STRUCTURAL AND FUNCTIONAL ANALYSIS OF

CLOSTRIDIUM DIFFICILE TOXINS A AND B

By

Rory Nelson Pruitt

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Microbiology and Immunology

December, 2011

Nashville, Tennessee

Approved:

Professor D. Borden Lacy

Professor Timothy L. Cover

Professor Eric P. Skaar

Professor John V. Williams

Professor Melanie D. Ohi

To my parents

ACKNOWLEDGEMENTS

I am extremely fortunate to have had the opportunity to do my thesis research in the laboratory of Dr. Borden Lacy. Borden is not only a great scientist, but also a great teacher and mentor. I would like to say a very special thank you to Borden for her enduring patience, support, and enthusiasm throughout my time in graduate school. I am also grateful to Borden for assembling a terrific group of researchers with whom I have had the pleasure of working for the past four years. Dr. Darren Mushrush, Dr. Kelly Gangwer, and Adarsh Godbole were in the lab when I joined and they were very patient and helpful in teaching me basic skills and assisting with experiments. Michael Cover, a summer student in the lab, contributed greatly to this work by obtaining the first crystals of the TcdA cysteine protease domain. Nicole Chumbler cloned and expressed the TcdB used in Chapters III and IV. These and the other members of the lab, John Schmitt, Michelle LaFrance, Desiree Benefield, Stacey Seeback, Dr. Melissa Farrow, and Ryan Craven, have all been helpful in many ways and have made the Lacy lab an enjoyable place to work.

Vanderbilt has been a fantastically collaborative and collegial research environment. In addition to my labmates, I am particularly indebted to Dr. Benjamin Chagot and Dr. Walter Chazin, who assisted with the NMR experiments on the TcdA cysteine protease domain (Chapter II); Melissa Chambers and Dr. Melanie Ohi, close collaborators for the EM analysis of TcdA and TcdB (Chapter III); Dr. David Friedman, who carried out MS experiments to detect modified Rap2A (Chapter IV); and Dr. Benjamin Spiller, who assisted with the X-ray crystallographic data collection and analysis (Chapter II and IV). In addition, Dr. Robert Carnahan, Tracy Cooper, and Graham Black at the Vanderbilt Monoclonal Antibody Core generated the antibodies used for domain localization (Chapter III). My thesis committee, Dr. Timothy Cover, Dr.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A_{280}	Absorbance at 280 nm
aa	Amino acids
BHI	Brain heart infusion broth
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
CDAD	*Clostridium difficile* associated disease
CDTab	*Clostridium difficile* binary toxin
CPD	Cysteine protease domain
DMEM	Dulbecco's modified Eagle's medium
DTT	Dithiothreitol
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
FBS	Fetal bovine serum
FSC	Fourier shell correlation
g-CPD	TcdA CPD plus a small part of the preceding GTD (aa 510-809)
GAP	GTPase activating protein
GDI	Guanosine nucleotide dissociation inhibitor
GEF	Guanine nucleotide exchange factor
GST	Glutathione S-transferase
GT-A	Glucosyltransferase A
GTD	Glucosyltransferase domain
GTPase	Guanosine triphosphatases
HRP	Horse radish peroxidase
HT	Holotoxin

His_{10}	Decahistidine tag
IL	Interleukin
InsP6	Inositol-hexakisphosphate
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilodaltons
LB	Luria broth
LCT	Large clostridial toxins
LR	Long repeat
m/*z*	Mass to charge ratio
MALDI-TOF MS	Matrix-assisted laser desorption/ionization, time-of-flight mass spectrometry
MARTX	Multifunctional autoprocessing repeats in toxin
MLD	Membrane localization domain
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
OD_{600}	Optical density, absorbance at 600 nm
PBS	Phosphate buffered saline
PBST	PBS plus 0.1% Tween-20
pdb	RCSB Protein data bank
PMC	Pseudomembranous colitis
ppm	Parts per million
RBD	Receptor-binding domain
rmsd	Root mean square deviation
SAXS	Small angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
s. d.	Standard deviation

SR   Short repeat

TcdA   *Clostridium difficile* toxin A

TcdB   *Clostridium difficile* toxin B

TcsH   *Clostridium sordellii* hemorrhagic toxin

TcsL   *Clostridium sordellii* lethal toxin

Tcnα   *Clostridium novyi* alpha toxin

TpeL   *Clostridium perfringens* large toxin

UDP-glucose   Uridine diphosphate glucose

UDP-GlcNAc   Uridine diphosphate *N*-acetyl-glucosamine

VcRTx   *Vibrio cholerae* repeats in toxin

CHAPTER I

INTRODUCTION

Clostridium difficile is the leading cause of antibiotic associated diarrhea and is among the most important of the nosocomial pathogens. Advances in the understanding of this bacterium have lagged far behind that of other important human pathogens, however. There is a notable lack of information about the factors involved in colonization and in the case of the two proteins, toxins A and B (TcdA and TcdB), known to be key factors in disease, their roles in pathogenesis and their mechanism of action are still unclear. In this work, I present structural and functional analyses of TcdA and TcdB. These studies begin to fill some of the critical gaps in our understanding of the pathogenic mechanism of these two toxins.

Historical overview

Clostridium difficile was first described in 1935 by Ivan Hall and Elizabeth O'Toole (1). The researchers discovered the bacteria while investigating the bacterial colonization of the intestinal tracts of normal infants during the first ten days following birth. In stools from four out of the ten infants being monitored, they isolated anaerobic, slow-growing bacteria that had not been described previously. The authors named the bacteria *Bacillus difficilis* on account of "the unusual difficulty which was encountered in its isolation and study." It would later be renamed *Clostridium difficile.*

Hall and O'Toole were surprised to discover that *C. difficile* was pathogenic towards animals. When bacterial culture was injected subcutaneously into rabbits and

1

guinea pigs, the animals developed subcutaneous edema, and most died within a few days (1). Similar results were seen when the animals were given sterile filtrates from the bacterial culture. Heating abolished the toxicity of the filtrate. Thus, Hall and O'Toole, in the first report on *C. difficile*, showed that it was pathogenic towards animals and that the disease process involved the secretion of heat labile toxin(s). Nevertheless, the bacteria had been isolated from normal infants, and there was no indication that the presence of the bacteria had any deleterious effects on the newborns. A few years later, it was shown that oral administration of the toxin containing filtrate to rats and guinea pigs did not cause disease, leading to the belief that the toxin(s) could not be taken up within the gastrointestinal tract (2). Thus, for the next four decades *C. difficile* remained a little known bacterium that was considered a part of the normal intestinal flora of infants.

C. difficile's rise from obscurity began in the 1970s when it was found to be the primary cause of pseudomembranous colitis (PMC). PMC is a severe condition characterized by inflammation of the colon with the formation of plaques, or pseudomembranes, along colonic mucosa. These pseudomembranes are typically a few millimeters in diameter and are composed of fibrin, mucin, necrotic epithelial cells, and neutrophils. (3,4). PMC had been described as early as 1893, but the causative agent was undefined for decades (5). Following the development and use of broad-spectrum antibiotics, the number of PMC cases increased dramatically (3). An especially large number of reported cases occurred in the 1970s following introduction of clindamycin (6,7). With this rise in incidence, the search for the cause of PMC intensified, and by the late 1970s there was strong evidence from a number of groups that *C. difficile* was the causative agent (3). Moreover, the toxic components that Hall and O'Toole noted in the filtrate of *C. difficile* were implicated in causing PMC. The toxic components were identified as two secreted proteins, toxin A (TcdA) and toxin B (TcdB).

Clostridium difficile associated disease

Symptoms

Although the pathogenicity of *C. difficile* towards humans was first discovered in relation to its ability to cause PMC, it is now known that the manifestations of *C. difficile* colonization can range from asymptomatic carriage, to mild diarrhea, to life-threatening conditions such as PMC and toxic megacolon. Collectively, the manifestations of disease caused by *C. difficile* are referred to as *C. difficile* associated disease (CDAD).

The most common symptoms associated with infection are mild to moderate diarrhea (4). More severe infections result in severe diarrhea with abdominal pain and swelling, along with systemic symptoms such as fever, dehydration, malaise, anorexia, and nausea. The raised plaques, or pseudomembranes, on the colonic mucosa that are hallmarks of PMC may or may not be present. When they are present, the symptoms are typically more severe (4). The disease can progress to life-threatening conditions where the colonic mucosa and the underlying muscularis propia are severely damaged leading to paralytic ileus and a decrease in the colonic muscle tone (8). As a result, there is a decrease in diarrhea and the colon can become grossly dilated, a condition known as toxic megacolon (4). Although toxic megacolon secondary to *C. difficile* associated colitis is rare (0.3-4% of cases), the mortality rate is staggering, with estimates ranging from 38-80% (8).

Risk factors and transmission

A number of risk factors have been implicated in CDAD including: immunosuppression, poor nutrition, use of gastric acid suppressants, and an interleukin 8 (IL-8) polymorphism (9,10). However, the most important factors are age, hospitalization, and use of broad-spectrum antibiotics (11). Interestingly, infants are

typically not susceptible to disease despite high colonization rates (often over 50%), and CDAD is rare in children and young adults (3,12). Individuals over the age of 65 are at highest risk of disease, being ~5 times more likely to develop CDAD than those in the 45-64 age group and more likely to die as a result of the infection (12,13).

C. difficile can be found in the environment, but it is more commonly encountered in health care institutions, where there are large populations of infected and at risk individuals. A small percentage (~3%) of the general population is colonized by C. difficile, and in most cases transiently (14,15). Hospitalized patients and nursing home residents have much higher colonization rates of 10-25 and 4-20%, respectively (15,16).

C. difficile's prevalence in health care institutions is due in part to the production of hardy spores. Spores are presumably the infectious form of the bacteria, as they are prevalent in the environment and are more resistant to the acidic environment of the stomach (17). Germination is triggered in the duodenum by bile components including glycine and taurocholic acid (17). The mechanisms of sporulation and germination are poorly characterized for C. difficile, but the importance of spores in human disease is obvious. The spore form of the bacteria is metabolically inert and highly resistant to environmental stresses allowing them to persist on surfaces for months or years (18-21). The spores can be destroyed by bleaching or autoclaving, but not by commonly used disinfectants, antimicrobial soaps, and alcohol-based hand products (22,23). Measures to control the spread of C. difficile include the isolation of symptomatic patients, use-dedicated equipment and personal protective clothing, and specialized environmental cleaning (24). These procedures add considerably to health care costs.

C. difficile in an opportunistic pathogen which rarely infects healthy individuals because the natural intestinal flora provides a strong defense against colonization and overgrowth. Individuals with depleted flora as a result of antibiotic treatment, however, are at great risk of developing CDAD. Symptoms typically begin during or shortly after

treatment with broad-spectrum antibiotics including cephalosporins, penicillins, lincosamides, and fluoroquinolones (4,9). It is estimated that 15-25% of cases of antibiotic-associated diarrhea are caused by *C. difficile* (25).

Pathogenic strains of *C. difficile* have been identified with resistance to many common antibiotics including erythromycin, clindamycin, tetracycline, moxifloxacin (26). Sequencing of the *C. difficile g*enome has revealed a large number of putative resistance genes. Many of these, including tetracycline and erythromycin resistance genes, are located on transposable elements, implicating an important role of horizontal gene transfer in acquisition of antimicrobial resistance (27,28). Vancomycin and metronidazole are the most commonly used antibiotics in the treatment of *C. difficile* infections. So far, no strains have been isolated that are resistant to vancomycin, but reduced susceptibility to metronidazole is emerging (26). These antibiotics are not effective against the metabolically inert spores (29).

Epidemiology

Sohn, *et al.* have estimated that there are 7 CDAD case patients per 1,000 admissions in acute care hospitals (30). It should be noted, however, *C. difficile* burden varies dramatically by geographic region, between institutions, and even between wards of the same hospital (3,15,30,31). Thus, it is unclear how accurately this estimate reflects the true burden of CDAD. A number of groups have calculated the costs accrued during a hospital stay by patients with CDAD compared to patients whose stay was not complicated by CDAD. These estimates range from $2,000-10,000 in additional costs, a 38-78% increase over matched case-control patients (32,33). Total costs of *C. difficile* to the US health care system are thought to exceed $3 billion per year (34).

Although the estimates vary concerning the costs and rates of *C. difficile* infection and disease, it is abundantly clear that there has been a dramatic increase in the

number of infections between the late 1990s and mid 2000s. Rates of CDAD more than doubled in many localities (35). Moreover, there has been an increase in the severity of disease and fatal outcomes (36,37). According to death certificate data, *C. difficile* related deaths in the US rose from 5.7 deaths per million in the population in 1999 to 23.7 in 2004 (13). In England, the number of death certificates listing *C. difficile* as the primary cause of death rose from 499 to 3,393 between 1999 and 2006, surpassing the number of deaths caused by methicillin resistant *Staphylococcus aureus* (35). This vast upsurge in CDAD has been primarily attributed to the emergence of a more virulent strain, NAP1/027. The NAP1/027 strain has been reported to have higher production of TcdA and TcdB (38), a more cytotoxic form of TcdB (39,40), production of an additional toxin (41), higher rates of sporulation (42,43), and increased antibiotic resistance (41). The emergence of this new strain underscores the significance of *C. difficile* as one of the most threatening nosocomial pathogens.

Treatment

Disease typically occurs following a disruption of the natural intestinal flora due to the use of broad-spectrum antibiotics. Upon diagnosis of CDAD, use of the offending antibiotic is stopped. In the case of mild disease, symptoms often naturally subside and no additional treatment is necessary (4,44). In most cases patients are administered metronidazole or vancomycin, and symptoms typically begin to resolve within 3 days.

