<td>Patient</td>
<td>2007</td>
<td>Eastern USA</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>BI-moxi</td>
<td>Patient</td>
<td>unknown</td>
<td>unknown</td>
<td>Yes</td>
<td>Yes, Moxifloxacin susceptible</td>
</tr>
<tr>
<td>M3</td>
<td>Asymptomatic patient</td>
<td>1989</td>
<td>Minnesota</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>M23</td>
<td>Asymptomatic patient</td>
<td>1991</td>
<td>Minnesota</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>T7</td>
<td>Asymptomatic patient</td>
<td>1986</td>
<td>Minnesota</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

All strains were routinely cultured in Brain-Heart Infusion (BHI) Broth or on BHI-agar plates (Difco, Buchs, Switzerland; 37g/L) in a Coy anaerobic chamber (Grasslake, MI) with 5% CO₂, 5% H₂ and 90% N₂. For all assays, *C. difficile* strains were
grown to saturation in brain-heart infusion broth (BHI, BD Biosciences, Boston, MA) overnight. 1 ml aliquots were clarified by centrifugation at 2000xg. Bacterial pellets were washed in phosphate-buffered saline (PBS), resuspended in fresh BHI broth at a ratio of 1 to 50, and allowed to grow without agitation under anaerobic conditions (90% N₂, 5%H₂ and 5% CO₂) in a Coy glove-box (Coy, Grasslake, Michigan). Identical inocula were used for all strains and all growth experiments. For growth curves, optical density readings (600 nm wavelength) were taken at intervals of one hour or less for the first 18 hours, and again at 24 and 48 hours.

*E. coli* Strains and Media

Unless indicated otherwise, *Escherichia coli* strains (Table 2) were grown in Luria-Bertani (LB) broth (1% w/v tryptone, 0.5% w/v yeast extract, 1% sodium chloride). For selection purposes during cloning and expression the following antibiotics were used at the following concentrations in liquid and solid media: Ampicillin, 200 ug/ml; Carbenicillin, 50 ug/ml; Kanamycin, 25 ug/ml; Chloramphenicol, 25 ug/ml.

Table 2. *E. coli* strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant characteristic(s)¹</th>
<th>Source or Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> Top-10</td>
<td><em>recA1, endA1</em></td>
<td>Invitrogen</td>
</tr>
<tr>
<td><em>E. coli</em> DH5α</td>
<td><em>recA1, endA1</em></td>
<td>Invitrogen</td>
</tr>
<tr>
<td><em>E. coli</em> Rosetta</td>
<td>Bl-21 derivative; pRARE plasmid contains 9 tRNA genes for rare codon expression</td>
<td>Novagen</td>
</tr>
</tbody>
</table>
Anaerobic Bacterial Adherence Assay

To quantitate CD attachment to human host cells, I devised an anaerobic bacterial adherence assay using a derivative of the Caco-2 human intestinal epithelial cell-line, Caco-2BBE. C2BBE host cells were cultured in high-glucose (25mM) Dulbecco’s Modified Eagle Medium (DMEM), 10% fetal bovine serum, 20mM HEPES, 100 IU/ml penicillin, and 100 mg/ml streptomycin at 37°C in the presence of 5% CO2. Cells between passages 25 and 45 were grown to confluent monolayers (1.2x10^6 cells) in 6-well plates, and transferred to antibiotic and serum-free DMEM 24 hours prior to adherence assays. All assay solutions were pre-reduced in the anaerobic chamber overnight. DMEM with 25mM CaCl₂ was made by adding 1 mL of 1M CaCl₂ to 40 mL of DMEM immediately before introduction to the anaerobic chamber. For the assays, C2BBE plates were introduced into the anaerobic chamber just before use, serum-free medium was removed, and exponential phase *C. difficile* applied at a multiplicity of infection of 20 in a total volume of 2mL DMEM. All bacterial strains used were washed and resuspended in anaerobic DMEM with 25mM CaCl₂ (800 uL into 10 mL of DMEM) prior to incubation with host cells. 1 ml of DMEM with 25mM CaCl₂ was applied to all wells, and additional 1 mL of inoculum was added to test wells and 1mL of DMEM with 25mM CaCl₂ to control wells.

Adherence was allowed to proceed under anaerobic conditions for 40 minutes. Host cells and adherent bacteria were then washed twice with 1mL of anaerobic phosphate-buffered saline (PBS), scraped, serially diluted and plated to enumerate adherent *C. difficile*. Each experiment was performed in quadruplicate, and repeated at
least three times in entirety. The percent adherence was calculated as the ratio of recovered *C. difficile* to input *C. difficile* multiplied by 100.

C2BBE cells survive the anaerobic conditions under which our experiments were performed, as confirmed by Live-Dead staining (Invitrogen).

**Immunofluorescence Microscopy**

C2BBE monolayers were grown on 12mm-diameter coverslips. Cells were placed in antibiotic- and serum-free cell culture medium containing 50% DMEM, 50% F-12, 1.2% sodium bicarbonate, 3.56% HEPES, and 0.5% mannose overnight and then used in anaerobic adherence assays as described above. For immunofluorescence staining, infected monolayers were rinsed in PBS and then fixed in 3.75% paraformaldehyde for 15 minutes. Fixed cells were quenched with PBS containing 75mM NH₄Cl and 20mM glycine for 15 minutes. Cells were then washed twice in PBS and permeabilized with PBS containing 0.5% TritonX-100 for 15 minutes. Cells were washed once more with PBS and blocked for 1 hour in 5% bovine serum albumin (BSA). For analysis of the zona occludens 1 (ZO-1) protein, monolayers were incubated with 1:50 dilution of the mouse anti-ZO-1 (Zymed Laboratories, South San Francisco, CA) for 1 hour and then probed with mouse-specific secondary antibodies coupled with Alexa Fluor 568 (Molecular Probes, Eugene, OR) for 1 hour at room temperature. For actin visualization, host cells were incubated with 1:100 dilution of BODIPY® 558/568 phalloidin (Molecular Probes, Eugene, OR) for 30 minutes at room temperature. For visualization of *C. difficile*, samples were incubated with 1:20 dilution of rabbit anti-Clostridial spp antiserum conjugated with FITC (ViroStat, Portland, ME) for 1 hour at room temperature. Stained
samples were mounted with Prolong Gold Antifade Reagent (Molecular Probes, Eugene, OR) supplemented with 1ug/mL 4',6-diamidino-2-phenylindole (DAPI). A Leica DM4000B microscope (Leica, Wetzlar, Germany) was used for visualization, and images were documented using SlideBook 4.2 software (Olympus Imaging Systems, Center Valley, PA). All images were captured at a magnification of 1000X.

**Total Surface Protein Interference Assay**

Using the anaerobic adherence protocol described above, confluent C2BBE monolayers were incubated for 20 minutes (prior to the addition of bacteria) with DMEM plus equal volumes of PBS containing increasing amounts of purified, neutralized, dialyzed, anaerobic total surface-layer protein (SLP) preparations (isolation described below). Exponential phase *C. difficile* were then added to the monolayers, and adherence allowed to proceed for another 20 minutes as described above.

**Antibody Interference Assay**

Exponential phase *C. difficile* were resuspended in DMEM with 25mM CaCl₂ and then incubated with anti-LMW SlpA or anti-HMW SlpA, or both, antisera at a dilution of 1:1000 for one hour before addition to confluent C2BBE monolayers. Control experiments using *C. difficile* enumeration on BHI-agar plates were performed to confirm that there was no further growth/death of bacteria during this one hour incubation. Control antisera included anti-6-histidine antiserum (Qiagen, Valencia, CA), anti-TraG antiserum (an irrelevant *E. coli* conjugation protein, non-commercial) and an anti-clostridial antiserum (Abcam, Cambridge, MA; antiserum does not contain *C. difficile*-
derived antibodies) were used as controls in the resulting adherence assays (at dilutions of 1:1000 each).

**Total soluble protein isolation**

To obtain total cellular protein, bacteria were grown to exponential phase (O.D.\textsubscript{600nm} 0.5) in 37g/L BHI broth, harvested, and lysed by sonication (55% power; 12 pulses of 15 seconds each). Cellular debris was removed by centrifugation at 6500g for 30 minutes at 4°C. A protease inhibitor solution (EDTA-free Complete Cocktail, 1X final concentration, Roche, Indianapolis, IN) was added to the resultant supernate. This supernate was then centrifuged at 265,000g for 2.5 hr at 4°C to fractionate the sample into soluble proteins in the supernate and insoluble proteins in the pellet. The pellet was washed with PBS, and proteins dispersed with gentle sonication (45% power, 3 pulses of 15 seconds each). Equal amounts (30ug) of total soluble proteins were subjected to SDS-PAGE on 15% Tris-HCl gels (Biorad, Hercules, CA), and stained with Gel-Code Blue (Pierce, Rockford, IL) to visualize protein bands.

**S-layer protein extraction**

SlpA and other surface-layer proteins (SLPs) were extracted from multiple *C. difficile* strains using 0.2M glycine pH2.2, as described by Calabi et al [234]. Briefly, 50mL of *C. difficile* culture grown in BHI was harvested at exponential phase by centrifugation (3000g for 20 minutes), washed in PBS, and resuspended in 200µl of 0.2M glycine pH2.2 and incubated at room temperature for 30 minutes. After centrifugation to remove the cell pellet (16,000g for 15 minutes at 4°C), the resultant supernate containing surface proteins (SLPs) was stored at -20°C until further use. SLPs used for adherence
interference assays were subsequently dialyzed into 10 volumes of PBS using 10kDa molecular weight cutoff centrifugation-based filters (Millipore, Billerica, MA).

**Protein Quantitation**

Total CD soluble protein, surface layer extracts, and recombinant protein were measured using the Bicinchoninic Acid (BCA) Protein Assay Kit (Pierce), according to the manufacturer’s instructions. A standard curve constructed using a gradient of known concentrations of bovine serum albumin (BSA) was prepared and tested for each assessment.

**DNA sequencing of slpA**

For sequencing, genomic DNA was isolated from exponential phase cultures using Qiagen DNeasy columns (Germantown, Maryland), according to the manufacturer’s instructions. DNA concentration was determined by a Beckman spectrophotometer, and DNA was aliquoted and frozen at –20°C.