One of the biggest difficulties in treating CDAD is recurrence of disease after antibiotic treatment is halted. It is estimated that 15-35% CDAD patients relapse following treatment (44,45). Subsequent recurrences are even more likely. If an individual has had two or more episodes, the relapse rate is 33-65% (44,46). The high recurrence rates following antibiotic treatment is likely because both vancomycin and metronidazole prevent the growth of a large number of bacteria, preventing the

reestablishment of the intestinal flora. Metronidazole and vancomycin have even been reported to initiate CDAD (45). Following cessation of antibiotic treatment, the ecologically unstable colon may be re-colonized by remaining spores, which are resistant to the antibiotics, or by re-infection (29,47).

The FDA has recently approved a third drug, fidaxomycin, for treatment of *C. difficile*, which may reduce the burden of recurrent infections. Fidaxomycin (Dificid™) is a minimally absorbed, narrow spectrum antibiotic that has been shown in clinical trials to be as effective as vancomycin at treating *C. difficile* infections. Importantly, because fidaxomycin is fairly specific for *C. difficile*, it does not have such deleterious effects on the gut microbiota. In clinical trials the recurrence rates were reduced to 13-16% compared with 25-27% for vancomycin treatment, and fidaxomycin is thus a promising new treatment option for CDAD. Several other antibiotics including fusidic acid, teicoplanin, rifampicin, and bacitracin have been reported to clear infection, but none of these are approved for use in treating CDAD (46,48).

In addition to antibiotics, several alternative treatments exist, and several others are under development. Because CDAD typically follows a disruption of the intestinal flora, probiotics and bacteriotherapy are of great interest. *Saccharomyces boulardii* and *Lactobacillus rhamnosus* have been reported to prevent CDAD, but reports are inconsistent (49,50). Fecal bacteriotherapy, or stool transplant, has been used with high success rates (~90% report no recurrence) but is an unpalatable option and thus unlikely to become a commonly used therapy (46). A number of adsorbent resins or polymers have also been used to bind toxins before they can bind enterocytes. Thus far, some (e.g. Tolevamer) have had moderate efficacy, but they are not as good as current antimicrobial treatments (46). Finally, a number of immune strategies targeting TcdA and TcdB are being explored including the use of intravenous immunoglobulin and vaccination (9,46). Toxoid-based vaccines are currently in phase II clinical trials (51).

Disease in animals

C. difficile has been reported to not only colonize but also cause disease in a large number of animals including horses, dogs, hamsters, guinea pigs, ostriches, and elephants (52). Symptoms often include diarrhea, but clinical manifestations are variable (52). Hamsters have been chosen as the model system of choice for studies of *C. difficile*, because experimental infection of hamsters recapitulates many features of human disease (3). Most notably, disease in hamsters can be induced by antibiotic treatment. Hamsters sometimes show the same symptoms seen in human CDAD such as diarrhea and, in some cases, formation of pseudomembranes similar to those in PMC (3). In addition, neonatal hamsters, like human infants, are not susceptible to disease. However, damage is caused to a different part of the intestine in hamsters, and disease is more severe, typically resulting in death.

C. difficile is also capable of colonizing important food animals including pigs, cattle, and poultry. In neonatal pigs *C. difficile* is an important cause of enteritis (53). *C. difficile* does not seem to be pathogenic towards poultry or cattle, but notably these animals are colonized by many of the same strains that are pathogenic in humans. The isolation of *C. difficile* from retail meats has raised the question of whether *C. difficile* may be transmitted by consumption of spore containing foods (52). Given the ubiquity of *C. difficile* and the relative infrequency of community acquired CDAD, it is unlikely that foodborne or zoonotic transmission play important roles in the epidemiology of CDAD. Nevertheless, the colonization of food animals by pathogenic strains of *C. difficile* is alarming, considering that many of these animals are raised in confined animal feeding operations where antibiotics are often used indiscriminately. These conditions are favorable for the generation of strains with increased antibiotic resistance.

Virulence factors of *C. difficile*

Several factors have been implicated in the virulence of *C. difficile* including adhesins (54), extracellular enzymes (55), a binary toxin (56,57), fimbrae (58), flagella (59,60), capsule (58), and a paracrystalline S-layer (27,61). Yet, these factors are poorly understood, and in many cases their roles in pathogenesis are highly speculative. Because *C. difficile* has been difficult to manipulate genetically, none of the aforementioned virulence factors have been mutated and tested in animal models of disease. Nevertheless, two *C. difficile* proteins clearly stand out as important virulence factors, toxins A and B.

Like many pathogenic *Clostridia*, *C. difficile* produces several potent toxins. Three have been identified for *C. difficile*: TcdA, TcdB, and the binary toxin CDTab. TcdA and TcdB are 308 and 270 kDa proteins with 49% identity and 63% similarity. They belong to a larger family of large clostridial toxins (LCTs) which includes lethal and hemorrhagic toxins from *C. sordellii* (TcsL and TcsH), α-toxin from *C. novyi* (Tcnα), and large cytotoxin from *C. perfringens* (TpeL) (Table 1). LCTs are homologous toxins that inactivate host Rho and Ras family guanosine triphosphatases (GTPase) by glucosylation. In *C. difficile* infections, TcdA and TcdB are secreted into the colon where they are cytotoxic and proinflammatory (62). All pathogenic strains of *C. difficile* that have been identified, thus far, produce TcdB; and the majority produce both toxins (63). CDTab is an actin-specific ADP-ribosyltransferase that is homologous to iota toxin from *C. perfringens* (56,57). Many pathogenic strains do not produce CDTab (64), and researchers have struggled to find an association between production of binary toxin and pathogenesis (57,65,66).

Table 1-1. Large clostridial toxins

Toxin	Species	Mol. weight	Lethal activity[1]	Targets
TcdA	*difficile*	308 kDa	50 ng	Rho, Rac, Cdc42, Rap
TcdB	*difficile*	270 kDa	50 ng	Rho, Rac, Cdc42
TcdB-F[2]	*difficile*	270 kDa	50 ng	Rac, Ras, Ral, Rap
TcsH	*sordellii*	~300 kDa[3]	75 ng	Rho, Rac, Cdc42
TcsL	*sordellii*	270 kDa	5 ng	Rac, Ras, Ral, Rap
Tcnα	*novyi*	250 kDa	5-10 ng	Rho, Rac, Cdc42
TpeL	*perfringens*	191 kDa	16 µg	Rac, Ras, Ral, Rap

[1] Amounts listed are for one mouse lethal dose (MLD), and were obtained by intraperitoneal injection (67,68).

[2] TcdB-F is a variant TcdB produced by the TcdA⁻TcdB⁺ strain 1470 (69).

[3] The amino acid sequence for TcsH is unknown. The estimated molecular weight is based on SDS-PAGE (70).

The *C. difficile* toxins are chromosomally encoded at two distinct loci (Figure 1-1). TcdA and TcdB are chromosomally encoded in a region known as the pathogenicity locus. The locus consists of five genes: the toxin-encoding genes, *tcdA* and *tcdB*; two regulatory genes, *tcdC* and *tcdD* (also called *tcdR*); and a putative holin, *tcdE* (Figure 1-1). TcdD is a sigma factor that positively regulates transcription of TcdA and TcdB. TcdC, on the other hand, is a negative regulator of toxin expression (71). Interestingly, in the hypervirulent NAP1/027 strain there is an 18 base pair deletion in *tcdC*, which is thought to result in a 16-23 fold increase in toxin production (38). TcdE has homology to holins, and it has been suggested that this protein may be involved in release of the toxins by permeabilization of the bacterial membrane (72). *In vitro*, expression of TcdA and TcdB occurs in stationary phase and correlates with a decline in the levels of TcdC and increase in TcdD (73). However, the factors regulating TcdC and TcdD are unknown. The expression profile of these proteins *in vivo* is unclear, but TcdA and TcdB can be detected in the stool in patients experiencing symptoms of CDAD.

Pathogenicity Locus   CDT Locus

tcdD   *tcdB*   *tcdE*   *tcdA*   *tcdC*   *cdtR cdtA cdtB*
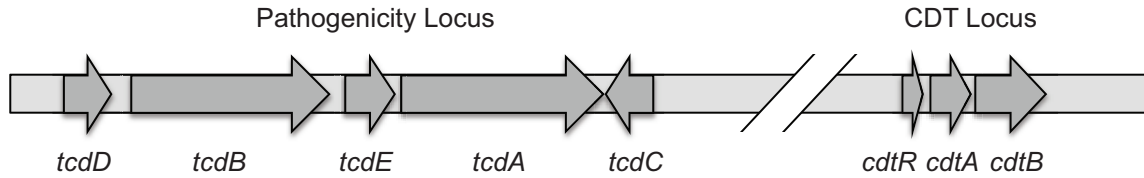
Figure 1-1. Organization of the toxin genes in *Clostridium difficile*. The large clostridial toxin and binary toxin genes are encoded in the chromosome at distant loci. The genes encoding TcdA and TcdB are found in a 20 kb pathogenicity locus along with genes encoding TcdD and TcdC, which are positive and negative regulators of toxin expression, respectively. *tcdE* encodes a holin-like protein that may be involved in release of the toxins. *cdtA* and *cdt*B encode the binary toxin genes; *cdtR* encodes a sigma factor that regulates their expression.

Roles of TcdA and TcdB in disease

Molecular targets of TcdA and TcdB

TcdA and TcdB harbor N-terminal glucosyltransferase domains which, upon being delivered into the cell, inactivate small GTPase proteins of the Rho family (Table 1-1). The Rho family includes the prototypical members RhoA, Rac1, and Cdc42 and belongs to the larger superfamily of Ras GTPases. The Rho and Ras family GTPases are master regulators of a number of vital cellular processes including cycle progression, cell-cell adhesion, cytokinesis, secretion, and maintenance of the cytoskeleton (74,75). Thus, it is not surprising that these GTPases are targets of a large number of bacterial toxins and effectors (76).

In an ordinary cell, GTPases cycle between active, GTP-bound and inactive, GDP-bound states (Figure 1-2). The cycle is controlled by regulatory molecules including guanosine dissociation inhibitors (GDIs), GTPase activating proteins (GAPs), and guanine nucleotide exchange factors (GEFs). Upon binding GTP, Rho undergoes a conformational change allowing it to interact with several effector molecules (75,76). The repertoire of effector molecules is extensive and includes protein and lipid kinases,

phophatases, lipases, and scaffolding proteins. The downstream signaling cascades induced by activation of these effectors are varied and complex. Most are poorly understood. Nevertheless, it is clear that these proteins have key roles in the regulation of the cytoskeleton and processes requiring cytoskeletal rearrangements.

TcdA and TcdB preferentially act on GDP-bound, membrane associated Rho (77,78). The toxins inactivate Rho proteins by monoglucosylation of Thr37 (Thr35 of Rac and Cdc42) (77,79). The Thr37 residue is located in the switch 1 region and participates in Mg^{2+} and GTP binding (80). This region is involved in binding effectors and regulatory molecules and undergoes a significant conformational change upon activation (74,76,78). Glucosylated Rho GTPases are no longer able to cycle between the soluble and membrane-associated states (78). Both GAP-stimulated hydrolysis and GEF-stimulated exchange of nucleotides are also blocked (81,82). Most importantly, Rho is no longer able to switch into its active conformation and bind many of its effectors and therefore, numerous downstream signaling pathways are interrupted (74,76).
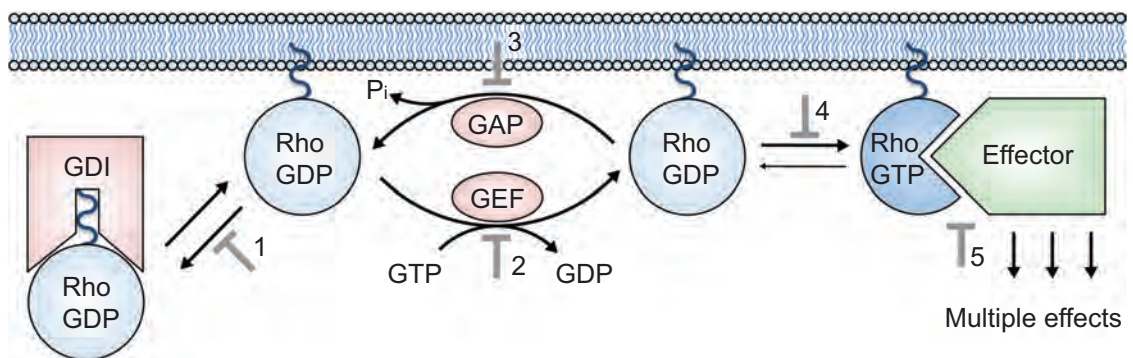
Figure 1-2. Inhibition of the GTPase cycle by TcdA and TcdB. GTPases (here represented by Rho in blue) cycle between active GTP-bound (darker blue) states at the plasma membrane and inactive states (lighter blue). The cycle is controlled by regulatory proteins (pink) including GDIs, GEFs, and GAPs. GAPs promote the hydrolysis of GTP to GDP. GDIs sequester inactive GDP-bound Rho in the cytosol. GEFs promote activation of Rho by stimulating exchange of GDP for GTP. Activated Rho binds to multiple effectors leading to a variety of downstream effects. Glucosylation of Rho by TcdA or TcdB results in an inability to interact with regulatory molecules (1-3), switch into the active conformation (4), and bind effector proteins (5).

Effects of the toxins on cells

Though the Rho GTPases affect many cellular processes, their principle role is the regulation of the cytoskeleton. Accordingly, the most notable change in cells treated with TcdA or TcdB is the loss of cytoskeletal structure. In cultures of adherent cells, treatment with toxin leads to rounding of the cells (Figure 1-3). Cell-rounding is one of the most commonly used techniques to characterize the presence and activity of TcdA and TcdB. Both toxins induce rounding in a wide range of cell types. The cytopathic potency of the toxins does vary between cell lines; however, TcdB is 100-10,000 times more potent than TcdA towards most cell types (83-86).

| Mock | TcdA | TcdB |
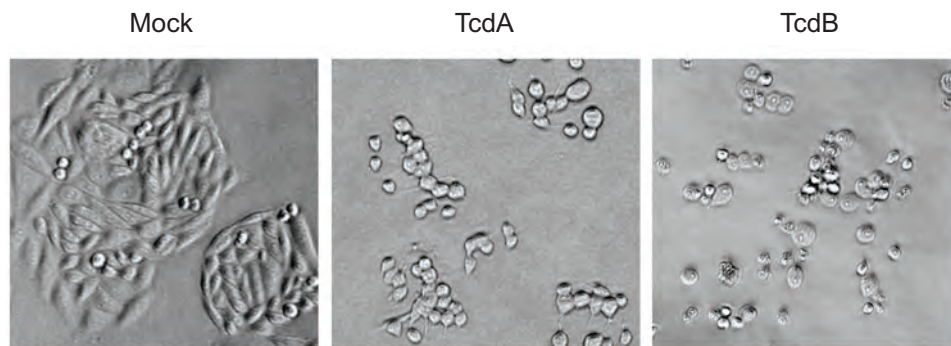
Figure 1-3: Cytopathic effects of TcdA and TcdB on cultured cells. CHO cells were plated in a 96 well cell culture dish at 1,000 cells per well in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS). Wells were treated with (A) buffer, (B) 0.2 µg TcdA, or (C) 0.2 µg TcdB for 4 hours and visualized by light microscopy.

TcdA and TcdB also cause cells to die. Cell death and cell-rounding are distinct events in intoxication (87). To distinguish between the two events, the cell rounding and cell death are sometimes referred to as the cytopathic effect and cytotoxic effect, respectively. TcdA and TcdB have been reported to cause death through a number of different mechanisms including p53-dependent and p53-independent apoptosis, caspase-dependent and caspase-independent apoptosis, as well as necrosis (87-92). It

is challenging to assimilate these and other data into a cohesive model of how *C. difficile* kills cells. It is apparent from these and other studies that inactivation of Rho proteins affects many cell-signaling processes in ways that are deleterious to the cell.

TcdA and TcdB have a number of effects on cells that do not necessarily result in rounding or death, yet may contribute to pathogenesis. Notably, Rho proteins have been demonstrated to have a role in regulation of tight junctions, and the perijunctional actin ring (93). Accordingly, inactivation of Rho GTPases by TcdA and TcdB results in the disruption of cell-cell junctions (94,95). Toxin stimulated disruption of cell-cell junctions likely contributes to the increased epithelial permeability and lumenal fluid accumulation associated with CDAD.

In addition, TcdA and TcdB induce the secretion of cytokines in epithelial and immune cells. TcdA has been shown to directly or indirectly induce the production of macrophage inflammatory protein 2, substance P, IL-1, IL-6, IL-8, transforming growth factor β, tumor necrosis factor, and intestinal secretory factor (95-100). Cytokine production in response to TcdB has not been characterized as extensively, but TcdB has been shown to induce the production of tumor necrosis factor, IL-1, IL-6, and IL-8 (101). IL-8, in particular, is thought to play a pivotal role in *C. difficile* pathogenesis. IL-8 is involved in the recruitment and activation of neutrophils, which are present in high amounts at sites of *C. difficile* associated inflammation. Furthermore, a polymorphism in the IL-8 gene has been associated with susceptibility to recurrent CDAD (102).

Putative roles of the two toxins in disease

The activities of purified TcdA and TcdB have been investigated in a number of different animal models including mice, rats, rabbits, and hamsters. Although the manifestations of disease vary in the different animals, in these animal models TcdA induces fluid accumulation and inflammation within the intestinal tract, whereas TcdB

14

causes minimal or no intestinal pathology (103-105). In the rabbit ligated ileal loops, for example, TcdA induces fluid accumulation with comparable activity to cholera toxin (105). Unlike cholera toxin, the toxin also causes extensive tissue damage and the accumulated fluid is hemorrhagic (3,105). TcdB has no effect in this assay. Likewise, when TcdA and TcdB are given to mice and hamsters intragastrically, TcdA causes intestinal hemorrhage, diarrhea, and death, but TcdB has no effect (104). TcdB, is however, a potent toxin towards these animals when administered by other routes. When given by intraperitoneal injection, TcdA and TcdB are both lethal toxins with similar potencies (Table 1-1) (67).

Based on these early studies TcdA was referred to as the enterotoxin, and TcdB, because of its ~1000 fold higher cytopathic potency toward cultured cells, was referred to as the cytotoxin. It was proposed that TcdA induced the initial damage in the colon where TcdB alone was relatively inert. After TcdA had disrupted the intestinal epithelium, TcdB might then access and act on other tissues (3). This model implicated TcdA as the key virulence factor in disease and TcdB as, perhaps, an accessory virulence factor. In support of this model, Lyerly *et al*. showed that if TcdB is given to hamsters with damaged intestines or together with sublethal doses of TcdA the animals die. Furthermore, it was shown that a humoral immune response against TcdA correlated with protection from disease both in humans and in animal models (106-109).

More recently, a number of pieces of evidence have shown TcdB has a much more important role in disease than previously appreciated. Since the 1990s, a number of pathogenic strains have been identified that produced TcdB, but not TcdA. These TcdA$^-$TcdB$^+$ strains produce the same range of symptoms as TcdA$^+$TcdB$^+$ strains in humans, and are pathogenic in animal models (110). Thus TcdA is not essential for pathogenesis. To date, no pathogenic TcdA$^+$TcdB$^-$ clinical isolates have been identified.

15

In addition, there have been reports that while TcdB may not be enterotoxic in animal models, it is toxic toward human colonic tissues. Riegler, *et al.* tested the effects of TcdA and TcdB on human colonic explants (111). Electrical resistance across the membrane was monitored in Ussing chambers following exposure to TcdA and TcdB. Damage to the membrane was confirmed by light, fluorescent, and scanning electron microscopy. Riegler, *et al.* found that TcdB was ~10 times more potent than TcdA at disrupting the integrity of the membrane (111). Savidge, *et al.* have demonstrated the toxicity of TcdB in a chimeric mouse model of disease where a human fetal intestinal xenograft is transplanted into an immunodeficient mouse (101). When challenged with either TcdA or TcdB, the human xenografts exhibited intestinal epithelial cell damage with marked necrosis, increased mucosal permeability, and acute mucosal inflammation.

Elucidating the roles of TcdA and TcdB has been slow partially due to the lack of genetic tools for the manipulation of *C. difficile* (112). In 2009, Lyras *et al.* reported the construction of the first isogenic *tcdA* and *tcdB* mutants (113). Clindamycin-treated hamsters were challenged with the isogenic strains lacking either toxin. Surprisingly, knocking out TcdA seemed to have no effect on the virulence of the strain. Strains lacking TcdB, however, were avirulent. Death of the hamster was the only reported phenotype, so it is not clear if the TcdA$^+$TcdB$^-$ strains had subtler pathological effects. Nevertheless, this study indicated that TcdB is the more important virulence factor. However, when the same experiments were carried out by another group, TcdA and TcdB deficient mutants were both shown to cause death in hamsters (114). This second group saw a loss of virulence only when both *tcdA* and *tcdB* were disrupted.

Thus, although TcdA and TcdB have long been accepted as the primary virulence factors of *Clostridium difficile*, the relative importance of these proteins in disease remains a matter of ongoing debate. Considering that TcdA and TcdB have different properties and that most pathogenic strains produce both toxins, we think it

unlikely that these proteins have identical roles. Understanding the nature of these roles will require further study.
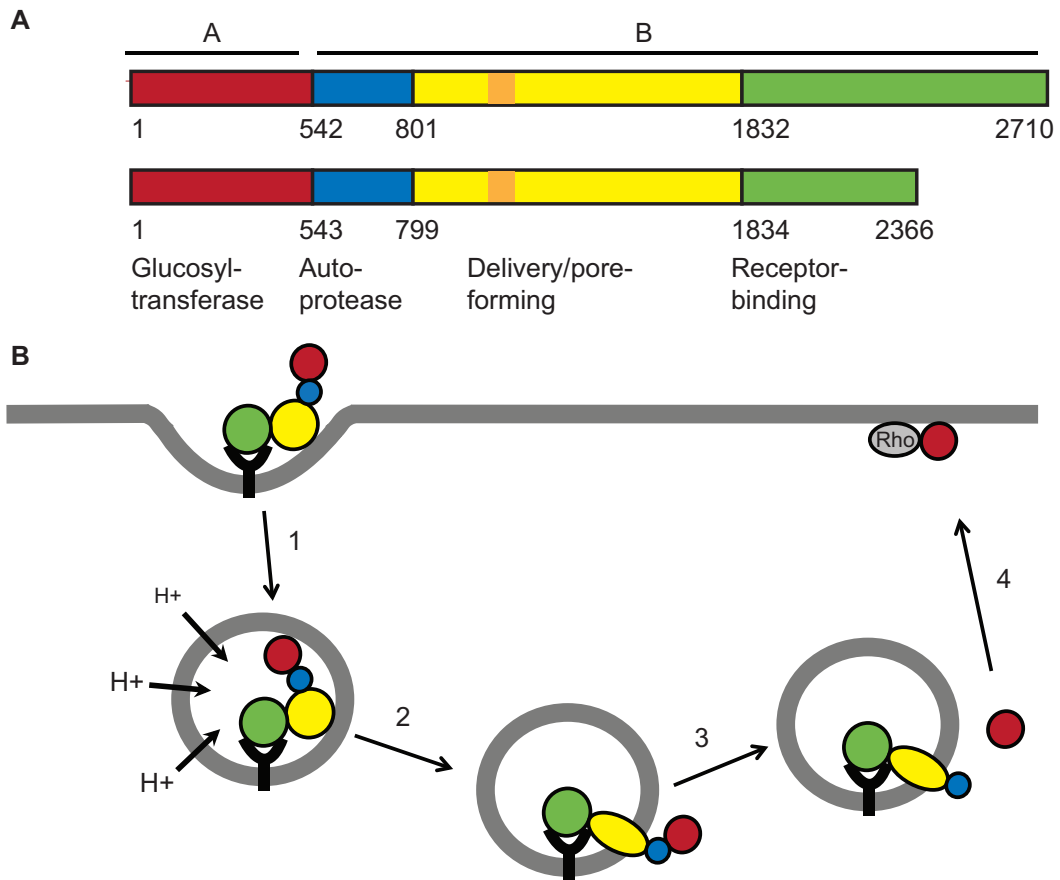
Figure 1-4. TcdA and TcdB primary structure and mechanism of cellular intoxication. (A) TcdA and TcdB are homologous AB toxins consisting of four domains. The enzymatic A component is an N-terminal glucosyltransferase domain (GTD) (red). The B component, involved in delivery of the GTD into the cell, has three identified domains: receptor-binding (green), 'delivery' or pore-forming (yellow), and autoprotease domains (blue). The orange box represents the hydrophobic region of the delivery domain that has been proposed to form part of the transmembrane pore (amino acids 956-1128 of TcdB). (B) The delivery process is divided into four main steps that are mediated by each of the four domains. 1) The toxin binds to the surface of the cell and is internalized by receptor-mediated endocytosis. 2) Acidification of the endosome triggers the formation of a pore through which the GTD is translocated. 3) The GTD is released into the cytosol by InsP6 dependent autoproteolysis. 4) The GTD glucosylates Rho-family GTPases at the cell membrane.

17

TcdA and TcdB structure and mechanism of action

TcdA and TcdB are homologous AB toxins. They can be divided into two components: an enzymatic A subunit and a B subunit involved in the delivery of the A subunit into the target cell (Figure 1-4). The A subunit is an N-terminal glucosyltransferase domain (GTD) that inactivates host Rho proteins by glucosylation. Three additional domains (B component) are responsible for delivery of the GTD to the cytosol of the host cell. These include a receptor-binding domain (RBD), a 'delivery' or pore-forming domain, and an autoprotease domain. The cytotoxic mechanism can be divided into four steps that are mediated by the four known domains: 1) binding/internalization, 2) pore-formation and translocation of the GTD across the membrane, 3) release of the GTD by autoproteolysis, and 4) inactivation of host Rho proteins by glucosylation (Figure 1-4B). Each of these steps is discussed in greater detail below.

Receptor-binding

The C-terminal RBDs of TcdA and TcdB consist of amino acids 1832-2710 and 1834-2366, respectively (115). The sequences that make up the domains are highly repetitive, consisting of multiple 19-24 amino acid short repeats (SR) and 31 amino acid long repeats (LR) (74,116). The TcdA RBD is made up of 32 SRs and 7 interspersed LRs (Figure 1-5B). The TcdB RBD is considerably shorter and contains 19 SRs and 4 LRs. Together these repeats form cell wall binding motifs that bind sugar moieties on the surface of host cells (117,118). Cell wall binding motifs from TcdA were initially shown to bind α-Gal-(1,3)-β-Gal-(1,4)-β-GlcNAc (119). However, this sugar is not present on human cells (120). TcdA has since been reported to bind to the human I, X, and Y blood antigens as well as a human glycosphingolipid (120,121). These molecules all have the

core β-Gal-(1,4)-β-GlcNAc structure found in α-Gal-(1,3)-β-Gal-(1,4)-β-GlcNAc. It is not know which of these, if any, serve as the native ligand in the human colon.

Two proteins have been implicated as receptors for TcdA. Rabbit sucrose-isomaltase was first shown to serve as a receptor for TcdA in the rabbit ileum (122). Binding to sucrose-isomaltase was inhibited by galactosidase treatment, indicating that the toxin binds the glycosyl modification on the protein, but the identity of these sugars is unknown (122). Rabbit sucrose-isomaltase cannot be the only receptor for TcdA, because many cells and tissues, including the human colonic epithelium, do not express this protein yet are sensitive to TcdA (122). More recently, the human protein gp96 has been reported as a receptor for TcdA (123). Like sucrose-isomaltase, gp96 is predicted to be glycosylated, but the identities of the saccharides have not been determined. It has not been shown for gp96 whether or not the saccharide modifications are involved in binding TcdA. No receptors have been described for TcdB.