Since SlpA is highly variable, primers to amplify and sequence the slpA gene were designed inward from neighboring conserved regions. The upstream primer (slpAF: ATGTTGGAGGAATTTAAGAAATG) was designed to include part of the conserved signal sequence of SlpA. The downstream primer (slpAR: ACCTCCACCAGTTTTTCATCTCTGC) was designed inside of SecA [252]. 100 ng of genomic DNA was used as template for PCR reactions with primers at 40 pmol concentration. I used the Failsafe PCR system (Epicentre Biotechnologies, Madison, WI) consisting of the Failsafe Polymerase and a buffer mix (Buffer E) demonstrated to amplify *C. difficile* DNA. Reactions of 50ul volume were amplified for 32 cycles
including a 94°C denaturation step (30 seconds), annealing from 45-53°C (2 min), and a 72°C extension step for 3 min. Annealing temperature were optimized for each strain using the temperature gradient feature of the BioRad i-cycler (Hercules, CA). PCR products were visualized using agarose gel electrophoresis, and purified by QiaQuick PCR purification columns (Qiagen, Germantown, MD), according to manufacturer’s instructions. If multiple bands occurred, the band of the appropriate size was gel purified using the Qiagen gel extraction kit. 20ng of purified PCR products were sent to the University of Florida Sequencing Core Laboratory. Each slpA gene from each strain was sequenced on both strands using initial amplification primers and subsequent primers derived from the sequences obtained. Sequence data was assembled and analyzed using Vector NTI software (Invitrogen, Carlsbad, CA).

Immunodetection

Western blotting experiments were performed on total soluble protein and surface protein extracts. 30ug of total soluble protein and 5ug of SLP extracts were electrophoresed on denaturing 4-20% gradient Tris-HCl acrylamide gels (Biorad, Hercules, CA) and transferred overnight to 0.45 uM nitrocellulose membrane at 50 volts using neutral 1x TG buffer (2.5 mM Tris-Cl, 19.2% Glycine) at 4°C in a Trans-Blot cell (BioRad, Hercules, CA). Membranes were blocked for 1 hour at room temperature or overnight at 4°C, using 1% blocker from the Roche Western Blotting Kit. Primary antisera to the HMW and LMW SlpA subunits were used at 1:100,000 dilution. Primary antibodies were incubated for 60 mins in 0.5% blocker in TBS, at room temperature with shaking. Membranes were washed three times with TBST (50 mM Tris-Cl, 150 mM NaCl, pH 7.5 and 0.1% Tween
20) for 10 minutes. The secondary antibody was goat anti-rabbit IgG-POD conjugate (Roche, Indianapolis, IN). Secondary antibodies were incubated for 30 minutes, then washed for four times with TBST for 15 minutes. Proteins were visualized using the POD chemiluminescent detection system in the Roche Western Blotting kit according to manufacturer’s directions (Roche, Indianapolis, IN).

**Proteomic Identification of *C. difficile* Surface Proteins**

Surface protein extracts were electrophoresed using denaturing 4-20% Tris-HCl PAGE (Biorad, Hercules, CA) and stained with Coomassie Brilliant Blue. Bands of interest were excised, and proteins identified using liquid chromatography/mass spectrometry (LC/MS) analyses. All mass spectrometry analyses were performed at the University of Minnesota Mass Spectrometry Consortium.

**Molecular Cloning Techniques**

To produce recombinant subunits, the portions of *slpA* corresponding to the LMW subunit and the HMW subunit were each cloned individually from three CD strains (Figure 8). Based on the crystal structure of SlpA [250], I also designed a construct of the LMW subunit from strain 630 lacking the region required for interaction with the HMW subunit (Figure 7). Gene fragments were amplified from CD genomic DNA using the Failsafe PCR system (Epicentre Biotechnologies, Madison, WI) as described above. Using the Gateway System of entry and destination vectors (Invitrogen), amplified products were ligated into the entry vector pENTR/SD/D-TOPO. The entry vector was recombined with the destination vector pET-DEST-42, which contains an IPTG-inducible promoter, and a C-terminal V5 and 6x-histidine tag.
Figure 8. Schematic representation of *slpA* constructs cloned for the expression and synthesis of recombinant protein.
Following subcloning into *E. coli* DH5α, pET-DEST-42 expression vectors with *slpA* constructs were transformed into Rosetta *E. coli*. The Rosetta strain of *E. coli* expresses several rare tRNA molecules that correct the codon bias of *E. coli*, making it more suitable for expression of *C. difficile* proteins.

**Table 3. Plasmids used in this work.**

<table>
<thead>
<tr>
<th>Name</th>
<th>Genotype/Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pENTR/SD/D-TOPO</td>
<td>Entry Vector, Kan+</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pET-DEST-42</td>
<td>Expression vector, Amp+</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>pET-DEST-42-630LMW</td>
<td>Strain 630 LMW subunit</td>
<td>This work</td>
</tr>
<tr>
<td>pET-DEST-42-630HMW</td>
<td>Strain 630 HMW subunit</td>
<td>This work</td>
</tr>
<tr>
<td>pET-DEST-42-630LMWtrunc</td>
<td>Strain 630 LMW subunit with 3’ truncation</td>
<td>This work</td>
</tr>
<tr>
<td>pET-DEST-42-K14LMW</td>
<td>Strain K14 LMW subunit</td>
<td>This work</td>
</tr>
<tr>
<td>pET-DEST-42-K14HMW</td>
<td>Strain K14 HMW subunit</td>
<td>This work</td>
</tr>
<tr>
<td>pET-DEST-42-BI17LMW</td>
<td>Strain BI17LMW subunit</td>
<td>This work</td>
</tr>
<tr>
<td>pET-DEST-42-BI17HMW</td>
<td>Strain BI17 HMW subunit</td>
<td>This work</td>
</tr>
</tbody>
</table>

**Purification of recombinant proteins**

The conditions for recombinant protein synthesis and purification differed for different constructs. Briefly, Rosetta *E. coli* containing the constructs was grown and induced using Novagen Overnight Express Autoinduction Terrific Broth (TB) medium. Following induction, bacteria were pelleted by centrifugation, and frozen at -80° C.
Pellets were later lysed using Bugbuster (Novagen) in the presence of protease inhibitors (EDTA-free Complete Cocktail, 1X final concentration, Roche, Indianapolis, IN). For the LMW subunit of 630, recombinant protein was purified from the soluble fraction of cell lysates using cobalt affinity chromatography (Talon-spin columns, Clontech). The HMW subunit of 630 and the LMW subunit of K14 both degraded in the soluble fraction, so these constructs were isolated from the insoluble fraction, based on the protocol of Fagan et al [250]. Briefly, cell pellets were lysed using BugBuster lysis buffer (Novagen), and insoluble protein from inclusion bodies was purified according to manufacturer’s directions. These inclusion bodies were dissolved in 8M urea, 150mM NaCl 10mM HEPES, and subjected to cobalt affinity chromatography. The purified recombinant proteins were refolded by step-wise dialysis to native buffer conditions (150mM NaCl 10mM HEPES). This method yielded soluble refolded proteins.

The K14 HMW construct degraded extensively in _E. coli_, so I employed an alternate approach to isolate this subunit from _C. difficile_ surface extracts. Surface proteins were extracted from an exponential phase culture of strain K14 using the acid-glycine method described above. These extracts were mixed with 8M urea and Laemmli loading buffer, heated at 70°C for 15 minutes and subsequently electrophoresed in an SDS-PAGE gel. Portions of the unstained gel corresponding to the HMW subunit were excised, minced, and placed in the vertical tubes of a BioRad electroelution apparatus, and allowed to migrate out of the gel slices over three hours. The electro-eluted protein solution was dialyzed against native buffer to remove SDS, denatured in 8M urea and refolded as described above.
**Recombinant SlpA Protein Interference Assay**

For technical reasons, I used the parental Caco2 cell line, which exhibits similar adherence to the Caco2-BBE line (not shown). To conserve recombinant protein, volumes were scaled down to conduct the assay in 24 well rather than 6 well culture plates. Similar to the Total Surface Protein interference assays described above, confluent Caco2 monolayers were incubated for 20 minutes (prior to the addition of exponential phase strain 630 bacteria) with 250 uL DMEM plus 25mM CaCl₂ and 250 uL of 150mM NaCl 10mM HEPES buffer containing increasing amounts of anaerobic recombinant protein. Exponential-phase *C. difficile* (MOI 20, 250 uL volume of inoculum prepared as described above) were then added to the monolayers, and adherence allowed to proceed for another 20 minutes as described above.

**Toxin ELISA**

For toxin testing using an ELISA, culture samples were clarified, and supernatant fluids were sterile filtered and frozen at -80°C prior to use. All growth experiments were performed in entirety at least three times. For toxin quantitation, culture supernatant fluids were collected at mid-exponential phase (OD₆₀₀ of 0.5), early stationary phase (defined by two consecutive non-increasing OD₆₀₀ readings; about 10-12 hours), mid-stationary phase (15hrs), and at 18, 24, and 48 hours of growth.

Toxin amounts were quantitated using the Wampole Tox A/B II kit (TechLabs, Inc, Blacksburg, VA). Purified Toxin B (provided by TechLabs, Inc) was used to construct a standard curve. Samples of the supernatant fluids were diluted to fall within the linear range of the standard curve, and this dilution was used to calculate the amount
of toxins present. All samples were tested in triplicate, and each experiment performed in entirety at least three times.

**RNA Isolation and cDNA synthesis**

For toxin gene expression assays, total RNA was isolated from exponential phase (OD$_{600nm} = 0.5$) and stationary phase (12 hours) cultures. Five milliliters of culture were harvested by centrifugation at 2800xg for 10 minutes at 4°C. Bacterial pellets were processed immediately or resuspended in lysis buffer, flash-frozen in liquid nitrogen and stored at -80°C. Lysis and extraction of RNA were performed using the Ambion RiboPure Bacteria RNA kit (Ambion, Austin TX) according to the manufacturer’s instructions, followed by DNaseI digestion using Ambion Turbo DNase. DNaseI digestion was repeated twice for all samples. The RNA obtained after each DNaseI digestion was purified using the Qiagen RNeasy RNA column purification kit, according to the manufacturer's instructions (Qiagen, Germantown, MD), quantitated using a NanoDrop spectrophotometer, aliquoted and stored at -20°C. RNA quality was assessed spectrophotometrically (260/280 nm) and by visualization on denaturing formamide/formaldehyde gels (not shown). 500 ng of pure RNA from each sample was converted to cDNA using random hexamers and the BioRad iScript cDNA synthesis kit (Biorad, Hercules, CA).