In 2005, Ho *et al*. published the crystal structure of a fragment of TcdA comprising amino acids 2573-2709 (TcdA RBD f1) (Figure 1-5A) (pdb 2F6E (115)). This fragment contains 4 SRs and 1 LR. The structure revealed that each repeating element consists of a β-hairpin followed by a loop. In the SRs, the loops are 7-10 amino acids in length, whereas the LR loop comprises 18 amino acids. The SRs are packed together in a regular fashion where each repeat is rotated by ~120° in relation to the previous one (115). The repetitive stacking of SRs forms straight, rod-like structures. The LR's packing with its adjacent SRs disrupts the regular repeating arrangement of the SRs, causing a ~30° kink in the rod-like structure (115). Using the structure of the TcdA RBD f1 as a template of SR-SR and SR-LR interactions, Ho, *et al.* constructed a model of the entire TcdA and TcdB RBDs (Figure 1-5C).
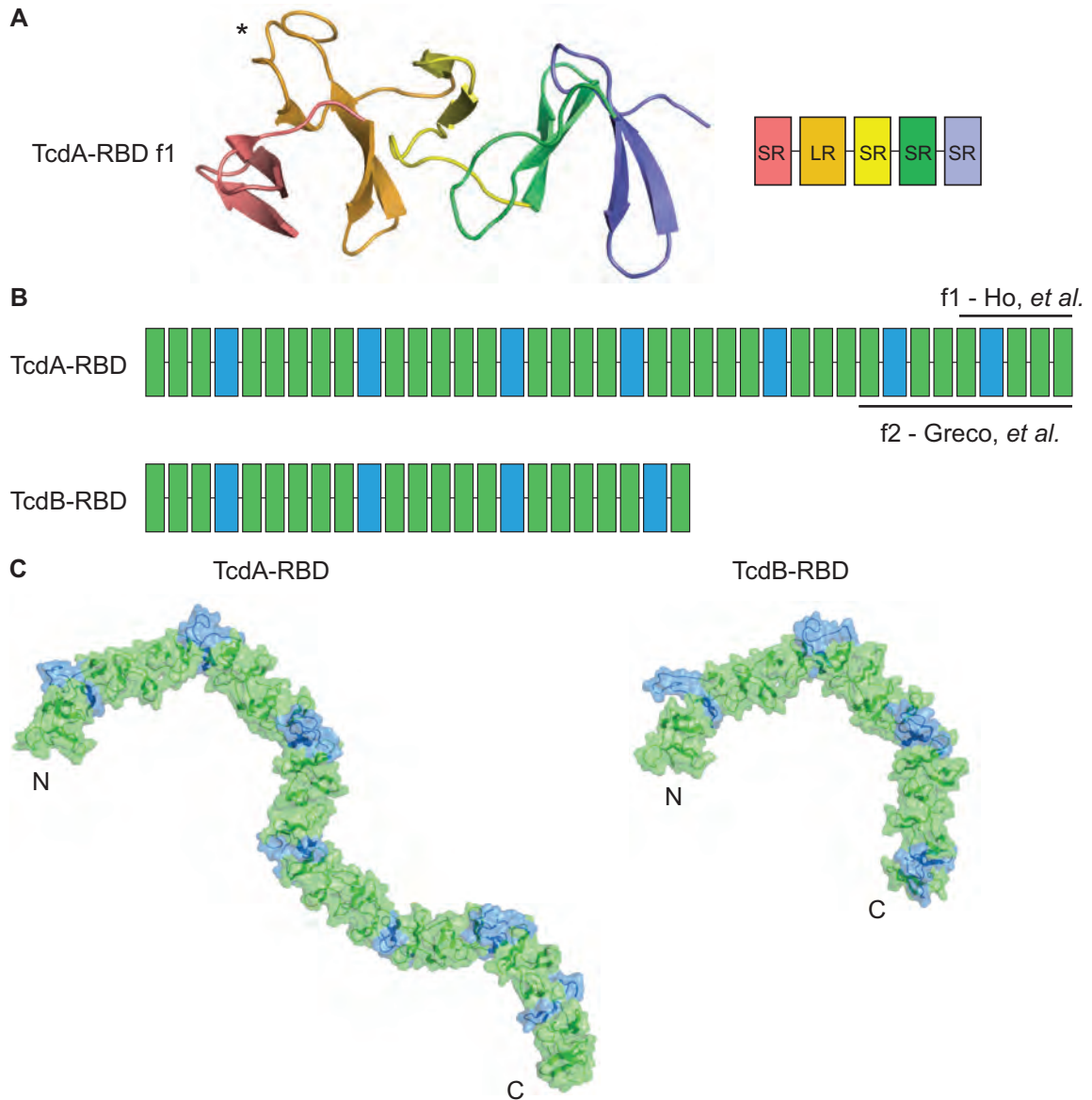
Figure 1-5. Structure of the TcdA receptor-binding domain. (A) Crystal structure of TcdA-RBD fragment 1 (f1), consisting of residues 2573-2709 (115). The structure consists of repeating elements that contain a β-hairpin followed by a loop. In short repeats (SR), the loops comprise 7-10 amino acids; in long repeats (LR) the loops consist of 18 amino acids (LR loop indicated by *). The SRs pack together into relatively straight rod-like structures; whereas the LR packing with its adjacent SRs results in a ~30° kink in the structure. (B) The entire TcdA RBD is made up of 32 SRs with 7 interspersed LRs, which are represented by green and blue boxes, respectively. The TcdB RBD consists of 19 SRs and 4 LRs. (C) Models of the TcdA and TcdB RBDs were constructed based on the structure of TcdA RBD f1. The models are colored as in (B).

The same group later published a co-crystal structure of a larger fragment (f2) of the RBD in complex with an α-Gal-(1,3)-β-Gal-(1,4)-β-GlcNAc derivative (pdb 2G7C (124)). This structure revealed that the saccharides bind at the junctions formed between LRs and SRs (Figure 1-6). Thus, each of the kinked regions in Figure 1-5C represents a putative saccharide binding side. TcdA has 7 such binding sites, whereas TcdB has 4. The model taken from these structures suggests multivalent binding along an extended binding domain (124).
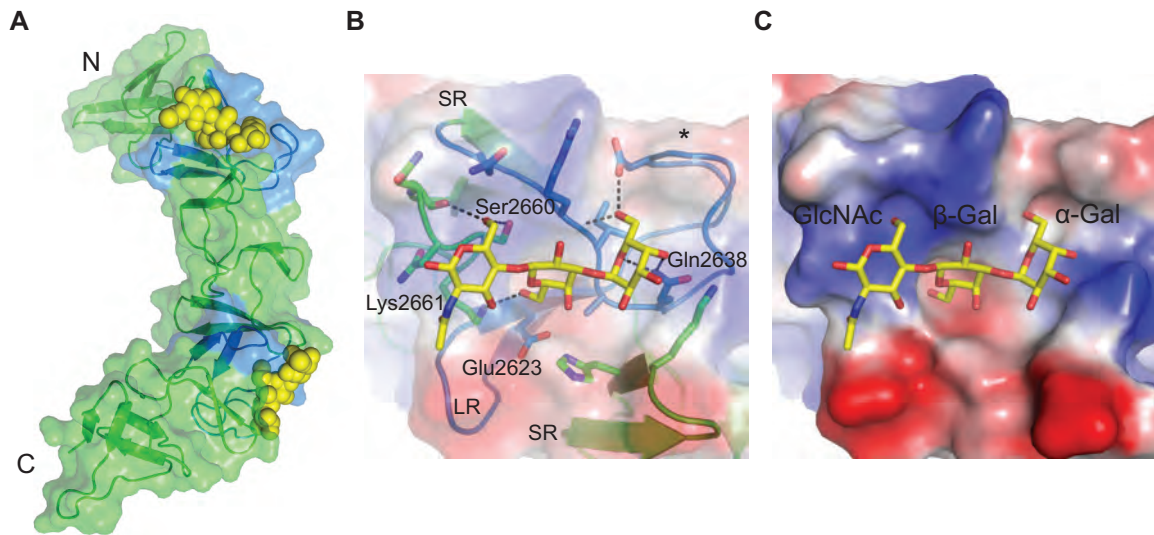
Figure 1-6. Carbohydrate recognition by the TcdA RBD. (A) The crystal structure of TcdA RBD f2 (amino acids 2456-2710) is shown with the SRs colored green and LRs colored blue. The head-group of the bound liposaccharide α-Gal-(1,3)-β-Gal-(1,4)-β-GlcNAcO$(CH_2)_8CO_2CH_3$ is shown in yellow. (B-C) A close-up view of the C-terminal trisaccharide-binding pocket formed by an LR and adjacent SR. The electrostatic surface potential is shown with negatively-charged surfaces in red and positively-charged surfaces in blue (transparent in (B) and opaque in (C)). In (B) residues that make up the trisaccharide binding pocket are shown as sticks, and hydrogen bonds are shown as dashed lines. The labeled residues are strictly conserved in the saccharide-binding repeats of the TcdA RBD, but are divergent in the TcdB RBD.

The binding of α-Gal-(1,3)-β-Gal-(1,4)-β-GlcNAc is shown in Figure 1-6. The terminal α-Gal abuts the large loop of the LR and makes several hydrogen-bonding contacts with residues of the loop (e.g. Gln2638) (Figure 1-6B). The central β-Gal moiety

sits with its methoxy group pointed down into a small basic cleft formed in part by Glu2623. The oxygen of the methoxy group hydrogen bonds with Lys2661 of the SR. The GalNAc group interacts with a number of SR residues including a hydrogen bond formed with Ser2660. The residues pointed out here (Glu2623, Gln2638, and Ser2660, Lys2661) are all strictly conserved in the saccharide binding repeats of the TcdA RBD. Yet, these residues are strikingly different in the TcdB RBD repeats. Glu2623, for example, which contributes to the small acidic pocket where the β-Gal binds is replaced by a bulky, basic residue (lysine) in each of the TcdB LRs. Thus, while the TcdB RBD is predicted to have a similar fold to the TcdA RBD, the saccharide binding sites are radically different. This provides a likely explanation for why TcdB does not bind the β-Gal-(1,4)-β-GlcNAc with the affinity of TcdA.

Many gaps remain in our understanding of the binding of the *C. difficile* toxins to target cells. Chief among these is the identity of the receptors. While there is some knowledge of TcdA binding targets, nothing is known about the nature of the TcdB receptor. Both TcdA and TcdB act on many cell-types with varying potency in tissue culture experiments, but the target cells *in vivo* are unknown. It is also unknown whether or not the carbohydrate-binding repeats are the only parts of the proteins involved in binding to cells. The homologous toxin TpeL from *C. perfringens* lacks the C-terminal repeats yet is still toxic, albeit less potent (68). Likewise, removing the RBD from TcdA or TcdB attenuates, but does not eliminate cytopathicity suggesting the existence of an additional binding activity outside of the RBD (125,126). Finally, it is not known how binding induces uptake. TcdA and TcdB were recently shown to be taken in by clathrin-mediated endocytosis, but the mechanism by which internalization is stimulated has not been discovered (127).

Pore formation

Once the toxins have been internalized, the glucosyltransferase must be delivered across the endosomal membrane. Based on analogy to many other AB toxins, endosomal acidification is thought to induce structural changes in the delivery domain that expose hydrophobic segments. These hydrophobic regions insert into the host membrane forming a pore through which the glucosyltransferase domain can pass. Although this has not been demonstrated for TcdA or TcdB, data from previous studies are consistent with this model.

Qa'dan, *et al*. have shown that acidification causes structural rearrangements in TcdB. The authors demonstrated that upon reduction of pH to 4.0, TcdB exhibits differences in native tryptophan fluorescence and protease susceptibility. As expected, these changes resulted in the exposure of hydrophobic surfaces as detected using an environment-sensitive fluorescent probe (128). Barth et al., have shown that acidification not only induces conformational changes, but also triggers pore formation. Chinese hamster ovary (CHO) cells were preloaded with $^{86}Rb^+$, and then treated with TcdB. Under acidic conditions, $^{86}Rb^+$ was released into supernatant when TcdB was present, indicating that TcdB inserts into the membrane and forms ion channels. The authors further showed that TcdB can induce ion channels in artificial lipid bilayers at low pH (129). TcdA has also been shown to form pores in cells and in artificial bilayers (130). For TcdA, but not TcdB, cholesterol is essential for pore-formation (130). Barth *et al.* and Qa'dan, *et al*. further demonstrated that these results are applicable in natural intoxication of cells. Treatment of cells with pharmacological inhibitors of endosomal acidification (*e.g.* bafilomycin A1) retards the toxic effects of TcdB, presumably by inhibiting pore formation (128,129).

Within TcdA and TcdB, amino acids ~956-1128 are particularly hydrophobic and have been hypothesized to compose at least part of the transmembrane pore (116,118).

Genisyuerek, *et al*. have recently constructed a series of TcdB truncations to delineate the region of the toxin responsible for pore-formation (131). Their results indicated that the amino acids 830-990 contain the minimal pore-forming region (131). This raises the question of what the role of the other 875 residues in the delivery domain is in intoxication. It has been suggested that the C-terminal part of the domain may have a role in binding to cells, but there is very minimal evidence for this (125,126). So far, nothing is known about the structure of the delivery domains of TcdA or TcdB in either the soluble or membrane states. It is not even known whether the toxins must oligomerize in order to form a pore. Understanding the structural changes that allow TcdA or TcdB to go from being soluble proteins to membrane pores is a challenging structural problem important for understanding the function of these proteins.