**Quantitative Real-time PCR.**

Primers used to amplify PaLoc genes are shown in Table 4, and were either synthesized using previously published sequences (for *tcdA*, non-HV strain *tcdB*, and *tcdC*; [103], or specifically designed for HV strain *tcdB* as well as for *tcdR* and *rpoA*.
Table 4. Primers used in this study for Quantitative Real-time PCR.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence 5’ to 3’</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>JRP3443Aup</td>
<td>GTCGGATTGCAAGTAATTGACAATA</td>
<td>tcdA-specific forward primer</td>
</tr>
<tr>
<td>JRP3444Adn</td>
<td>TAAACAGTCTGCAACCTTTTGAGA</td>
<td>tcdA-specific reverse primer</td>
</tr>
<tr>
<td>JRP4055Bup</td>
<td>ACCATATAGCTTTGTAGATAGTGAAGGAAA</td>
<td>tcdB-specific forward primer non-HV strains</td>
</tr>
<tr>
<td>JRP4056Bdn</td>
<td>AAGAACTACATCGGTCAAATTCAGATACAAAA</td>
<td>tcdB-specific reverse primer Non-HV strains</td>
</tr>
<tr>
<td>tcdB4B1up</td>
<td>AGCTGCTTCAGTGAGAGAAA</td>
<td>tcdB-specific forward primer HV strains</td>
</tr>
<tr>
<td>tcdB4B1dn</td>
<td>AATCAATTGCTCCCCCTCT</td>
<td>tcdB-specific reverse primer HV strains</td>
</tr>
<tr>
<td>JRP4053Cup</td>
<td>AGGTATTGCTCTACTGGCATTTTATT</td>
<td>tcdC-specific forward primer</td>
</tr>
<tr>
<td>JRP4054Cdn</td>
<td>CCTCATGGTCTTCAGAACAAGCT</td>
<td>tcdC-specific reverse primer</td>
</tr>
<tr>
<td>tcdR3up</td>
<td>ATCAAGTAAAGTCTGTTTTTGAGGAAG</td>
<td>tcdR-specific forward primer</td>
</tr>
<tr>
<td>tcdR3dn</td>
<td>TGCTCTATTATTAGCCTTATTAACAGC</td>
<td>tcdR-specific reverse primer</td>
</tr>
<tr>
<td>rpoA2up</td>
<td>TCATTACCGGTGAGCAGTGAA</td>
<td>rpoA-specific forward</td>
</tr>
</tbody>
</table>
All primers were tested in genomic DNA amplification for all test strains to confirm specificity and amplicon size, before use in quantitative reverse-transcriptase (qRT) PCR reactions. To test for DNA contamination, RT-PCR reactions were performed for all samples using RNA alone, and found to be negative (not shown). Expression levels of all PaLoc genes tested in each isolate were normalized using the *C. difficile* housekeeping gene *rpoA* as a reference [102]. Transcription rates for individual PaLoc genes relative to each other among different isolates were not determined since the efficiency of reverse transcription varies for different genes, making between-gene transcription rate comparisons inaccurate.

A gradient of *C. difficile* genomic DNA was used to determine the efficiency of amplification of each primer set. Efficiencies ranged from 90-101% and R² values were at least 98%. For all samples tested, 25 µl reactions were performed in triplicate using 5PRIME Real Master mix SYBR ROX (Fisher, Pittsburgh, PA), 1 µl of cDNA, and 250 nM primers in an Eppendorf MasterCycler equipped for qRT-PCR detection [102]. Specificity of product was determined by dissociation (melt) curve analysis. PaLoc gene expression was analyzed relative to that of the reference gene using the ΔCt method, according to the following formula: $2^{\Delta C_{\text{Ref}} - \Delta C_{\text{Test}}}$ [253].

**Sporulation Assays.**
To assess the presence of spores over time, I employed plating and microscopy. At 8, 20, 28, and 48 hours post-inoculation, 1 mL samples of each strain tested were clarified by centrifugation, bacterial pellets washed in PBS, heat-shocked at 65°C for 15 minutes (to kill vegetative cells), serially diluted and plated on taurocholate (a spore germinant) fructose agar (TFA) plates to enumerate spores [221]. Plates were incubated anaerobically for 48 hours, after which CFUs were enumerated.

For microscopy, 10 µL samples of 48 hour cultures were applied to microscope slides, oven-dried at 85 °C for 5 minutes, and Gram-stained. Spores were defined as all refractile bodies, whether free or still attached to mother cell material. For each strain, ten distinct fields were photographed and enumerated at 100X magnification under oil-immersion.

Statistical Analyses

The SPSS (SPSS, Chicago, IL) and StatView (SAS, San Francisco, CA) software packages were used for statistical analyses. Significance was determined using analysis of Variance (ANOVA) to enable comparison between multiple groups of continuous numerical data. The Protected Least Significant Difference test was used for posthoc analyses.
CHAPTER THREE

RESULTS

The Adherence of CD to Host Cells and the Role of SlpA in this Interaction

Introduction

From the perspective of a commensal bacterium, the gut is a rich but unstable environment. To avoid being completely shed in the fecal stream by the action of peristalsis, bacteria must adhere to the host. Adherence can be direct between the bacterium and the host cell, or extracellular structures such as biofilms can accomplish this task. As an opportunistic pathogen, \textit{C. difficile} faces the same problem. Additionally, if CD produces toxins to destroy host cells for the purpose of extracting nutrients, they must maintain proximity to host epithelia to do so. If adherence is an adaptation, are more adherent strains more successful? In particular, if hypervirulent CD strains are successful pathogens, do factors in colonization or adherence contribute to this phenotype?

To investigate the role of non-toxin proteins in hypervirulent strains, my approach was to examine colonization using an in vitro model and, in tandem, to examine total protein profiles, and discover factors that might vary in hypervirulent strains.

To achieve this first goal, I developed an anaerobic assay to test the adherence of \textit{C. difficile} strains to human intestinal epithelial cells grown in culture. For \textit{in vitro} studies, few reports appear in the literature for which live, whole \textit{C. difficile} bacteria have been used. For example, most studies exploring \textit{C. difficile} attachment to host epithelia
have been performed using non-intestinal cell lines, and under aerobic conditions that fail to recover viable bacteria [227, 230, 254].

At the molecular level, mechanisms of *C. difficile* pathogenesis have been difficult to elucidate since the organisms are not readily genetically tractable. *C. difficile* cannot be reliably electroporated or transformed, and conjugation-based methods of DNA introduction into *C. difficile* have been hampered by lack of suitable cloning vectors, and the presence of restriction/methylation barriers that hinder heterologous DNA maintenance [255, 256]. In recent years, two genetic tools for *C. difficile* gene disruption have been developed, one based on an unstable *C. perfringens* vector and another based on a targetable Group II intron [102, 257]. While these tools represent a watershed advance in CD research, they were not well developed at the initiation of this work. Further, these tools are not suitable for genome-wide genetic screens. Given these limitations, we decided to use a protein-based approach to initially identify unique and divergent proteins that might be involved in hypervirulence.

**C. difficile strains differ in adherence to host epithelia in culture**

Using immunofluorescence microscopy, I visualized the adherence of bacteria with host cells, as shown in Figure 9. Bacteria are distributed over the whole monolayer, and do not seem to be associated with particular structures such as the cell-cell borders. This assay also confirmed host cells exhibited morphology consistent with viability under anaerobic conditions (well rounded nuclei with uniform staining, normal actin stress fibers and uniformly distributed ZO-1 around host-cell peripheries).
Figure 9: Microscopy-based visualization of *C. difficile* adherence. Bacteria (green) were stained with anti-clostridial sp antiserum conjugated to FITC; host-cell nuclei (blue) were stained with 4',6-diamidino-2-phenylindole (DAPI), host cell actin (red) was stained with BODIPY® 558/568 phalloidin (panel A) and host-cell ZO-1 (red)
was stained with anti-ZO-I antiserum conjugated to Alexa Fluor 568 (panel B). Images shown were captured at a 1000X magnification, are representative of multiple fields visualized, and are from two independently performed experiments.
I began my comparison of HV strain to other toxigenic CD strains by developing a quantitative in vitro test of adherence to host cells. Hypervirulent strains were compared to the toxigenic but non-hypervirulent *C. difficile* strains K14, J9 and J32. As shown in Figure 10, I found that three out of five BI strains tested exhibited $\geq 100\%$ higher adherence to C2BBE cells than toxigenic but non-hypervirulent strains ($p \leq 0.0001$). From these data, it appeared that hypervirulent *C. difficile* strains had increased adherence to human host cells. This was the first indication that non-toxin proteins, particularly those involved in bacterial attachment, might be different in these strains.
Figure 10: Adherence of *C. difficile* strains to C2BBE monolayers. Three non-hypervirulent strains (K14, J9, J32) and five hypervirulent strains (BI-6, BI-8, BI-17, BI-23 and a moxifloxacin-susceptible BI isolate) were tested. Percentage adherence with standard error of the mean is depicted; all experiments were performed in quadruplicate, and repeated in entirety at least three times.
Adherence of CD strains over the growth cycle

The behavior of CD changes with nutrient availability. When nutrients are limited, bacteria begin to produce toxins. Does adherence of CD change over during different growth phases? To answer this question, I tested the adherence of several strains of CD during stationary phase. I determined the adherence of strains 630, K14 and BI17 during exponential growth and during early stationary phase. Figure 11A shows the mean adherence of multiple independent assays. HV Strain BI17 showed a significant decrease in adherence in stationary phase (p<.05). Strain 630 showed a trend to decreased adherence (p=.07), while strain K14 did not appreciable decline. Figure 11B shows the means of single assays of four other strains, which also show trends to exhibit lower adherence in stationary phase in two (J32 and M3) of the four strains.
Figure 11. Adherence of *C. difficile* strains to C2BBE monolayers. A. Three strains 630, K14 and BI17 were tested at exponential phase and early stationary phase. Percentage adherence with standard error of the mean is depicted; experiments were performed in triplicate, and repeated in entirety two or three times. B. Adherence of four CD strain. Percentage adherence of a single quadruplicate assays are shown.
Protein profiles of HV CD and the identification of SlpA

In parallel, I wanted to discover other factor that might be different in HV CD as compared to toxogenic CD. To explore potential protein variation in hypervirulent *C. difficile*, I examined total soluble protein profiles of hypervirulent strains, as well as those of the toxigenic (but non-hypervirulent) *C. difficile* strains 630, J9 and K14. Total soluble protein profiles on SDS-PAGE gels were compared between strains. Bands that appeared to differ between strains were subjected to MALDI mass spectrometry analyses for protein identification. One strong band that differed between strains was identified as the *C. difficile* Surface Layer (S-layer) protein SlpA. As shown in Figure 12A, a prominent protein band in the BI-6, BI-8 and BI-17 strains that appeared altered in amount compared to that of strains 630 and J9. If a surface protein was variable, this might be related to factors involved in colonization and host interaction.
Figure 12. *C. difficile* protein studies. Panel A: Total soluble protein profiles of *C. difficile* strains. 630, J9 and K14 are toxigenic (but non-epidemic strains); BI-6, BI-8 and BI-17 are hypervirulent strains. Arrows indicate bands excised for mass spectrometry analysis. Panel B: Extracted surface-layer protein (SLP) profiles of *C. difficile* strains. Eleven BI-17 bands were identified by mass spectrometry (one round; peptide mass determination); identities of 7 bands and their Genbank Accession numbers are shown in Panel D. Panel C: Western blot analyses of SLP preparations from *C. difficile* strains. For all *C. difficile* strains tested, 30ug of total soluble and 5ug of SLP preparations were electrophoresed; antisera were used at a 1:100,000 dilution.
As mentioned earlier, S-layer proteins are secreted by many bacterial species and self-assemble into an ordered lattice on the cell surface. The *C. difficile* S-layer contains two subunits [247], the high-molecular weight (HMW) and the low molecular weight (LMW) proteins, which are cleaved from a common precursor, SlpA [234]. I extracted the *C. difficile* S-layer and other surface-associated proteins (SLPs) from multiple *C. difficile* strains using an acid-glycine method [234]. Figure 12 B shows the S-layer and other surface-associated proteins from the hypervirulent strains BI-6, BI-8 and BI-17 and also the toxigenic (but non-hypervirulent) K14 and 630 strains. I found that hypervirulent strains produced surface layer proteins that were distinct in both size and number from non-hypervirulent strains. The HMW S-layer protein appeared as a doublet of approximately 48kDa, and the LMW S-layer protein appeared as a doublet of approximately 30kDa. Western blotting experiments using anti-HMW SlpA and anti-LMW SlpA antisera showed that the bands in the doublets were indeed SlpA (Figure 12 C). The LMW subunit has been previously noted to vary more in size and antigenicity [233, 234, 252]. To complete this initial *C. difficile* surface protein analysis, several SLP bands were excised and subjected to MALDI mass spectrometry analyses, and showed that all were S-layer or surface-associated proteins of *C. difficile* (Fig. 12 D) which shared greatest similarity to those of the recently sequenced hypervirulent *C. difficile* strain from Quebec QCD-32g58 (Genbank accession #AAML00000000; GI:145694830).
Surface protein extracts block adherence of CD to host cells

*Clostridium difficile* SLP preparations have previously been demonstrated to adhere to host epithelial cells [232]. To determine if there was a link between SlpA and adherence of *C. difficile* strains, I employed two approaches: protein interference and antibody interference. First, I performed anaerobic adherence assays where confluent C2BBE monolayers were pre-incubated with increasing amounts total surface-layer (SLP) preparations (such as those used in Figure 12 B). There was a dose-dependent reduction in *C. difficile* adherence with increasing amounts of total SLP protein (up to 80% adherence inhibition; \( p \leq 0.0001 \)) (Figure 13). This indicated that pre-incubation of C2BBE host cells with the total SLP preparation blocked adherence presumably by competing for a host-cell receptor(s). In control experiments, PBS alone or PBS with 50ug bovine serum albumin did not interfere with *C. difficile* binding to C2BBE cells in a statistically significant manner.
Figure 13. *C. difficile* adherence interference assay using total SLP protein from strain BI-17 to coat C2BBE cells, and then testing adherence of the same strain (BI-17). For comparative purposes, data were converted to percent adjusted adherence, with adherence of the control strain set to 100%. Asterisks indicate significant differences in raw data ($p \leq 0.01$).
**Surface protein interference is not strain specific**

In the hamster model of CDI, colonization with a non-toxigenic strain of CD efficiently prevents colonization by a toxigenic *C. difficile* strain (36), although the mechanism for this effect is not known. In addition to the array on the surface, SlpA is secreted into the extracellular milieu (Figure 14 A). Could the presence of SlpA provide a mechanism for this effect? To test if the adherence interference observed above was strain-specific, I used total SLP preparations in cross-interference of adherence assays. SLPs prepared from the phylogenetically unrelated (by multilocus and microarray analyses) non-toxigenic *C. difficile* strain M3 significantly inhibited BI-17 adherence, and SLPs from *C. difficile* strain BI-17 significantly inhibited M3 adherence (Figs. 14 B and C). The degree of inhibition was almost identical in both sets of assays (≥85%; \( p \leq 0.0001 \)).
Figure 14. Panel A. Supernatants of CD indicate the presence of secreted SlpA Marker indicates kD Lane 1 Strain M3. Lane 2 Strain J9 Lane 3 Strain BI17. Panel B and C: Cross-interference of adherence with non-cognate SLP preparations. B. Strain BI-17 adherence tested on C2BBE host cells pre-adsorbed with total SLPs from either cognate
(BI-17) or non-cognate (M3) strains. C. Strain M3 adherence tested on C2BBE host cells pre-adsorbed with total SLPs from either cognate (M3) or non-cognate (BI-17) strains. All experiments were performed in quadruplicate.
Antibodies to SlpA interfere with adherence

Surface extracts contain other proteins in addition to SlpA. To implicate SlpA specifically, I pre-incubated the *C. difficile* strain 630 bacteria with anti-LMW SlpA or anti-HMW SlpA antisera for one hour before addition to confluent C2BBE monolayers. These antisera were raised against strain 630 SlpA subunits specifically. The presence of both antibodies significantly reduced adherence by approximately 50% (p<0.02), indicating that SlpA was indeed involved in *C. difficile* adherence (Figure 15, \( p \leq 0.0001 \)). The presence of an irrelevant antibody to an *E. coli* protein, anti-TraG, did not significantly reduce adherence. While the antisera were not raised against the HV strain BI17, other assays indicated that both antibodies also significantly reduced the adherence of this strain (not shown).
Figure 15. Adherence interference assay using anti-SlpA antisera. *C. difficile* strain BI-17 incubated with a 1:1000 dilution of non-SLP or anti-SlpA antisera prior to exposure to C2BBE host cells. Adherence of strain 630 set to 100%. Means and standard errors of three replicates are shown. Anti-TraG antiserum acted as a negative control.
Survey of Multiple CD strains reveals wide range of adherence values

Further work detailing the adherence of multiple *C. difficile* strains demonstrated a great range of adherence values, as shown in Figure 16. Genetic groups are not necessarily homogenous in adherence values. Across all strains, the mean adherence value was 6.39%. As a group, the hypervirulent strains exhibited a mean adherence value of 4.63%.
Figure 16. - Adherence of multiple *C. difficile* strains to C2BBE monolayers. Strains in the same genetic group are shown in the same bar color. Percentage adherence with standard error of the mean is depicted; all experiments were performed at least in quadruplicate.
Sequencing of \textit{slpA} from multiple \textit{Clostridium difficile} strains.

Even though further data indicated that HV strains were not markedly more adherent than other CD strains, investigating the contribution of SlpA to adherence remained interesting. I wanted to further characterize how SlpA mediates adherence, and to describe the variation in SlpA that may contribute to the increased adherence of some strains than others.

Many \textit{slpA} genes have been sequenced [233, 252, 258, 259], but observed variations have not been analyzed in the context of the genetic relatedness (both within and between genetic groups) or adherence phenotypes of the strains. I determined the sequence of the \textit{slpA} genes of hypervirulent, toxigenic and non-toxigenic \textit{Clostridium difficile} strains, so I could employ this data when evaluating SlpA function.

I chose 4 strains from the hypervirulent clade: BI1, BI6, BI8 and BI17. BI1 is a historical isolate in the Gerding collection, which predates the hypervirulent epidemics. I also chose toxigenic strains J9 and K14, and two non-toxigenic strains, M3 and T7. As mentioned earlier, microarray data indicate that the HV strains cluster in the same clade, termed HY. Strains J9 and K14 fall into the HA1 clade. Non-toxigenic strains M3 and T7 reside in a third clade, A-B+, which contains toxin variant and non-toxigenic strains.

When comparing the HV strains to other CD strains, there were striking differences in \textit{slpA} sequence identity, depending on the subunit. As shown in Figures 17 and 18, the HMW subunit is more conserved in amino acid sequence, while the LMW subunit is more divergent. This is consistent with previous surveys of \textit{slpA} [252].
LMW subunit shows areas of conservation at the N-terminal signal sequence for export, and in the C-terminal end involved in interaction with the HMW subunit. The HMW subunit is conserved over the whole subunit, consistent with its role as the peptidoglycan anchor.
Figure 17. Low Molecular Weight Subunit Alignment. Alignment of strains 630, K14, M3, J9 and BI17. Yellow indicates complete consensus, Blue indicates high consensus, Green indicates similar residue.
Figure 18. High Molecular Weight Subunit Alignment. Alignment of strains 630, K14, M3, J9 and BI17. Yellow indicates complete consensus, Blue indicates high consensus, Green indicates similar residue.
Figure 19 shows the percentage of amino acid identity between strains, by subunit. Data indicate that the *slpA* sequence was identical (99-100%) at the amino acid level among the four HV strains. Strain J9, despite being in a different clade, showed the highest amount of sequence similarity to HV strains. The non-toxigenic strain M3 showed the lowest amount of sequence identity, consistent with the phylogenetic divergence of non-toxigenic strains.
Figure 19. Sequence Analysis of S1pA subunits of 6 CD strains. A. LMW sequence identity. B. HMW sequence identity. Numbers indicate percent identity at the amino acid level.
SlpA binding to host cells of *C. difficile* SlpA subunits independently

If SlpA mediates adherence, then it is reasonable to assume that one or both of the subunits should be able to mediate adherence. The conservation of the HMW subunit and the divergence of the LMW subunit may indicate a distinct function of each subunit. Further, is the variability in SlpA sequence responsible for increased adherence seen in some strains? To answer these questions, I assessed the contribution of recombinant SlpA subunits to adherence using protein interference assays. I chose strain 630 because it is highly adherent, (~10 %, see Figure 16), and strain K14 because it has a low level of adherence (~2%, see Figure 16). If SlpA plays a role in adherence, then the subunits of a highly adherent strain should affect adherence more than those of a less adherent strain.

To produce recombinant subunits, the portions of *slpA* corresponding to the LMW subunit and the HMW subunit from CD strain 630, and the LMW subunit from strain K14 were each cloned individually and expressed in *E. coli* and purified (Figure 20). The HMW subunit from strain K14 was purified from *C. difficile* surface extracts. In the protein interference assays, confluent Caco-2 monolayers were pre-incubated for 20 minutes with increasing amounts of recombinant SlpA subunits. Exponential phase strain 630 *C. difficile* were then added to the monolayers, and adherence was allowed to proceed for 20 more minutes.
Figure 20. A. SlpA crystal structure, adapted from Fagan et al 2009. B. Recombinant proteins from strains 630 and K14 synthesized and characterized in this work.
The results obtained with the recombinant subunits correlate well with the antibody interference studies, which indicate that both subunits play a role in adherence. The presence of recombinant protein decreases adherence in a dose-dependent manner, but the degree of decrease differs for the subunits from different strains, as shown in Figure 21. From the range of values, I calculated the 50% Inhibitory Concentration (IC50) of each subunit, as shown in Table 4. Trends in the data suggest that the LMW subunit has a greater impact on adherence. Strain 630 is highly adherent, and the recombinant 630 LMW showed the lowest inhibitory concentration. Strain K14 exhibits a low level of adherence, and the IC50 values of K14 recombinant subunits reflect this. However, because of the limited number of observations, especially for the K14 subunits, the confidence intervals of the curves for each subunit overlapped, indicating that they are not significantly different.