Autoproteolysis

LCTs are made and secreted as single polypetide chains. It was unknown for many years whether the toxin remained intact or if the A component (GTD) was released into the cell as is typical for other AB toxins. It was first shown in 2003 that TcdB is proteolytically processed and that only the N-terminal GTD is released into the cytosol (132). In 2005, the cleavage site was localized between Leu543 and Gly544, and it was revealed that cleavage was mediated by a component of the target cell cytosol (133). It was thought that this proteolytic event was mediated by a host protease (133). Surprisingly, when Reineke, *et al.* tried to identify the host factor required for cleavage of TcdB, they found that the host factor was not a protein but a small molecule, one of the inositol-phosphates (134). The most active of the inositol-phosphates was inositol hexakisphosphate (InsP6) (134). When mixed with purified LCTs, InsP6 alone was enough to induce cleavage. Reineke, *et al.* further showed that treatment with 1,2-epoxy-3(p-nitrophenoxy)propane, a covalent inhibitor of aspartate proteases, modified

Asp1665 and inactivated the toxin. Thus, the authors proposed that Asp1665 is the catalytic residue of an aspartate activity encoded in the central region of the toxin. However, this could not explain InsP6-induced cleavage of Tcnα, which does not have an aspartate in this position.

Just a few months later, Egerer, *et al.* published a second paper describing the autoproteolytic processing of TcdA and TcdB (135). They confirmed that cleavage was stimulated by InsP6, and showed that reducing agents enhanced the rate of cleavage. The major finding, however, was that cleavage was mediated by a cysteine protease domain (CPD) adjacent to the GTD. This finding has been borne out in a number of subsequent studies including our own (Chapter II), whereas there have been no further reports of aspartate protease activity for any LCT.

The autoprotease domain was identified based on its homology to a CPD that had just been discovered in a large multifunctional autoprocessing toxin (MARTX) from *Vibrio cholerae* (136). As in the LCTs, the *V. cholerae* MARTX (VcRTx) toxin CPD facilitates the release of enzymatic moieties into the target cell, and it is also activated by InsP6 (137). In 2008, the structure of the VcRTx CPD bound to InsP6 was published (pdb 3EEB (138)). It was shown that InsP6 binds in a basic pocket and allosterically activates cleavage at a separate catalytic site containing a catalytic triad of cysteine, histidine, and aspartate residues. The TcdA CPD and VcRTx CPD are only 19% identical, yet the residues that make up the putative catalytic triad are strictly conserved among all LCTs and necessary for InsP6-induced proteolysis of TcdB (135-137,139).

The discovery of the autoproteolytic processing of LCTs and identification of the responsible domain have filled a critical gap in the model of GTD delivery into target cells. Moreover, this activity presents a possible new therapeutic target for inhibiting toxin action. A deeper understanding of the structural and mechanistic details is an

important next step in advancing our understanding of the mechanism by which these proteins intoxicate cells. This is the subject of Chapter II of this thesis.

Glucosyltransfer

The N-termini of the LCTs contain a 63 kDa GTD responsible for cytopathicity (140). As discussed previously, once the GTD is translocated into the cytosol it inactivates small Rho GTPases by glucosylation of a threonine in the switch region. UDP-glucose serves as the source of glucose for TcdA, TcdB, TcsH, and TcsL (77,141). The homolog Tcnα, however, uses UDP-GlcNAc (142). TpeL can use either UDP-glucose or UDP-GlcNAc as the co-substrate (143).

Based on *in vitro* experiments, TcdA and TcdB can target RhoA, RhoB, RhoC, RhoG, Rac1, Cdc42, and TC10 (74,144). TcdA has also been reported to modify the additional substrates Rap1A and Rap2A which are more closely related to Ras family proteins (84). TcsH and Tcnα also target Rho family substrates (Table 1-1) (142,145). TcsL and TpeL, on the other hand, act mostly on Ras family proteins such as H-Ras, Ral and Rap (143). They can also target Rac but not Rho. Like the Rho proteins, the Ras GTPases are involved in a large number of cellular signaling pathways. They have important roles in the regulation of cell-cell junctions, cell proliferation, and survival (146).

A number of strains have been identified that produce a variant TcdB which modifies Ras family substrates (69,147,148). These are sometimes referred to as functional hybrids because most of the TcdB protein is the same as in prototypical strains (e.g. 630), but the GTD has the substrate recognition properties of TcsL. Interestingly, all of the TcdA$^-$TcdB$^+$ strains that have been characterized so far carry these variant toxins. It is not known why some toxins preferentially target Rho family

proteins while others target Ras family proteins, but it is clear that the inactivation of these different proteins has different effects on the cells (69,147-149).

In 2005, Reinert *et al.* determined the crystal structure of the GTD of TcdB in complex with UDP-glucose (pdb 2BVM (150)). The structures of the GTDs from two other LCTs, TcsL and Tcnα, have also been determined (pdb 2VKD and 2VK9 (151)). These structures coupled with biochemical data have helped shed some light on the enzymatic mechanism of glucosyltransfer.

The structure of the TcdB GTD is presented in Figure 1-7. At the core of the structure is a Rossman fold similar to what is seen in other glycosyltransferases belonging to the glycosyltransferase A (GT-A) family (150). In addition to the common GT-A family fold, TcdB has a number of α-helical additions (red). These include an N-terminal 90-residue subdomain at the 'bottom' of the GTD, as presented in Figure 1-7A, and several other large protuberances at the 'top right' and 'top left'. The N-terminal subdomain has recently been shown to target the GTD to the plasma membrane, the site of the target GTPases (152,153). This region is, therefore, referred to as the membrane localization domain (MLD). The role of the other α-helical additions is unknown, but it has been suggested that these residues may be involved in substrate binding (150).

The structure of the TcdB GTD facilitated the identification of important residues in the reaction mechanism. In the TcdB GTD structure, UDP-glucose is hydrolyzed into UDP and glucose, but both molecules are bound along with the essential cofactor Mn^{2+}. As in other GT-A proteins, UDP-glucose is bound in a pocket formed by the edge of the β-sheet and several α-helices. A loop consisting of residues 517-523 overlays the binding pocket and two residues within this loop (Ser518 and Trp520) are involved in binding the phosphate group of UDP (Figure 1-7C). Figure 1-7C shows several other key residues that are involved in binding UDP-glucose. The activity of the enzyme is reduced

or abolished when these residues are mutated, with the possible exception of Glu515 and Asn139 which have not been tested (140,154). The structure was also used as a platform to discover residues that are involved in binding target GTPases. Jank, *et al.* have found 5 residues (Glu449, Arg455, Asp461, Lys463, and Glu472) that when mutated result in loss of substrate modification. These residues are located immediately to the 'left' of the UDP-glucose binding pocket (cyan in Figure 1-7B). The transferred glucose is only accessible from this 'front' view of the GTD, thus the GTPases must bind to this side of the molecule. A detailed model of substrate binding, however, is lacking.
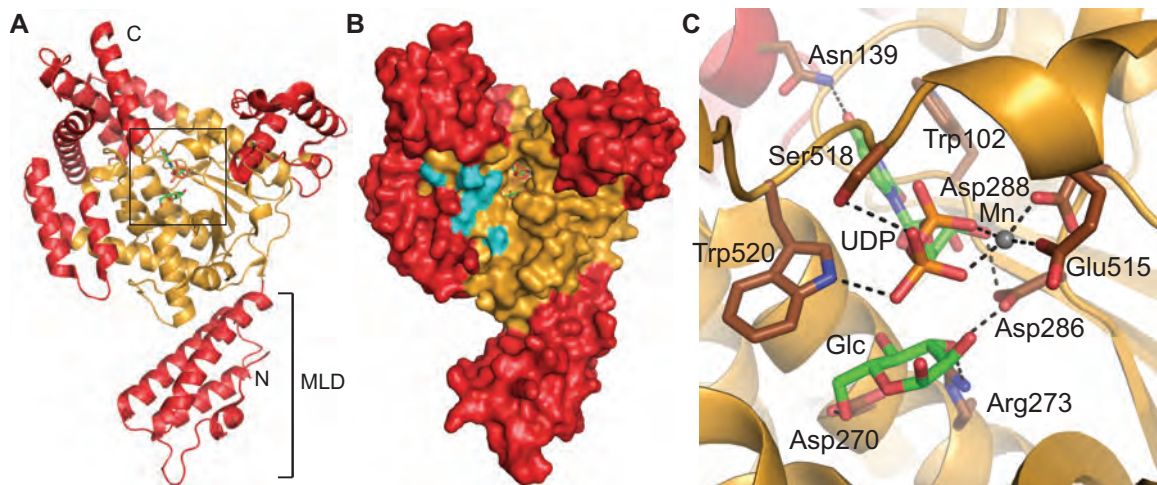
Figure 1-7. Structure of the TcdB glucosyltransferase domain. (A) The core GT-A fold is shown in orange, and the α-helical additions are shown in red. The N-terminal four-helix bundle makes up the MLD. (B) Surface view of the TcdA-GTD as shown in (A). Glu449, Arg455, Asp461, Lys463, and Glu472, residues that have been shown to be involved in GTPase binding, are colored cyan. (C) Close up view of the catalytic core. UDP, glucose, and a manganese ion are bound on the surface of the core GT-A fold. The binding pocket is overlayed by a loop comprising residues 517-523. Some of the residues involved in coordinating UDP, Glucose, and Mn^{2+} are shown as sticks.

Much less is known about the glucosyltransferase activity of TcdA. In one of the few studies comparing the glucosyltransferase activities of TcdA and TcdB, Chaves-Olarte, *et al.* showed that TcdB was a much more active enzyme than TcdA (84). TcdB

was ~100 fold more potent at modifying RhoA, Rac1, and Cdc42. TcdB also had a 6.5 fold higher rate of UDP-glucose hydrolysis in the absence of substrate. TcdA only outperformed TcdB in the modification of Rap2A, but even against this target, TcdA was a poor enzyme. When the toxins were microinjected into cells, TcdB had a 100 fold higher cytopathic potency than TcdA, correlating with the ~100 fold difference in glucosyltransferase activity. The authors concluded that this difference in glucosyltransferase activity was the major determinant in the 1000 fold difference in cytopathicity that had been reported previously.

Many questions remain regarding the enzymatic activities of TcdA and TcdB such as why TcdA has a lower activity than TcdB and why it is able to modify additional substrates. The structural determinants of substrate binding and substrate specificity have not been worked out for any of the LCT/GTPase interactions. Nothing is known about the molecular mechanism of membrane localization for either TcdA or TcdB. Localization of the GTD to certain parts of the cell membrane and the potential presence of unknown cellular cofactors could significantly impact the actual array of *in vivo* targets. Moreover, the substrates of TcdA and TcdB have only been described from *in vitro* reactions. What substrates are actually being glucosylated in cells? In Chapter IV of this work, I describe the structure of the TcdA GTD and my efforts to begin addressing some of these questions.

Research objectives

Crystal structures of a fragment of the TcdA RBD and the TcdB GTD have revealed a great deal about the molecular details of receptor binding and enzymatic inactivation of the target proteins. However, these two structures represent only a small part of the toxins. When we initiated this study, there was no structural information about

how the domains relate to each other or about the central 1300 amino acid region between the GTD and RBD for either toxin. This central region, which makes up about half of the toxin mass, contains the pore-forming and autoprotease domains responsible for the delivery of the GTD into the target cells. Furthermore, important differences have been noted in the glucosyltransferase and receptor-binding activities of TcdA and TcdB, highlighting the need for complementary structures of the TcdB RBD and the TcdA GTD.

In addition, I have determined high-resolution structures of the TcdA cysteine protease and glucosyltransferase domains. Chapter II describes the 1.6 Å X-ray crystal structure of the TcdA CPD in complex with the activating molecule InsP6. Complementary functional studies demonstrate an intra-molecular mechanism of cleavage and highlight specific residues required for InsP6-induced TcdA processing. In Chapter IV, I present the structure of the TcdA GTD. This structure and accompanying biochemical data show that TcdA is not a defective enzyme but has comparable activity to TcdB when modifying Rho proteins. I confirm that TcdA does modify Rap proteins both *in vitro* and in cells whereas TcdB does not. Finally, I show that in the context of the holotoxin structure, the glucosyltransferase activity of TcdA and TcdB is inhibited. These structures provide a framework for understanding the molecular mechanism of cellular intoxication, not only for TcdA and TcdB, but for all members of the LCT family.

CHAPTER II

STRUCTURE-FUNCTION ANALYSIS OF AUTOPROCESSING IN *CLOSTRIDIUM DIFFCILE* TOXIN A

Introduction

Most AB toxins release their enzymatic A component into the cytosol of the target cell. The botulinum neurotoxin, for example, contains a zinc metalloprotease that is tethered to the rest of the toxin by a disulfide bond. Upon being delivered into the target cell, the disulfide bond is broken in the reducing environment of the cytosol, and the protease is liberated. In the case of anthrax toxin, the A components, edema factor and lethal factor, are expressed separately from the B component. After being delivered across the membrane, the unattached A components are free to enter the cytosol. Unlike other known AB toxins, the A component (GTD) of LCTs is released from the B subunits by an autoproteolytic cleavage event (134). The autoproteolytic release of the GTD is a recently described property of LCTs and represents a novel strategy employed by AB toxins to discharge their payload.

Cleavage is triggered by host inositolphosphates such as InsP6 and the reducing environment of the cytosol (134). Autoproteolysis was initially attributed to an aspartate protease activity near the C-terminal RBD, the active aspartate being Asp1665. However, shortly after we initiated our studies on TcdA and TcdB, Egerer, *et al.* reported that cleavage was mediated by a CPD domain located within the N-terminal region of the B subunit (135). The putative autoprotease domain was identified based on homology with the CPD found in the MARTX toxins from gram-negative bacteria (136). Based on sequence homology between MARTX toxins and LCTs, the CPD was predicted to

comprise amino acids 543-769 of TcdB (135). Egerer, *et al.* identified three conserved residues within the CPD (Asp587, His653, and Cys698) that when mutated abrogate the autoprotease activity of TcdB. A subsequent study demonstrated that a protein containing TcdB residues 1-955 is sufficient for autoprocessing, showing that the CPD, not the reported aspartate protease activity, is responsible for InsP6-inducible cleavage.