Table 5. Inhibitory concentrations producing a 50% decline in adherence.

<table>
<thead>
<tr>
<th>Strain and Subunit</th>
<th>630 LMW</th>
<th>630 HMW</th>
<th>K14 LMW</th>
<th>K14 HMW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non Linear IC50</td>
<td>1.1 uM</td>
<td>3.7 uM</td>
<td>2.2 uM</td>
<td>4.4 uM</td>
</tr>
</tbody>
</table>
Figure 21. Adherence of strain 630 in the presence of recombinant LMW or HMW SlpA subunits from strains 630 or K14. Each dot indicates the mean of one triplicate test. Conditions with buffer alone were set to 100% adherence. Curves indicate non-linear least squares fit.
Summary

*C. difficile* strains vary significantly in their ability to adhere to host cells. The surface protein SlpA varies significantly between strains, and the antibody and protein interference experiments suggest that SlpA is involved in *C. difficile* adherence. Antibody interference and recombinant protein interference studies indicated that both subunits are involved.
Toxin Production and Gene Expression in Hypervirulent \textit{C. difficile}

\textbf{Introduction}

As mentioned earlier, toxins are produced in stationary phase in response to nutrient limitation [151]. The increased morbidity and mortality associated with HV CDI [5] first suggested that toxin production might be increased. This hypothesis was first tested in vitro by Warny et al, using ELISA to determine the levels of toxin production in a group of HV and non-HV strains [194]. They found that the median values of toxins A and B were 16 and 23 fold higher, respectively, than the median (not mean) for a group of non-HV strains. They measured toxin production at 0, 24 and 48 hours, and concluded that the high amounts of toxin accumulated by 24 hours were produced in exponential phase, though their growth curves do not define the inflection points of a typical sigmoid growth curve.

Taken at face value, this interesting result had the potential to be useful to our characterization of HV CD. If toxins were produced in exponential phase, this could serve as a positive control in investigating the overall proteome differences in HV strains by Stable Isotope Labeling of Amino Acids in Culture (SILAC), an approach not detailed in this work. However, I was unable to replicate the production of toxin during exponential phase by HV strains. Using a commercial ELISA for both toxins, I found no toxin present in supernatant from an HV strain. This discrepancy led me to want to systematically define the production of toxin by HV strains over the growth cycle, to determine if toxin production occurred earlier in the growth cycle, and determine if levels are increased when mean rather than median values are used for analysis.
To answer these questions, I performed growth curves and tested samples of supernatant collected over the whole growth cycle and subjected them to ELISA for toxins A and B. If toxin production were indeed dysregulated in HV strains, then I would expect to see toxin production during a well-defined exponential phase. I compared four HV strains to four non-HV strains. The non-HV strains included strain 630, the sequenced strain and a known low toxin producer, and strain VPI 10463, a known high-toxin producer, and strains, J9 and K14, which have been associated with previous outbreaks, [95, 189].

Hypervirulent CD strains do not secrete toxins during exponential growth

HV strains showed no statistically significant growth defects or advantages over non-HV strains (Figure 22 A). Some strains exhibited a steeper decline in optical density readings in late stationary phase, but this finding was not consistent or specific to HV or non-HV strains.

Evaluation of toxin production by ELISA indicated that no toxin was detectable during the exponential phase of growth for all strains tested (Figure 22 B). The sensitivity of the ELISA is approximately 0.8 ng/ml for Toxin A and 2.5 ng/ml for Toxin B. The highest toxin producer was the non-HV strain VPI 10463 (~2000 ng/mL total toxin at 48 hrs). Toxin began to accumulate in culture supernatant fluids between 12 and 15 hours in VPI 10463 as well as in three of the HV strains. For all strains tested, the majority of toxin accumulated between 24 and 48 hours. When averaged, there were no statistically significant differences between the group of four HV strains and the four non-HV strains. At 24 and 48 hours, strain VPI 10463 exhibited significantly higher toxin production than
all other strains. At 24 hours, BI8 was significantly higher than strains 630, J9 and BI23. Interestingly, while strain BI23 was significantly lower than BI17, BI8 and VPI 10463 at 24 hours, at 48 hours, it was significantly higher than all strains, other than VPI 10463.
Figure 22. A. Growth curves of the eight strains used in this study. Absorbance at 600nm wavelength measured over 72 hours of growth. B. Total toxin levels over time. Toxin levels indicate combined TcdA and TcdB levels by ELISA. Exponential Phase samples were taken when each individual culture reached an OD600 of 0.5. Means and standard error from three biological replicates are shown. The sensitivity of the ELISA is approximately 0.8ng/ml for TcdA and 2.5ng/ml for TcdB.
PaLoc Gene Expression Over the Growth Cycle

The previous data address the production of secreted Toxin A and Toxin B proteins in combination. If both were undetectable by ELISA, then one would assume that gene expression is similarly repressed, and that high toxin production in stationary phase would be correlated with high gene expression. However, each toxin may be produced in differing amounts, and protein synthesis and gene expression are not perfectly correlated. Also, I wanted to determine the expression of tcdR and tcdC over the growth cycle. Given my preliminary data, I hypothesized that the lack of TcdC, the negative regulator, would not be sufficient to account for toxin production during exponential growth, because transcription of the toxin genes requires TcdR and is also dependent on other stationary phase regulatory factors such as CodY [122, 125].

To test my hypothesis, I performed transcriptional analyses of 4 PaLoc genes over the growth cycle. I examined the transcription of the two toxin genes, tcdA and tcdB, and the regulatory factors tcdC and tcdR, in the 8 strains of CD, during exponential versus stationary phase, to distinguish between two possible effects of TcdC truncation on the production of Toxins A and B: dysregulation in exponential phase, and increased levels in stationary phase.

HV strains exhibit high tcdA and tcdB expression

Transcriptional analyses using quantitative real-time PCR were performed on all eight strains tested above for both tcdA and tcdB, as well as for the regulatory factor-encoding genes tcdC and tcdR, at both exponential and early stationary growth phases. Based on the ELISA data, I expected that the expression of tcdA and tcdB in all strains
would be low/undetectable in the exponential phase of growth, and high in stationary phase, and this was indeed the case. The levels of the *rpoA* reference gene expression were comparable across all strains in the same growth phase. Overall levels of *rpoA* transcription were higher in exponential phase than in stationary phase, as would be expected under rapid growth conditions.

For *tcdA*, and consistent with the expectation described above, low levels of expression were observed during exponential growth (Figure 23 A). Levels of detectable *tcdA* transcripts were, on average, 100-fold lower than those of *rpoA* (the reference gene) and were not significantly different between HV and non-HV strains as a group. If *tcdA* expression were dysregulated in exponential phase, one would expect to see much higher expression in the HV strains. In this sensitive assay, I also determined that strain 630 was significantly lower in *tcdA* gene expression than strains VPI 10463, K14, BI17, and BI23 (*p*≤0.004), while as noted earlier, no toxin protein was detectable during exponential phase.

In contrast, high relative expression of *tcdA* was seen in stationary phase samples (Figure 23 B). Strain VPI 10463 showed significantly higher *tcdA* expression compared to all other strains (*p*≤0.05), which was consistent with the ELISA results for toxin production. The expression of *tcdA* of all HV strains was, on average, not significantly different from all non-HV strains together.
Figure 23. A. Transcription of \textit{tcdA} during exponential phase, relative to transcription of the housekeeping gene, \textit{rpoA}. Means and standard errors of 3 biological replicates are shown. B. Transcription of \textit{tcdA} during stationary phase, relative to transcription of the housekeeping gene, \textit{rpoA}. Means and standard errors of 3 biological replicates are shown.
A similar pattern of expression was observed for the \textit{tcdB} gene in exponential phase (Figure 24 A). There was a small but statistically significant ($p \leq 0.02$) 3-fold higher expression in the HV strains versus the non-HV strains in exponential phase. During stationary phase, strain VPI 10463 exhibited the highest \textit{tcdB} expression level of all strains tested (Figure 24 B), consistent with toxin levels as determined by ELISA.
Figure 24. A. Transcription of *tcdB* during exponential phase, relative to transcription of the housekeeping gene, *rpoA*. Means and standard errors of 3 biological replicates are shown. B. Transcription of *tcdB* during stationary phase, relative to transcription of the housekeeping gene, *rpoA*. Means and standard errors of 3 biological replicates are shown.
TcdR is the sigma factor that directs toxin gene expression. If toxin production were indeed dysregulated in HV strains, and occurred during exponential growth, then the expression of \textit{tcdR} should be expected to be dysregulated as well, since this protein is the positive activator of toxin gene expression. However, consistent with the above results, relative expression levels of \textit{tcdR} were low, and not significantly different between HV and non-HV strains during exponential phase (Figure 26 A).

However, during stationary phase, I observed that \textit{tcdR} expression was high in all strains, and correlated with high toxin levels. High toxin producing strains tended to have higher levels of transcript of this sigma factor than low toxin producers (Figure 26 B). Expression of \textit{tcdR} in the HV strains was, on average, 1.2 fold higher than in all non-HV strains.
Figure 25. A. Transcription of *tcdR* during exponential phase, relative to transcription of the housekeeping gene, *rpoA*. Means and standard errors of 3 biological replicates are shown. B. Transcription of *tcdR* during stationary phase, relative to transcription of the housekeeping gene, *rpoA*. Means and standard errors of 3 biological replicates are shown.
Expression of tcdC occurred in exponential phase, and contrary to previous data obtained with strain VPI 10463 [119], this expression increased, rather than decreased, in all strains, including VPI 10463 during stationary phase (Figures 27 A and 27 B). Although there was evidence of tcdC transcription in the HV strains, it should be noted that this expression does not result in a functional protein; thus the relevance of tcdC expression in these strains is unclear. Expression of tcdC during exponential phase was higher in strain K14 than in the other strains (p ≤ 0.001).
Figure 26 A. Transcription of *tcdC* during exponential phase, relative to transcription of the housekeeping gene, *rpoA*. Means and standard errors of 3 biological replicates are shown. B. Transcription of *tcdC* during stationary phase, relative to transcription of the housekeeping gene, *rpoA*. Means and standard errors of 3 biological replicates are shown.
Summary

Consistent with my hypothesis about the role of TcdC, toxins were undetectable during exponential growth, and $tcdA$, $tcdB$ and $tcdR$ genes showed low, basal levels of transcription in both hypervirulent and non-hypervirulent strains. During the stationary phase of growth, hypervirulent strains produced robust but not significantly increased amounts of toxin as compared to the average of four non-hypervirulent strains Total toxin amounts were directly proportional to $tcdA$, $tcdB$ and $tcdR$ gene expression. Contrary to previous work, I found that $tcdC$ expression did not diminish in stationary phase.
Sporulation

Introduction

*Clostridium difficile* has several strategies to adapt to nutrient limitation or non-optimal environmental conditions. In a host environment, the production of toxins damages host cells, releasing nutrients for uptake. *C. difficile* also sporulates in response to starvation or the presence of oxygen. Both these stationary phase responses are tightly regulated by factors that assess environmental conditions. One of these is the response regulator CodY, which represses sporulation and toxin production during exponential growth [125]. In *Bacillus subtilis*, a complex network of two-component systems regulates stationary phase phenomena such as competence, antibiotic production, and sporulation. These phenomena can be mutually exclusive, as the accumulating of nutritional signals activates successive regulatory switches, culminating in the commitment to sporulation [260].

In CD, the relationship between toxin production and sporulation is the subject of debate. They may be somewhat mutually exclusive, as an inverse relationship was noted between toxin yield and spore counts in a survey of toxigenic strains [203]. For example, if bacteria sporulate early, there would be less time in which to produce toxin. Conversely, an extended time of toxin production in limiting conditions might leave the cell vulnerable to death before sporulation is complete.

The HV strains raise some interesting questions concerning this potential inverse relationship between toxin production and sporulation. If HV strains are high toxin
producers, then does the previous relationship between toxin and sporulation hold? As mentioned earlier, HV strains have spread widely and quickly. Does an increased ability to sporulate confer an adaptive advantage that promotes epidemics by HV strains? To answer these questions, I determined the course of sporulation over the growth cycle by selective treatment and plating of samples taken over the growth cycle.

**HV strains sporulated more efficiently than non-HV strains**

Since the *C. difficile* spore is the etiologic, transmissible agent, any alteration in sporulation efficiency can impact the degree of environmental dissemination. I determined the accumulation of spores over the growth cycle, and found that in HV strains, spores appeared earlier, and more spores accumulated per total volume of culture than non-HV strains. HV strain spore accumulation commenced at 28 hours in BHI broth, prior to any non-HV strain (Figure 27 A). At 48 hours, HV strains BI6 and BI8 had accumulated significantly more spores than all other strains. HV strains BI17 and BI23 had accumulated significantly more spores than strains 630, VPI 10463, J9 and K14. When calculated as the number of spores formed per total number of vegetative cells at the end of exponential growth (efficiency), HV strains also had the greatest efficiency (up to 3.55%), as compared with non-HV strains (≤0.66%; not shown).

Sporulation efficiency evaluated using microscopy also indicated similar trends (Figure 27 B). The highest numbers of spores were seen in HV strains BI6 and BI17, which were not significantly different from each other, but which accumulated significantly more spores than the other six strains (p<0.05).
Figure 27. A. Spores as heat-resistant colony forming units over 48 hours of growth. Means and standard errors of three biological replicates are shown. B. Mean spores (defined as all refractile bodies, including sporangia) counted per field by microscopy.
CHAPTER FOUR

DISCUSSION

Commensals and Pathogens

Why do pathogens evolve? Of course, they evolve because the particular environment requires them to do so to survive, or a new niche provides an opportunity for a particular variation to expand. If pathogenesis is an adaptation, what drives how virulent a given pathogen is? If there is a range of pathogenicity, it has been hypothesized that it is ecologically most optimal for a pathogen to be moderately virulent, and extremes on the continuum are not as successful. Using viruses as an example, it has been argued that the Ebola virus is not well-adapted to humans, because it is so virulent it kills its host too quickly to spread effectively [261]. When a pathogen makes the jump to a new host, there is often a high-attrition evolutionary war in which both host and pathogen adapt, with the pathogen eventually becoming less virulent. One example is the emergence of the recent highly virulent human pathogen HIV from the more benign simian SIV. Recent estimates date the acquisition of SIV by non-human primates at 32,000 years ago, which is more consistent with the evolution of the virus to a less virulent form [262].

Similarly, among bacteria, there may be situations in which it may be more adaptive to be less pathogenic. For example, Vibrio cholerae can exist as a human pathogen or as a
non-pathogenic free-living freshwater organism. Disease-causing bacteria promote their expansion by transferring pathogenic genes to previously non-pathogenic strains. However, improved water sanitation practices can select for the survival of non-pathogenic strains [263].

In contrast to pathogens, gut commensal bacteria lead a domesticated life. In return for abundant nutrients, innate immune factors and abundant secretory IgA contain them in the gut compartment. However, commensals may have pathogenic capacity if they escape and grow elsewhere in the body. For example, *Bacteroides fragilis* is one of the more commonly identified members of the gut biota, but it can cause infections [264]. True pathogens such as *Shigella* are not normal parts of the microbiota. They can evade the physical and humoral barriers to invade tissues and cause systemic infections. However, asymptomatic carriage can follow infection, which illustrates the fact that a robust immune response can turn a pathogen into a temporary commensal. Of course, asymptomatic carriage can act to a pathogen’s advantage by promoting spread to new hosts.

The pathogenicity of *C. difficile* prompts similar questions. Where does CD fall on this continuum of commensal to opportunistic pathogen? Humans can carry CD asymptptomatically. In a survey of hospital patients, being asymptptomatically colonized with CD actually decreased the likelihood that a patient would exhibit symptomatic disease [265]. These asymptomatic carriers most likely have a robust immune response to Toxin A [266].
Further, why do some strains contain the PaLoc while others do not? The *C. difficile* genome has many genes acquired through horizontal gene transfer [182], and it appears the PaLoc has been acquired more than once, and is not the result of a single founding event [267]. Also, it appears that the PaLoc has been lost in non-toxigenic CD in multiple independent events, rather than that it was never acquired [200]. Non-toxigenic strains are certainly successful colonizers of humans. One non-toxigenic strain, M3 was the most commonly found non-toxigenic CD strain in a survey of one hospital in MN [221]. Strain M3 also sporulates well (Susan Sambol, personal communication), and it can out-compete other toxigenic strains in the animal model [221]. It appears that pathogenicity is not the only successful path.

**Adaptation of CD to Animal Hosts**

*C difficile* is well adapted to the animal gut. As an obligate anaerobe, CD limits its germination to environments which not only contain nutrients required for growth, but specific indicators of the human host, namely bile salts and a lack of oxygen. The presence of these features initiates germination and vegetative growth, but not necessarily pathogenesis.

In many viruses, growth and virulence often occur concomitantly. In contrast, in bacterial pathogens the signal to express virulence factors does not necessarily come from the engagement of a surface receptor. Instead the triggers for the expression of virulence factors come from metabolic cues rather than contact. For example, *Listeria monocytogenes* differentiates the human host environment versus the saprophytic environment, and the activation of a master regulator of virulence factors, by the
availability of carbon sources characteristic of animal cells versus plant cells [268]. The mutualist/pathogen *Photorhabdus luminescens* also switches lifestyles based on cues from growth phase and nutrient availability, utilizing a metabolic switch in the TCA cycle [269]. Bacterial virulence factors are often not expressed until nutrients are limited. In CD, toxins are induced during starvation, which is consistent with the idea of toxin production as a nutrient scavenging strategy.

The mechanism of the action of Toxin B provides evidence of an intimate adaptation to animal hosts. First, the acidification of the host endosome allows Toxin B to change conformation and insert into the membrane [144]. Second, the presence of a host co-factor, inositol hexakisphosphate, induces autocatalytic cleavage and allows the enzymatic domain to enter the cytoplasm and glucosylate its targets [149]. This feature is unique among bacterial toxins, and suggests that a direct animal interaction has driven the adaptation of toxin production.

We can only speculate about the role of humans as a CD host before the antibiotic era. Colonization may have been limited to infants or those with gut related injuries. However, it is undeniable that the invention of antibiotics has allowed CD to become successful professional pathogen. As mentioned earlier, any antibiotic that disrupts the gut microbiota can allow CD to colonize, but antibiotic resistance offers a distinct advantage. CD is also well-adapted to the pharmaceutical landscape as it has acquired multiple types of antibiotic resistance genes [32, 182, 270-272]. This is true even for non-toxigenic strains [219].
Whether CD has adapted to particular species is under investigation. CD affects humans, pigs, cows and horses, with varying pathogenicity [273, 274]. In pigs, CDI is not fatal, but does cause wasting, in which piglets do not gain weight. Most, but not all, non-human animal derived CD isolates cluster into a distinct phylogenetic group by microarray [200]. A further microarray study indicated that different sets of genes were conserved in CD strains from specific host species, which suggests some species adaptation [275]. Current studies in the Vedantam lab are testing whether human, porcine or bovine strains adhere better to cells in culture from the host organism in which they were isolated.

Adherence of the bacterium to host cells is another feature for which there may be a continuum of adaptation. We might assume that it is better to be more adherent, but this must be tested empirically. Do commensals adhere more or less than pathogens? Certainly, pathogens that invade the body from the gut, or replicate intracellularly, must make intimate contact with the host cells. What kind of adherence is adaptive for CD, as a non-invasive toxin producer? When I began my work on hypervirulent CD, these were the questions I wanted to investigate.

**Adherence Varies Between Strains and Is Multifactorial**

My initial comparisons were between HV strains and other toxigenic CD strains that had caused epidemics, to compare the previously most successful CD strains to HV strains. While these results indicated that HV strains might be more adherent than other strains, my survey of a larger group of strains indicated a wide variety of adherence values, with the HV strains averaging slightly below the mean of all strains tested. The
HV strains were not homologous as a group in their level of adherence. Other genetic groups, such as the J and K group, in which I tested two strains, also indicated variability.

Considering that it is typical for bacteria to have a whole repertoire of adhesins, and that the expression of these adhesins can be regulated, it is not surprising that even closely related strains of CD can vary so widely in their adherence. Adherence is a multifactorial phenomenon, and does not appear to be even broadly correlated, either positively or negatively, with levels of toxin production (at least within the limits of my in vitro assay). While the non-toxigenic strains such as M3, M23, and T7 were highly adherent, so were several toxigenic strains.

**Contribution of SlpA to Adherence and Colonization**

My antibody and protein interference assays indicated that the surface layer protein SlpA is involved in adherence. Most adhesins are typically surface associated molecules that have regulated expression. For example, flagella are employed for surface contact and then shed when a sessile life style is achieved, such as in a biofilm [276]. One example from *C. difficile* is the adhesin Cwp66, which is down regulated during stationary phase, but up-regulated in the presence of certain stressors, such as antibiotics [277, 278].