Autoprocessing in the MARTX toxin from *Vibrio cholera* (VcRTx) is also stimulated by InsP6 (137). A recent crystal structure of VcRTx CPD bound to InsP6 suggests a novel mechanism of InsP6-induced allosteric activation (138). The CPDs of TcdA and VcRTx share only 19% sequence identity. To gain insight into the mechanistic commonalities between these entirely different toxins and to delineate the LCT-specific modes of InsP6-induced processing, we performed structural and functional analyses on the cysteine protease from TcdA.

Methods

Plasmid construction and point mutants

The nucleotide sequence coding for amino acids (aa) 543-809 of TcdA (TcdA CPD) was amplified from *C. difficile* strain 10463 genomic DNA. The DNA was cloned into a modified pET27 vector such that the resulting protein contains an N-terminal His_{10} tag followed by a 3C protease cleavage site. The sequence preceding the TcdA CPD sequence is MGSSHHHHHHHHHHGSSLEVLFQGPGS. Following 3C cleavage, only non-native residues GPGS remain. The extended construct, TcdA g-CPD, encodes aa 510-809. This plasmid was constructed similarly except the last two residues of the leader sequence are VD rather than GS. Point mutants were generated by QuikChange site-directed mutagenesis.

Protein Expression and Purification

Transformed *E. coli* BL21(DE3) cells were grown in terrific broth containing 50 mg/L kanamycin. 10 mL of overnight culture was used to inoculate 1 L of media and the cultures were placed at 37°C and 230 rpm. When the cultures reached OD_{600} = 0.6, the temperature was changed to 16°C and expression was induced by the addition of 0.5 mM IPTG. After 16 h the cells were harvested by centrifugation and resuspended in 100 mM NaCl, 20 mM Tris, pH 8.0. Following French Press lysis, the lysates were centrifuged at 48,000 *g* for 20 min. Protein was purified from the supernatant by Ni-affinity chromatography. For crystallization trials, the His_{10} tag was cleaved from TcdA CPD by 3C protease overnight at 4°C. The mixture was then run back over a Ni-nitrilotriacetic acid agarose column to remove the cleaved His_{10}-tag and the His-tagged 3C protease. Cleaved TcdA CPD was further purified by gel filtration chromatography in 50 mM NaCl, 20 mM Tris, pH 8.0. Selenomethionine substituted TcdA CPD was prepared using *E. coli* BL843(DE3) in minimal media containing 40 mg/L L-selenomethionine following the same procedure, except 5 mM methionine and 1 mM DTT were added to the buffers to prevent oxidation of the selenium. TcdA g-CPD and mutants used for cleavage assays were expressed as above and purified by Ni-affinity and ion exchange chromatography. The His_{10}-tag was not cleaved for these proteins.

Crystallization

TcdA CPD was concentrated to 60-80 mg/mL in 50 mM NaCl, 20 mM Tris, pH 8.0. InsP6 (Sigma) was added to 10 mM. TcdA CPD was crystallized by the hanging-drop method at 21°C with a 1:1 ratio of protein and mother liquor containing 100 mM Tris, pH 8.0-9.0, 25-30% PEG 8000, and 200 mM guanadinium chloride. Crystals were mounted on cryo loops and flash cooled in liquid nitrogen. No cryoprotectants were required.

Structure Determination and Refinement

X-ray data were collected from single crystals on LS-CAT beamlines 21-ID-D and 21-ID-G at the Advanced Photon Source (Argonne, IL). Diffraction data were indexed, integrated, and scaled using HKL2000 (155). Phases were determined from anomalous scattering data using autoSHARP (156). The model was built using Coot (157) and refined using Phenix (158) with riding hydrogens and seven TLS groups per chain (Table 2-1). The final model contains 2 CPD molecules (consisting of residues 547-570 and 575-801 for chain A and 548-570, 578-655, and 661-803 for chain B), 2 InsP6 molecules, and 649 water molecules.

InsP6 Induced Cleavage Assays

TcdA g-CPD and g-CPD mutant proteins were exchanged into 60 mM NaCl, 250 mM sucrose, 20 mM Tris, pH 7.5. 100 µL protein at 5 µM was mixed with 1 µL buffer or InsP6 stock solution and incubated at 37°C. After 2 h, the reaction was stopped by the addition of loading buffer and heating. For the time course reactions, 1 mL 5 µM protein was treated with 10 µL 500 µM InsP6. At the indicated time-points, 50 µL aliquots were removed; SDS loading buffer was added; and the samples were heated for 5 min. The samples were analyzed by SDS-PAGE with Coomassie staining.

NMR spectroscopy

TcdA CPD was expressed and purified as described for crystallography, except that for 2D NMR, TcdA CPD was expressed in M9 minimal media with $^{15}NH_4Cl$ as the only nitrogen source. The protein was exchanged into a 50 mM NaCl, 50 mM potassium phosphate, pH 7.0 buffer by gel filtration chromatography and concentrated to 0.6 mM +/- 5 mM InsP6. NMR data were recorded at 25°C on a Bruker DRX600 spectrometer

with a 5 mm TXI-Z cryoprobe and processed with Topspin 2.0b (Bruker). 2D images were prepared with Sparky.

Table 2-1. X-ray data collection and refinement statistics for the crystal structure of the TcdA CPD.

Data Collection	Native	Selenium
Wavelength, Å	0.9784	0.9784
Resolution (outer shell)[1], Å	50-1.6 (1.66-1.60)	50-1.8 (1.86-1.80)
R_{merge}[2], %	4.3 (26.8)	5.5 (26.6)
Mean $I/\sigma I$	23.5 (2.4)	24.9 (6.7)
Completeness, %	96.2 (81.9)	99.6 (96.5)
Redundancy	4.8 (2.8)	7.6 (6.9)
Unique observations	68,137 (5,753)	47,174 (4,525)
Anomal. phasing power		1.226
Refinement		
R_{cryst}/R_{free}[3], %	16.54/19.67	
No. protein atoms	4007	
No. ligand atoms	84	
No. solvent waters	649	
Bond length rmsd, Å	0.011	
Bond angle rmsd, °	1.469	
Avg. protein B, Å^2	39	
Ramachandran plot[4], %		
Most favored	417 (92.3%)	
Allowed	33 (7.3%)	
Generously allowed	1 (0.2%)	
Disallowed	1 (0.2%)	

[1] Outer resolution bin statistics are given in parentheses.

[2] $R_{merge} = S_{hkl}(S_i|I_{hkl,i} - <I_{hkl}>))/S_{hkl,i}<I_{hkli}>$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices h, k and l, and $<I_{hkl}>$ is the mean intensity of that reflection.

[3] $R_{cryst} = S||F_{obs, hkl}| - |F_{calc, hkl}||/|F_{obs, hkl}|$, where $|F_{obs, hkl}|$ and $|F_{calc, hkl}|$ are the observed and calculated structure factor amplitudes. R_{free} is equivalent to R_{cryst} but calculated with reflections (5%) omitted from the refinement process.

[4] Calculated with PROCHECK (159).

Results

Defining the TcdA cysteine protease domain

TcdA and TcdB holotoxins undergo auto-proteolysis in the presence of InsP6 and DTT to release an N-terminal glucosyltransferase domain (residues 1-542 for TcdA) (133). In TcdB, the autoproteolytic activity and InsP6 binding have been mapped to fragments corresponding to residues 1-955 and residues 544-955, respectively (139). Sequence alignment with VcRTx suggests domain boundaries of 543-769 (TcdA) and 543-767 (TcdB) (135). To define a structural domain required for TcdA auto-proteolysis, we designed a construct for the recombinant expression of TcdA 510-809. The 33 residues from the glucosyltransferase domain C-terminus were included to provide a substrate for the enzyme and to permit visualization of cleavage by SDS-PAGE (Figure 2-1A). Incubation of this 37 kDa protein, g-CPD, with InsP6 causes proteolysis resulting in two cleavage products of 30 kDa (CPD) and 7 kDa (g). When InsP6 and TcdA g-CPD were mixed in equal molar ratios, proteolysis occurred quickly with 50% of the protein cleaved in about 10 minutes (Figure 2-1B). A shorter construct corresponding to TcdA residues 510-769 did not undergo cleavage, even when InsP6 was added in excess (data not shown).

Structure of the TcdA cysteine protease domain

TcdA amino acids 543-809, henceforth TcdA CPD, was crystallized in the presence of InsP6 and the structure was determined at 1.6 Å (Table 2-1, pdb 3HO6). The asymmetric unit contained two CPD molecules, which align with an rmsd of 0.236 Å based on the backbone α-carbon positions. For the remainder of this chapter we refer to chain A. The TcdA CPD is composed of a nine stranded β-sheet flanked by 5 α-helices (Figure 2-2). The closest structural homolog is that of the VcRTx CPD (pdb 3EEB, (138))

which aligns with an rmsd of 2.9 Å over 187 residues (Figure 2-3), while the next closest homolog is caspase-7 (pdb 1SHJ, (160)), aligning with an rmsd of 4.8 Å over 135 residues.
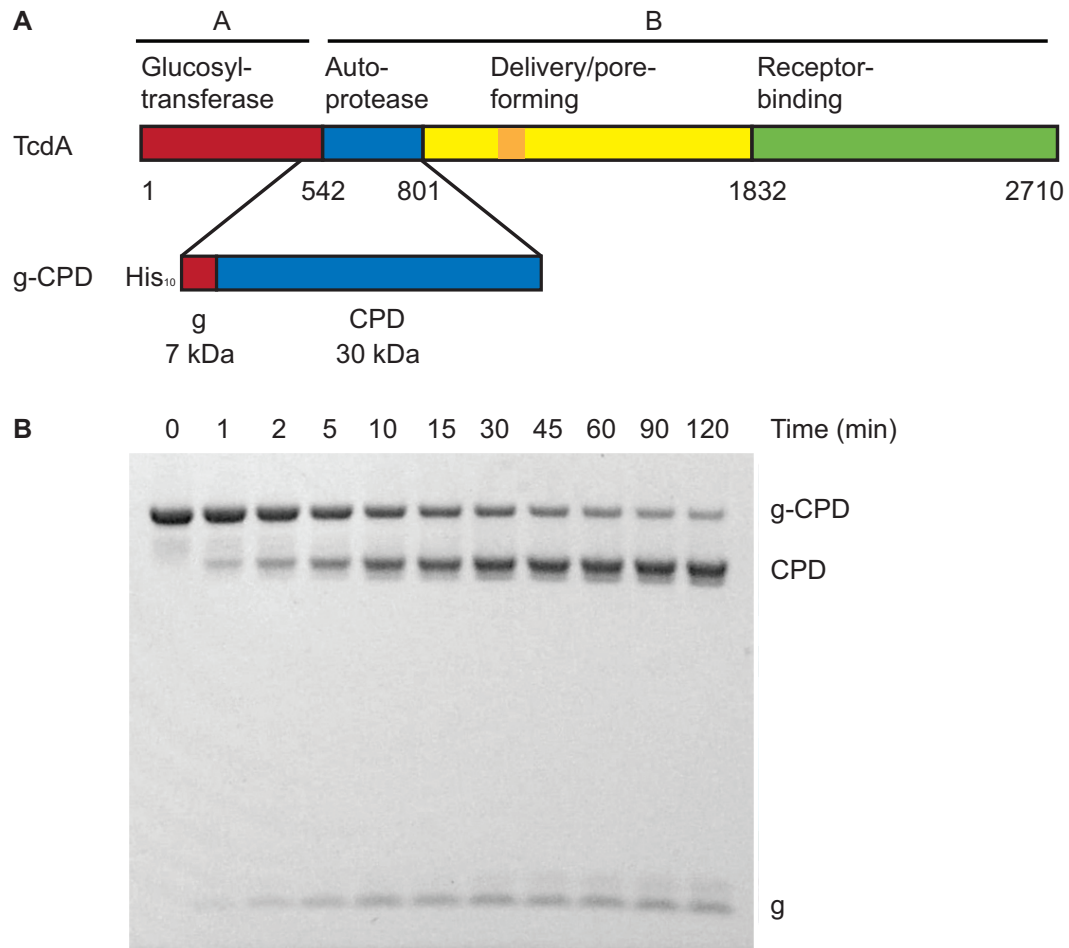
Figure 2-1. Defining the TcdA CPD. (A) TcdA is an AB toxin with four functional domains. The CPD is defined in this work as residues 543-809 based on the InsP6-induced activity of g-CPD, residues 510-809. (B) Time course of g-CPD (5 µM) cleavage in the presence of 5 µM InsP6 shows 50% cleavage within 10 minutes.
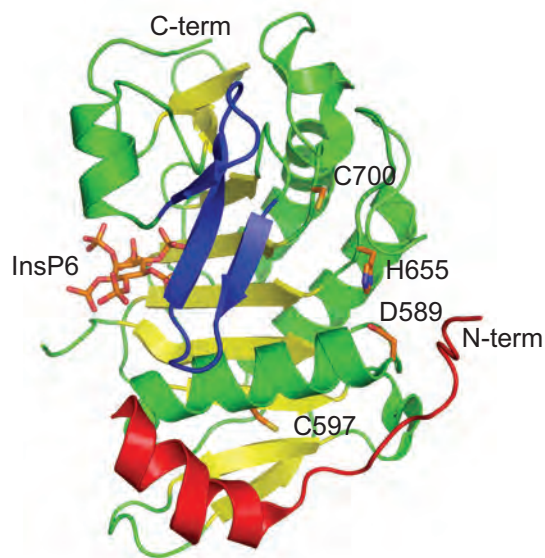
Figure 2-2. Structure of the TcdA CPD bound to InsP6. TcdA CPD is shown as a ribbon diagram with InsP6 and the side chains of Cys700, His655, Asp589, and Cys597 shown as sticks. The InsP6 binding site and catalytic site are separated by a three-stranded β-flap (blue). The central β-sheet is shown in yellow. The N-terminus (red) wraps around the domain with the most N-terminal residues near the catalytic site.

In both the VcRTx and TcdA CPD structures, InsP6 binds on one face of the β-sheet and is separated from the proposed active site (Cys700, His655, and Asp589 in TcdA) by a three-stranded β-hairpin structure termed the β-flap (Figure 2-2, 2-4). The N-terminus of the protein wraps around the exterior of the domain with the most N-terminal of the resolved residues (Gly547) near the proposed catalytic site (Figure 2-2). The TcdA CPD contains about fifty more residues than the VcRTx CPD, which leads to additional β-strands at the top and bottom of the β-sheet and two additional α-helices (Figure 2-3, 2-4). The two cysteines of the TcdA CPD domain (Cys597 and Cys700) are more than 20 Å apart, consistent with the observation that reductants are not needed for InsP6-induced auto-processing of g-CPD (Figure 2-1). The fact that reducing agents enhance the auto-processing of holotoxin (135) suggests the presence of a disulfide involving at least one cysteine from another domain.
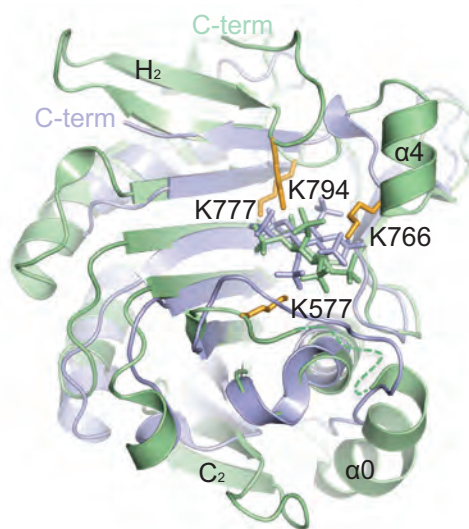
Figure 2-3. Overlayed comparison of the TcdA and VcRTx CPD structures. The TcdA (green) and VcRTx (blue) CPDs are shown as overlayed ribbon diagrams with InsP6 shown as sticks. Lysine residues 577, 766, 777, and 794 from TcdA CPD are shown as orange sticks. Secondary structure elements that are present in the TcdA CPD but not the VcRTx CPD are labeled.
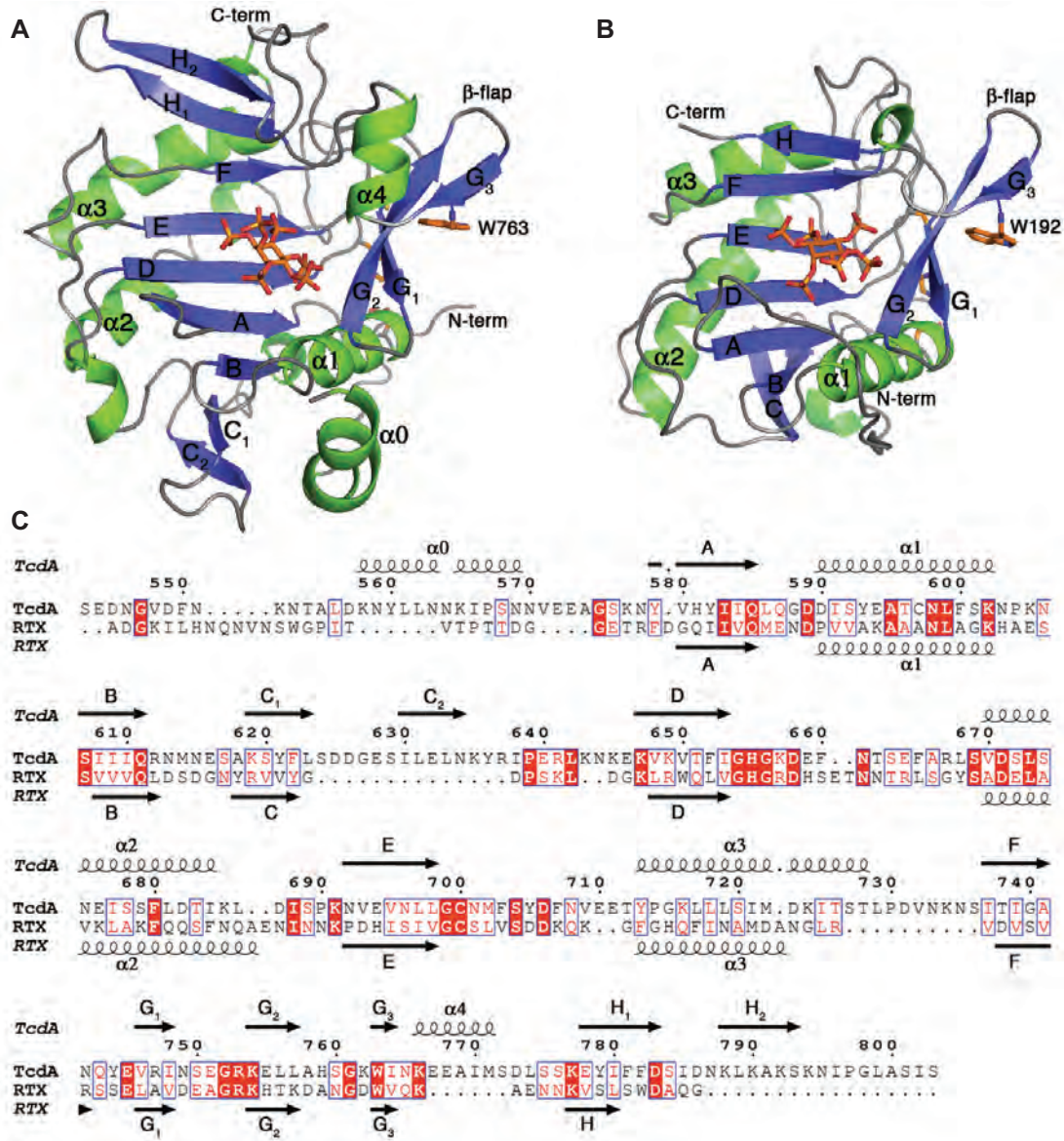
Figure 2-4. Comparison of the TcdA and VcRTx CPD structures and sequences.
(A-B) The structures of the TcdA (A) and VcRTx (B) CPDs are shown in the same orientation with the secondary structural elements annotated and color coded (α-helices in green; β-strands in blue). We have labeled the secondary structural elements to maintain consistency with the VcRTx CPD and caspase-7 literature. (C) The TcdA and VcRTx CPD sequences were aligned based on their structures (161) and displayed using ESPript (162). Strictly conserved residues are shown in white letters with a red background. Residues that are similar are boxed in red letters with white backgrounds. The secondary structural elements are marked above and below the sequences and the numbering corresponds to that of TcdA.

TcdA CPD active site

In cysteine proteases, the nucleophilicity of the active site cysteine is typically enhanced by hydrogen bonding to an adjacent histidine. Often, an additional negatively charged residue, such as aspartic acid, is present to stabilize the positive charge of the histidine. TcdA Cys700, His655, and Asp589 align to residues that have been implicated as catalytic residues in TcdB CPD (135,139) and VcRTx CPD (137). To address the role of these residues in autoproteolysis of TcdA, each was mutated within g-CPD and the proteins were tested for InsP6-inducible cleavage. Mutation of Cys700, His655, or Asp589 abrogates InsP6-inducible cleavage (Figure 2-5A). The adjacent Asp590 cannot substitute for Asp589, and mutation of Asp590 does not cause a defect in protease activity (Figure 2-5A). While these results are consistent with a catalytic triad, the structure indicates that while His655 and Asp589 are in hydrogen bonding contact, Cys700 is > 6 Å from His655 and >10 Å from Asp589 (Figure 2-5A). This arrangement, which we also observe in the CPD structure of VcRTx, suggests that the mechanism of catalysis in this class of self-cleaving proteases may be different from one in which a catalytic triad is in place.

Intramolecular cleavage by CPD

To test whether InsP6-induced cleavage in TcdA is inter- or intramolecular, we mutated the conserved P1 residue from leucine to alanine (L542A) in g-CPD. This mutation prevented InsP6-inducible cleavage at the normal cleavage site (Figure 2-5C). When this mutant was mixed with the catalytic cysteine mutant, g-CPD C700S, no cleavage was observed. The g-CPD L542A was unable to cleave the normal cleavage site of g-CPD C700S suggesting that intermolecular cleavage does not occur for TcdA. In the crystal structure of TcdA CPD, residues 543-546 are unstructured. In TcdA CPD containing an intact cleavage site, the unstructured residues 543-546 likely extend into

41

the catalytic site. The structure of the TcdA CPD suggests that Leu542 may bind in a hydrophobic pocket located between the catalytic cysteine and the β-flap (Figure 2-5B,D). This pocket is made up of the residues Ile591, Ala 595, Ile653, Leu698, Val746, Ile748, and Trp763 (shown in green in Figure 2-5), which are mostly conserved among LCTs (Figure 2-6).
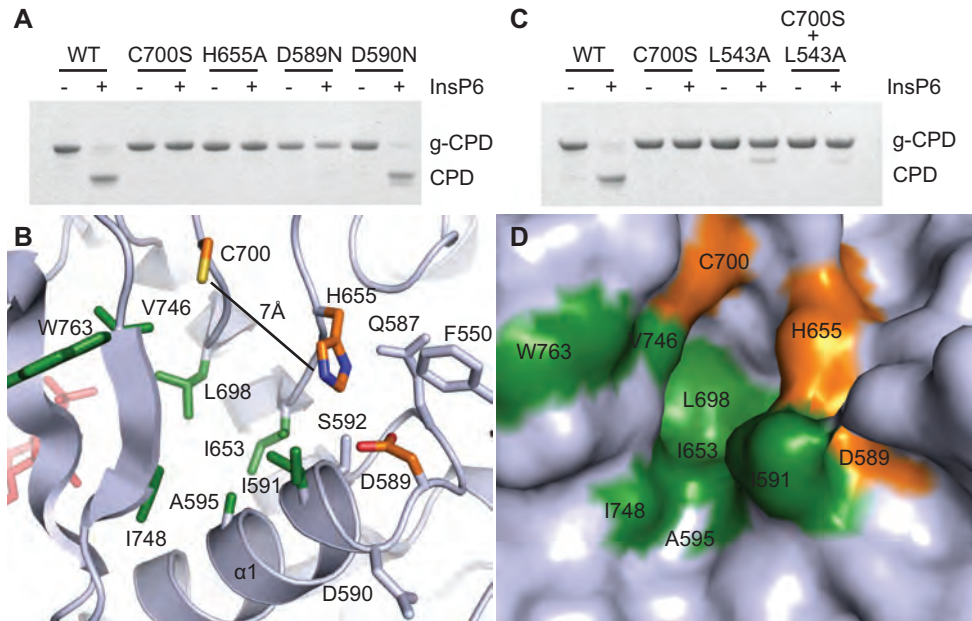
Figure 2-5. Catalytic site of the TcdA CPD. (A) Wild-type, C700S, H655A, D589N, and D590N g-CPD at 5 μM were incubated with 100 μM InsP6 for 2 h at 37°C. In contrast to wild-type and D590N, the C700S, H655A, and D589N mutants were unable to undergo InsP6-inducible cleavage. (B) Ribbon diagram of the catalytic site of CPD (gray) with residue side chains of the catalytic site (orange) and hydrophobic pocket (green) shown as sticks. Other side chains within the active site are colored gray. (C) L542A g-CPD (cleavage site mutant) does not undergo cleavage upon addition of InsP6. When C700S and L542A g-CPD are mixed, no InsP6-dependent cleavage is observed. (D) Surface representation of the catalytic site reveals a likely substrate-binding pocket. The D589, H655, and C700 residues are colored orange and the hydrophobic residues are colored green. The orientation is identical to that in (B).
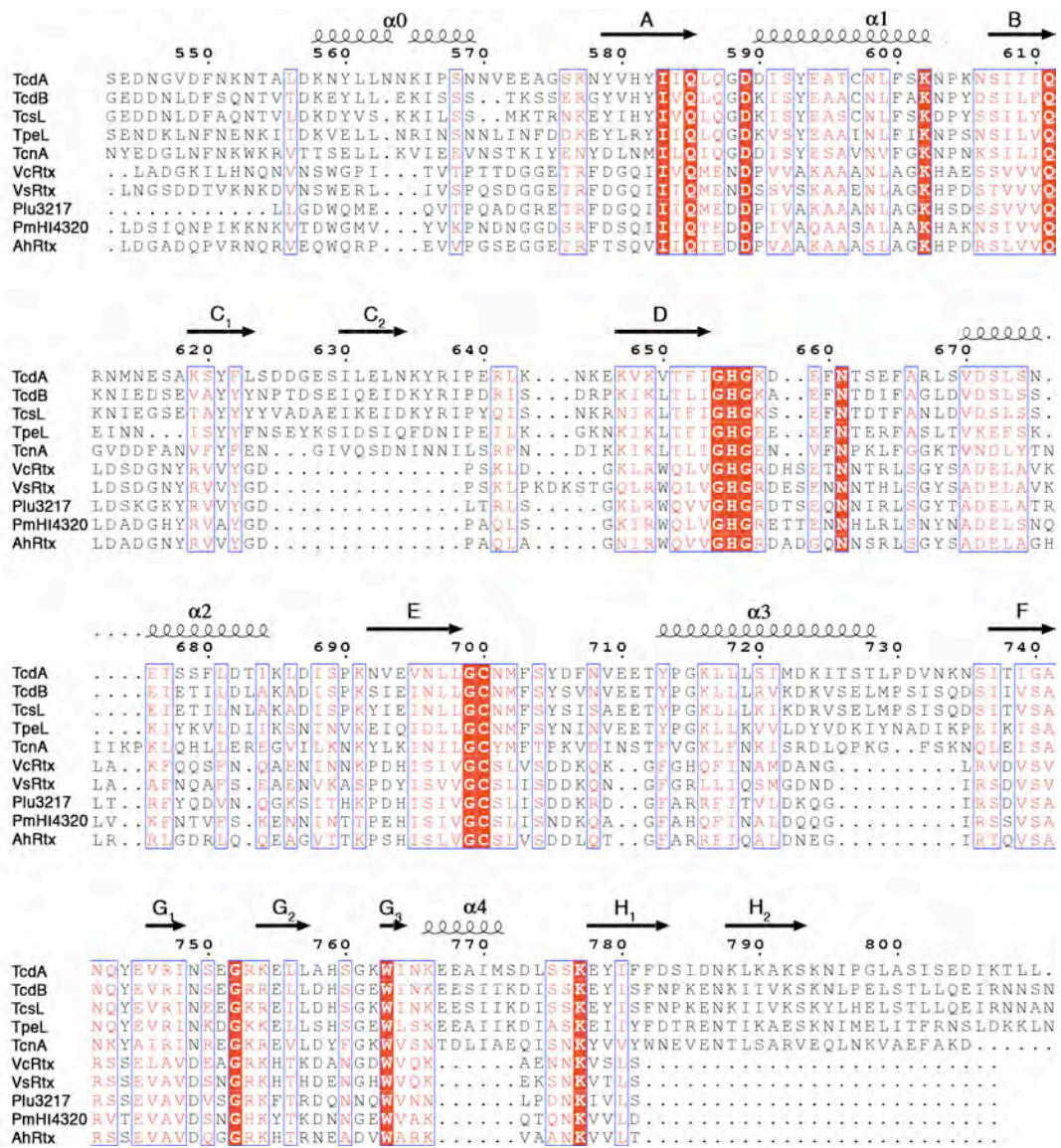
Figure 2-6. Conserved residues of the CPD. The TcdA CPD sequence (YP001087137) was aligned to sequences from four LCTs (TcdB (CAC19891), TcsL (CAA57959), TpeL (ACF49258), Tcnα (CAA88565)) and five diverse homologs found in gram-negative bacteria (VcRTx (ZP01975348), VsRTx (ZP00989505), Plu3217 (NP930444), PmHI4320 (YP002151764), AhRTx (YP855898)) using ClustalW (163). There are 14 strictly conserved residues shown in white letters with a red background.