SlpA is not a typical adhesin, in that the lattice of the S-layer covers the whole cell. As the cell grows, new layers grow and overlap the old in sheets. So increased expression of *slpA* would not provide additional binding epitopes on the surface; the cell is already covered with them. So how does SlpA serve as an adhesin? Its stability and ubiquity may be advantageous under certain conditions. For example, if other adhesins
are down regulated during stationary phase, SlpA is always present to serve as an alternate, high avidity adhesin.

Microarray studies indicate that slpA expression is increased during stationary phase [278]. My cross-interference assays confirmed that adherence inhibition occurred when using SLP preparations from non-cognate \textit{C. difficile} strains. I also found that SlpA was secreted into the extracellular milieu. Taken together, these findings hint at another function of SlpA. As mentioned earlier, in the animal model of CDI, the colonization of one CD strain prevents the subsequent colonization of another. The mechanism for this exclusion is not known. However it is likely to be different from the mechanism by which the normal gut microbiota exclude CD, since it would be maladaptive to inhibit the outgrowth of cognate spores. If it were merely competition for nutrients, one would expect to see a more graded outcome of competition, rather than all-or-none exclusion.

My results suggest that secreted surface proteins could prevent \textit{C. difficile} colonization by competitively excluding bacterial adherence. This idea is supported by studies of other mucosally associated bacteria. For example, surface layer extracts from \textit{Lactobacillus helveticus} inhibit the adherence of enterhemorrhagic \textit{E. coli} O157:H7 to epithelial cells (19).

\textbf{SlpA Subunit Sequence and Function}

I sequenced slpA from eight strains and compared them to the published sequence of strain 630. Consistent with previous studies [233, 252], my sequencing results indicate that the LMW subunit was more variable, and the HMW subunit was more conserved.
Because of the small sample size, the percentages of identity do not dovetail with the four clade grouping of CD described in the introduction, except for the HV strains. The four HV strains were identical. Members of the other clades, HA1 and A-B+, were no more similar to each other than to members of the other clade, which reflects not only my small sample size, but also the diversity of these large clades.

SlpA, especially the LMW subunit, in addition to being highly variable, is also antigenic. This antigenicity has led some to suggest that SlpA could be included in multi-component vaccine candidates for the prevention of CDI [67, 279, 280]. However, SlpA may actually be too variable to provide effective immunity to a broad range of strains. This high variability and antigenicity suggest that SlpA may be under positive selection for evasion of the host immune response. The hypothesis of positive selection imposed by host immune system can be tested examining the ratio of non-synonymous changes (dN) to synonymous changes (dS) [281]. Using this analysis, other slpA sequencing projects have noted that the HMW peptide is more divergent than the downstream gene, a SecA homologue, consistent with positive selection [252]. The LMW peptide sequences are so variable that divergence is difficult to assess using a limited number of strains, so similar analyses were not been performed in that work [252]. A more recent analysis, using whole genome sequencing of thirty CD strains, found evidence of 12 positively selected genes, shown in Table # [267]. While the authors admit this may be an underestimate, neither SlpA nor any of its surface associated neighbors were in the positively selected group [267]. The two-subunit structure of SlpA, with its variable and conserved regions, may render such analysis difficult.
Table 6. Positively selected core genes in *C. difficile*. Adapted from He et al, PNAS 2010.

Gene name refers to systematic identifiers in strain 630.

<table>
<thead>
<tr>
<th>Name</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD0195</td>
<td>Putative membrane protein</td>
</tr>
<tr>
<td>CD0707</td>
<td>Putative signaling protein</td>
</tr>
<tr>
<td>CD1068</td>
<td>Putative polysaccharide biosynthesis/sporulation protein</td>
</tr>
<tr>
<td>CD1755</td>
<td>Putative ABC transporter, permease protein</td>
</tr>
<tr>
<td>CD1989</td>
<td>Putative membrane protein</td>
</tr>
<tr>
<td>CD2022</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>CD2316</td>
<td>Two-component response regulator</td>
</tr>
<tr>
<td>CD2454</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>CD2468</td>
<td>Putative exported protein</td>
</tr>
<tr>
<td>CD3094</td>
<td>Putative sigma-54-dependent transcriptional regulator</td>
</tr>
<tr>
<td>CD3248</td>
<td>Putative polysaccharide deacetylase</td>
</tr>
<tr>
<td>CD3558</td>
<td>BirA bifunctional protein</td>
</tr>
</tbody>
</table>

The LMW subunits of 630 and K14 were only 20% identical, while the HMW subunits were nearly 80% identical. While different, my recombinant protein interference assays indicate that the LMW subunits from both strains impacted host binding. How is this possible? The divergence of the LMW subunit may have multiple constraints: changing enough to evade the immune system, while maintaining its function in host binding. This may be why one explanation of why it is not positively selected; instead, SlpA may be subject to balancing selection.
There are multiple described adhesins in CD, and some are variable, some are conserved. Cwp66 has also been shown to vary at the amino acid level between strains [231], although it is not known how this variation might contribute to colonization. In contrast, one other adhesin, the flagellar cap FliD, has a highly conserved sequence, which is notable in a highly immunogenic structure. Thus it has been argued that this indicates stabilizing selection (i.e., selection against any directional changes) for an adhesive function [282].

My recombinant protein experiments suggest that the HMW subunit may not contribute not as much to adherence, although further work will be required to confirm any significant differences. This is in contrast to the data of Calabi et al, who found that the HMW bound to gastrointestinal tissues, while the LMW did not [232]. That and other work had supported the idea of the HMW subunit as the conserved member of the complex, and the one responsible for mediating adherence. However, my data are more consistent with the crystal structure model of the SlpA molecule, which shows the LMW subunit more surface exposed, and the HMW subunit more closely associated with the peptidoglycan.

Sequence variation may not only affect affinity of SlpA for its target, it may also change the nature of the target entirely. In a study using the hamster model, Goulding et al found that strains 630 and B1 (a highly toxigenic but non-HV strain) differed in localization of adherence in the hamster gut. Strain 630 was found more frequently in the crypts, while strain B1 adhered more to the mucosal surface.
Considering the multifactorial possibilities, it seems unlikely that the variation in SlpA would be responsible for a significant proportion of the variance in adherence. But I wanted to test whether highly adherent CD strains had highly adherent SlpA. My recombinant protein interference assays showed trends suggesting that both subunits of strain 630 had lower IC50 values than both subunits of strain K14. However, strong conclusions cannot be drawn from this limited data set. Further work will be required to determine if strains which are highly adherent have SlpA molecules, both the LMW and the HMW subunits, which are more adherent.

Summary

Taken together, the results from the studies presented here indicate that that some C. difficile strains have increased adherence to human intestinal epithelial cells that is mediated in part by surface-layer protein A (SlpA). The LMW subunit, despite high variability, may be more involved than the HMW subunit in mediating adherence of SlpA to host cells.

Pathoadaptation of HV strains

Our hypothesis at the beginning of this work was that the HV phenotype was due to more than a simple disruption in toxin regulation. The wealth of genomic data that has accumulated in the past several years supports the idea of the micro-evolution of HV strains.

A change in hosts or a shift in virulence usually requires a pathogen to acquire new genes, either by horizontal gene transfer or by mutation. Surveys of HV strains indicate a gain of 234 additional genes as compared to strain 630 [283]. These genes not
only include typical horizontally transmitted genes such as antibiotic resistance and phage elements, but also include new regulators and a two-component system. One of the features of pathoadaptation is the development of new regulatory controls for new or pathoadapted genes.

Another feature of pathoadaptation is the loss of genes that may interfere with new function. These are termed “antivirulence genes” [284]. HV CD has lost 13 genes [200](Table 7), which include a membrane associated gene, and one surface anchored gene, but again, do not include genes from the cluster in which \textit{slpA} resides. \textit{slpA} itself could not be lost, as it is essential [Julian Rood, personal communication]. The relationship between genes lost and virulence may not be obvious. To demonstrate that a particular lost gene is an antivirulence gene, experiments would have to be conducted to re-introduce the gene and then demonstrate a loss in virulence. These genes would be interesting to investigate in the future, especially in the animal model of CDI.

<table>
<thead>
<tr>
<th>Loci</th>
<th>Deletion(s)/divergent genes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD0630-CD0719</td>
<td>Methenyltetrahydrofolate cyclohydrolase</td>
</tr>
<tr>
<td>CD0630-CD0720</td>
<td>Putative FolD bifunctional protein (includes methylenetetrahydrofolate dehydrogenase and methenyltetrahydrofolate cyclohydrolase)</td>
</tr>
</tbody>
</table>

TABLE 7. Gene deletions specific to the Hypervirulent (HY) clade. Adapted from Stabler 06.
Toxin production in HV strains of *C. difficile*

In the past decade, there has been great interest in elucidating the molecular mechanisms contributing to the hypervirulent phenotype, particularly those involved in toxin production. My data indicate that in *C. difficile*, the four HV strains tested do not
make more toxin than all non-HV *C. difficile* strains tested, and my data also clarify the timing of the onset of toxin production. My evaluation of sporulation found a significant difference in the production of infectious particles by HV strains, the magnitude of which exceeds any difference in the production of toxins.

In this work, I have shown that HV *C. difficile* strains do not make the toxins TcdA or TcdB during exponential growth. However, HV strains do produce toxin amounts comparable to, or exceeding, those of other strains during stationary phase. Even this small sample of non-HV strains exhibited a wide range of toxin production. This variance illustrates the importance of the choice of comparator strains when making generalizations about the HV phenotype. Previous studies [194] may have underestimated the significance of these variances in toxin production, leading to misattributions about the potential mechanisms of hypervirulence. If there have been other high toxin producers, but that have not become epidemic, then there must be other hypervirulence mechanisms. For example, Strain VPI10463, or other isolates from its genetic REA group (Z) are not often found in surveys of patient isolates (Dale Gerding, unpublished data).

I conducted my experiments in BHI medium, which contains glucose, a known repressor of toxin gene expression. The kinetics of the onset of toxin gene production might vary in different media. For example, when CD is grown in media which lacks glucose, toxin production does appear earlier [128]. Unfortunately, direct comparison with the medium used in the study by Warny et al is not possible because they used a proprietary medium. However, the point is not the exact time of onset, but the difference
between HV and non-HV strains relative to the $tcdC$ deletion. As discussed further in later sections, this difference should not affect the HV strains differently, because the nutritional regulators of toxin production, such as CodY do not differ between HV and non HV strains.

The negative regulator TcdC has been at the center of the debate about toxin production in HV strains. A previous study of the transcription of PaLoc genes in strain VPI 10463 indicated that $tcdC$ was highly expressed during exponential phase, and that this expression diminished as growth slowed in stationary phase [119]. These data, along with those indicating that TcdC prevents TcdR from complexing with RNA polymerase, led to the model in which TcdC acts as a negative regulator of toxin gene expression in exponential phase [124]. Most HV clade $C. difficile$ isolates have a frame-shift mutation in $tcdC$ which would result in a truncated 65 amino acid protein [208]. Warny et al hypothesized that the lack of a functional TcdC results in transcription of the toxin genes and production of toxin protein during exponential growth [194]. However, my results clearly demonstrate that there is no TcdA/B protein detectable in exponential phase, and that $tcdA$ is no more highly expressed in HV strains than non-HV strains in exponential phase, and $tcdB$ expression is only slightly higher in HV strains.

These data are consistent with the absolute requirement of TcdR for toxin gene expression, and the strong repression of $tcdR$ (observed by us and others) in exponential phase [119, 122, 125]. Thus, the absence of TcdC is likely not sufficient to permit toxin production during exponential growth, and underscores the critical requirement of TcdR for TcdA/B expression.
The prevailing model of TcdC as a negative regulator of toxin gene expression describes TcdC protein diminishing (by an unknown regulatory mechanism) in stationary phase. However, I observed relatively high transcription of \textit{tcdC} in stationary phase. This finding is consistent with recently published data. In particular, while examining the role of CodY (a global regulator of gene expression in gram positive bacteria) in toxin gene repression, Dineen and colleagues found that in strain 630, \textit{tcdC} was expressed during both exponential and stationary phase [125]. Karlsson \textit{et al} also demonstrated that \textit{tcdC} expression increases in stationary phase, and that this expression was suppressed by nutrient supplementation similar to that seen for the toxin genes [131].

I thus hypothesize that TcdC, while not required for exponential phase \textit{tcdA/B} repression, may exert a modulatory effect on toxin production. The lack of this modulator in HV strains may be one reason they show relatively high toxin production in stationary phase. I hypothesize that in exponential phase, small amounts of TcdC may inhibit any rare upstream read-through transcription, which may explain the slight but significant increase in exponential phase \textit{tcdB} expression in the HV strains.

However, there are likely multiple influences on the quantity of toxins produced by various strains. TcdC protein levels may be one factor; the affinity of TcdC for its targets may be another. For example, strains 630 and VPI 10463 both have functional TcdC proteins, but exhibit vastly different toxin production.

My transcriptional studies showed significantly different basal levels of transcription of the four PaLoc genes for different strains. We hypothesized that this may be due to differences in upstream elements. To begin to explain some of these trends in
gene expression, my co-authors in the publication of this work, Michael Mallozzi and Bryan Roxas, examined genetic differences that might produce variations in transcription and translation of toxin genes and proteins, respectively. They performed an alignment of the PaLoc DNA sequences of two closely related strains, 630 and VPI 10463, which exhibited the lowest and highest levels of transcription of toxin and toxin regulating genes, respectively, across all growth phases. The sequence differences in these strains are likely to be responsible for the differences in the amount of toxin produced since they share a common ancestor (as they both fall within the HA1 clade [200]). Briefly, Bryan and Michael aligned the sequences using the ClustalW algorithm [285] in the BioEdit program [286], predicted and analyzed existing promoters [119], predicted transcriptional terminators [287], and annotated ribosomal binding sequences manually based on the *E. coli* consensus sequence. The free energy of RNA secondary structures was estimated using the RNAfold webserver [http://rna.tbi.univie.ac.at/].

This analysis discovered several interesting differences in the sequences of these strains relevant to PaLoc expression. First, the negative regulator (*tcdC*) of VPI 10463 contains a mutation that changes Aspartic Acid 7 of the TcdC protein to Glutamic Acid, which may change its affinity for TcdR or RNA polymerase (potentially lessening its negative-regulating capacity). Second, while the promoter sequences of all of the PaLoc toxin and toxin regulators are identical, the Ribosomal Binding site (Shine Dalgarno sequence) of the *tcdR* gene in VPI 10463 contains a single nucleotide difference as compared to strain 630, which may increase translation of the *tcdR* transcript. Further, two other genetic variations might lead to increased read-through transcription of
the \textit{tcdB} and \textit{tcdR} genes from upstream promoters (Figure 28). For \textit{tcdR}, the predicted transcriptional terminator for \textit{cdu1} (the gene upstream of \textit{tcdR}) lies within the coding sequence of \textit{tcdR}. Therefore, if the \textit{tcdR} RBS mutation in VPI 10463 leads to increased translation, this could also interfere with the transcriptional termination of the \textit{cdu1} gene, leading to read-through transcription of \textit{tcdR} (from the \textit{cdu1} promoter), which would increase the amount of \textit{tcdR} transcript in turn allowing for increased transcription of the toxin genes. For \textit{tcdB}, the secondary structure of the RNA encoded by the \textit{tcdR} and \textit{tcdB} intergenic region has a lower minimum free energy in strain 630 (24.71 kcal/mol) than in VPI10463 (-10.92 kcal/mol). This suggests that there may be more read-through transcription from the \textit{tcdR} promoter in strain VPI 10463 than in strain 630, which may increase the level of the \textit{tcdB} transcript in VPI 10463. Determining whether these differences are responsible for expression levels cannot be determined from strain or sequence comparisons alone, and will require transcriptional studies.
Figure 28. Schematic of the *C. difficile* PaLoc region. Coding regions and predicted regulatory elements are depicted. The *cdul* gene has three predicted Rho-independent terminators (indicated by hairpins). The formation of any given hairpin usually prevents the formation of others downstream. Further, if ter2 forms, transcription from ptdcR would be blocked. “ter”, terminator; “p”, promoter; RBS, ribosome binding site. Not to scale.
Toxin production and sporulation are both responses to nutrient limitation, and the relationship between them is a subject of debate. One survey of toxigenic *C. difficile* strains indicated an inverse relationship between toxin yield and spore counts [203], suggesting that if bacteria sporulate early in stationary phase, there is less time in which to produce toxin. However, an extended time of toxin production under nutrient-limited conditions may result in bacterial death before sporulation is completed. For example, strain VPI 10463, a known high toxin producer, sporulates very poorly [203]. Other studies, however, indicate that mutants of stationary phase regulator Spo0A and its associated sensor kinase are impaired not only in sporulation but also in toxin production [177]. Further research is thus required to elucidate the links between the two systems, and in particular, the role of the *tcdC* mutation in HV strains. One might speculate that TcdC may also disrupt sporulation-associated sigma factors or other ECF sigma factors involved in stationary phase phenomena that may impact survival [116, 288].

The data clearly demonstrate that the four strains in the hypervirulent clade of *C. difficile* that I tested not only sporulate with greater efficiency than other strains, but also produce robust amounts of toxin. These data are consistent with another study of different hypervirulent isolates [213]. HV strains also produce toxin B with different intoxicating potential [202]. While the correlation between *in vitro* toxin production and clinical outcomes is not consistent [203], it is possible that the altered toxin B phenotype may influence disease severity. Enhanced sporulation may
increase the likelihood of disseminating infectious particles into the environment, acting synergistically with toxin to give an adaptive advantage for hypervirulent C. difficile in terms of pathogenesis.

Summary

Most likely, a synergistic confluence of multiple factors has allowed the C. difficile hypervirulent phenotype to emerge. Human factors such as the increased use of antimicrobials in the hospital and community settings as well as inability to control environmental contamination have undoubtedly contributed to CDI outbreaks. Bacterial factors that have been described include genetic alterations resulting in fluoroquinolone resistance, Toxin B enzymatic variation, high toxin production, and increased sporulation. However, my data and that of others indicate that there is much more to discover, about hypervirulent C. difficile and about C. difficile pathogenesis in general.

Future Directions

The variability of the LMW subunit raises interesting questions about the relationship of structure, function and selection. It would be interesting to see if the LMW subunit could evolve in a way that conserves structure and function, but also shifts to avoid the host immune system. Now that there is a crystal structure of the LMW subunit from one strain it would be illuminating to thread other sequences over this structure and assess how well they match, if at all. Extremely diverse sequences can produce very similar structures, one example is the structural similarity of
eukaryotic actin and the bacterial proteins ParM and MreB [289]. However, it is unknown if such adaptations could evolve on a relatively short time scale.

To further investigate the pathoadaptation of HV strains in terms of their interaction with the host, some of the genes that have been lost in the HV clade could be re-introduced, to test if they are indeed anti-virulence genes. These genes could be re-introduced either on a plasmid or on a transposon, and then any differences could be noted in adherence, toxin production, sporulation, or virulence in the animal model.

To further investigate the regulation of toxins, it would be informative to re-introduce a functional copy of tcdC, either on a plasmid or on a transposon. If the re-introduced gene was placed under the control of an inducible promoter, then my hypothesis that TcdC acts as a modulator could be tested. If the hypothesized relationship is correct, then one would expect that toxin transcription and protein production would be inversely correlated with the amount of TcdC.

Using the same approach, one could also test for effects on sporulation. If a relationship between TcdC and sporulation efficiency was found, then an investigation of other potential ECF sigma factor targets in the CD genome could begin.
REFERENCE LIST


100. Kuijper, EJ, J de Weerdt, H Kato, N Kato, AP van Dam, ER van der Vorm, J Weel, C van Rheenen, and J Dankert. Nosocomial outbreak of Clostridium


164. Feltis, BA, SM Wiesner, AS Kim, SL Erlandsen, DL Lyerly, TD Wilkins, and CL Wells. Clostridium difficile toxins A and B can alter epithelial permeability


267. He, M, M Sebaihia, TD Lawley, RA Stabler, LF Dawson, MJ Martin, KE Holt, HM Seth-Smith, MA Quail, R Rance, K Brooks, C Churcher, D Harris, SD


VITA

Michelle Merrigan was born and raised in southern California. She attended the University of San Diego where she earned a Bachelor of Arts, *magna cum laude*, in Psychology in 1990. She went on to earn a Masters Degree in 1995 from Tulane University in Psychology, with an emphasis on Neuroscience and Animal Behavior. After moving to Chicago, she began her studies of *Clostridium difficile* as a research assistant in the laboratory of Dr. Dale Gerding and Dr. Stuart Johnson at Northwestern University and later at Loyola University Chicago.

Michelle began her doctoral studies in the Microbiology and Immunology Department at Loyola University Chicago in 2004. The following year, she joined the laboratory of Dr. Gayatri Vedantam, and continued her research examining the pathogenesis of hypervirulent strains of *C. difficile*. Michelle received the Arthur J. Schmitt dissertation fellowship in 2009.

Michelle has been active in the Chicago chapter of the Association for Women in Science, and is currently the Vice President of Communications. After completing her PhD, Michelle will begin a position as a consultant and medical writer for Boston Strategic Partners.