Binding of InsP6

The InsP6 molecule is highly charged, bearing six negatively charged phosphate groups, and accordingly, binds within a positively charged pocket of the TcdA CPD (Figure 2-7). Binding of InsP6 results in a significant stabilization of TcdA CPD as reflected by an increased resistance to digestion by chymotrypsin (Figure 2-8). Efforts to obtain crystals of the TcdA CPD in the absence of InsP6 were unsuccessful, so NMR spectroscopy was used to assess the nature of InsP6-induced structural changes. The 1D ^{1}H NMR spectrum of apo TcdA shows excellent dispersion of the signals indicative of a well-folded globular domain (Figure 2-9A, lower trace). Addition of InsP6 into the sample resulted in a substantial perturbation of the signals with an even greater dispersion (Figure 2-9A, upper trace). This is most evident in the very high field methyl region (-0.4 – 0.4 ppm) and the very low field amide region (8.4-10.4 ppm, inset) of the spectrum. 2D ^{1}H-^{15}N NMR spectra of the CPD also reveal a larger number of shifted peaks in the presence InsP6 (Figure 2-9B). These data indicate that the TcdA CPD is folded in the absence of InsP6 and imply that InsP6 binding induces significant conformational reorganization and structural stability within TcdA CPD.
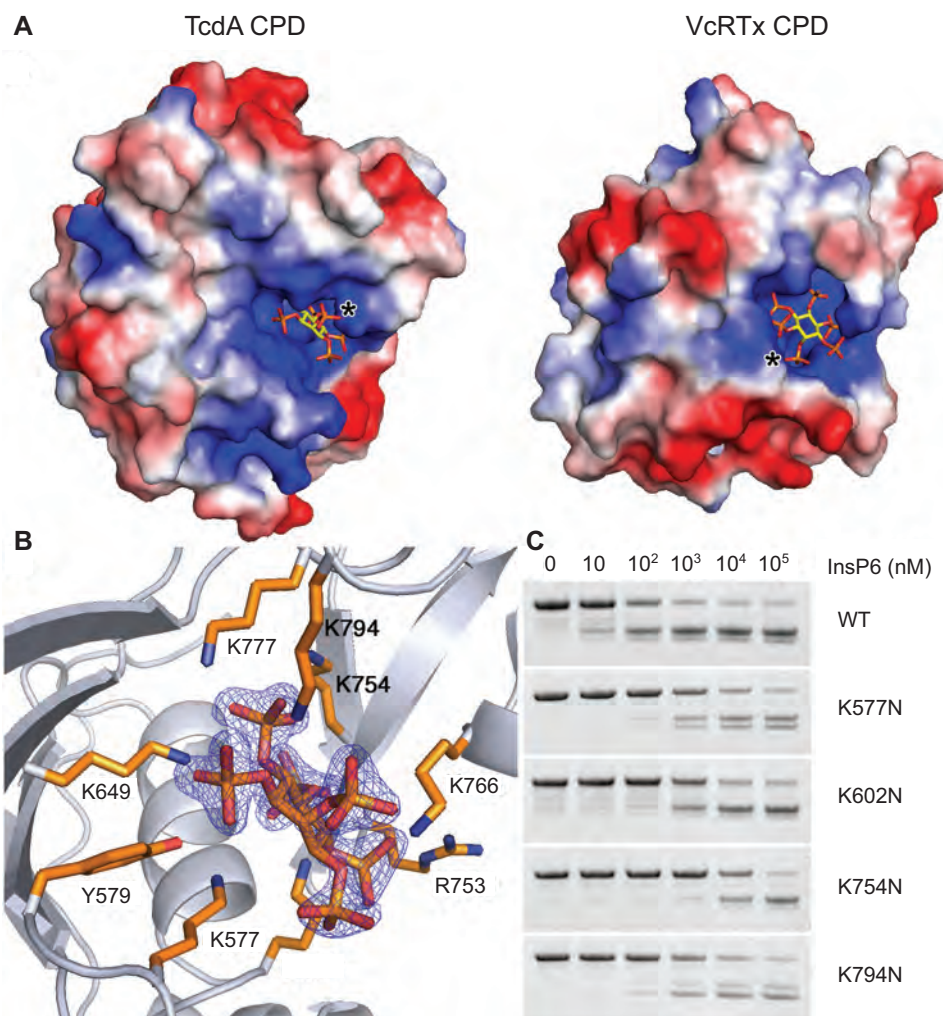
Figure 2-7. InsP6 binding pocket. (A) The TcdA CPD (left) and VcRTx CPD (right) surfaces colored by electrostatic potential (negatively charged surfaces are red; positively charged regions are blue, and neutral regions are shown in white) demonstrate that InsP6, shown as a stick model, binds in positively charged pockets. The protein molecules were aligned and are shown in identical views but the orientation of the InsP6 molecules and the position of the axial phosphate groups (denoted with an asterisk) differs. (B) InsP6 and the side chains of residues that directly interact with InsP6 are shown as sticks. The orientation is identical to that in (A). The displayed electron density is a $2F_o$-F_c map contoured at 5σ in which InsP6 was omitted from the phase calculations. (C) WT, K577N, K602N, K754N, and K794N g-CPD at 5μM were tested for autoproteolysis over a range of InsP6 concentrations. Each mutant required 10-1000 fold more InsP6 in order to undergo autoproteolysis at levels comparable to wild-type.
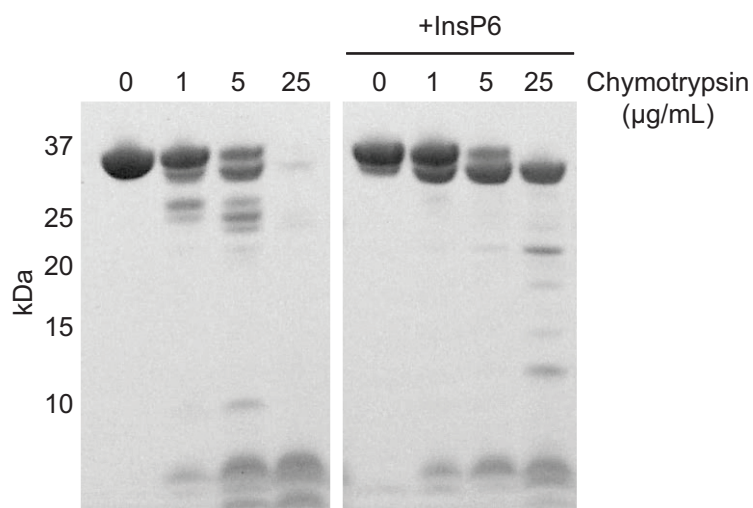
Figure 2-8. Effect of InsP6 on the structural stability to the TcdA CPD. TcdA CPD (at 1 mg/mL in 50 mM NaCl, 20 mM Tris, pH8.0) was incubated with chymotrypsin for 1 hour at 25°C. The presence of 1 mM InsP6 provides protection from proteolysis. Similar results were obtained with trypsin (data not shown).
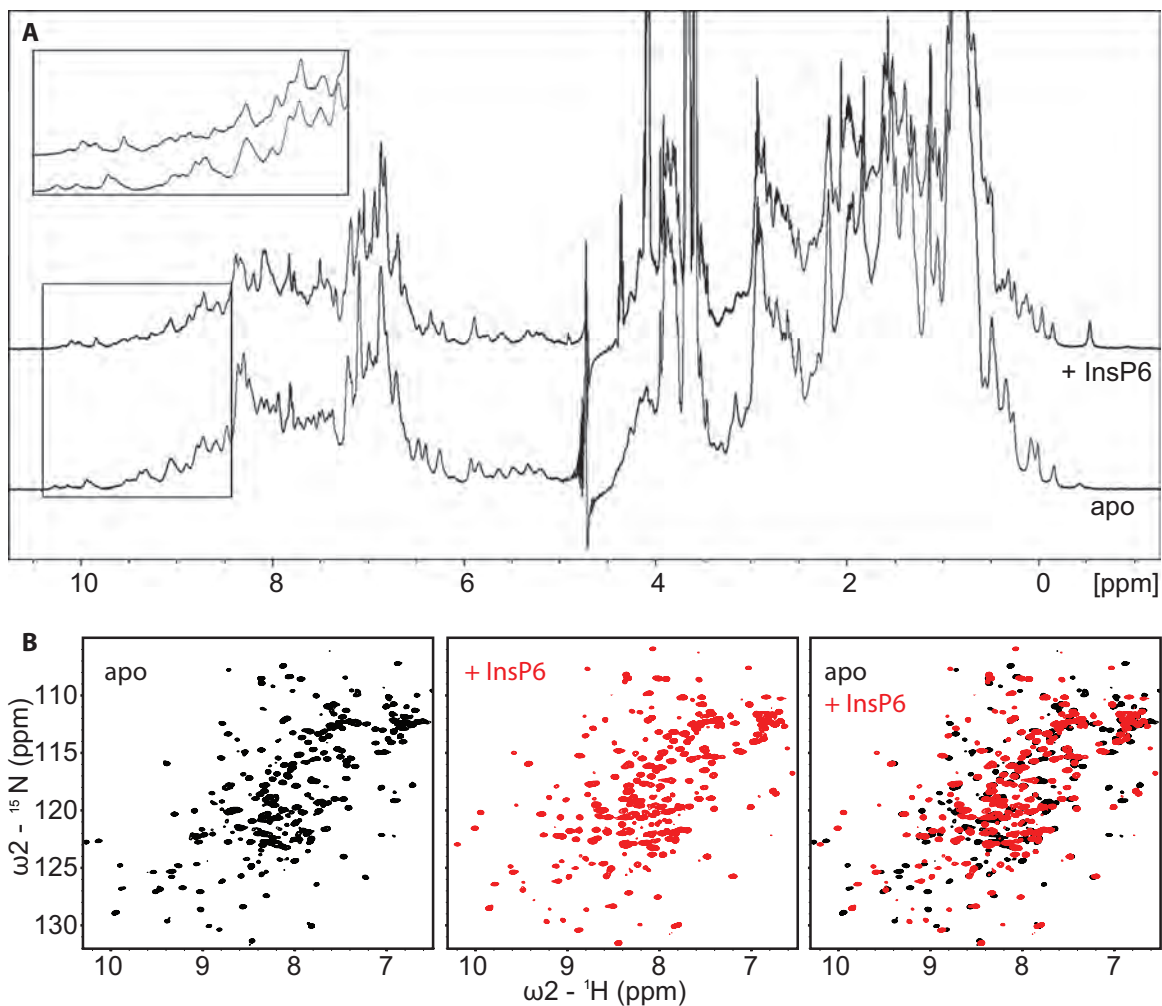
Figure 2-9. InsP6-induced structural changes. (A) 1D ^{1}H-NMR spectra of the TcdA CPD in the presence (upper trace) and absence (lower trace) of InsP6 show clear dispersion of peaks indicative of folded proteins. The inset shows major changes in the peaks suggesting that significant structural changes are induced by InsP6 binding. (B) An ^{15}N-^{1}H spectrum in the absence of InsP6 shows a clear dispersion of peaks indicative of a folded protein. More peaks are visible in the presence of InsP6. An overlay suggests that significant structural changes are induced by InsP6 binding.

Insights into the effects of InsP6 binding on TcdA CPD are provided by the high resolution structure of the complex. The TcdA CPD has nine residues that make direct contacts with the InsP6 (Figure 2-7B). These include one arginine (Arg753), one tyrosine (Tyr579), and seven lysines (Lys577, 602, 649, 754, 766, 777, and 794). Thus, binding of InsP6 involves residues spanning the entire domain, consistent with the significant reorganization of the structure upon binding. Lys794, which is conserved among all LCTs except TpeL of *C. perfringens* (Figure 2-6), coordinates the axial phosphate group of the InsP6. This residue is located on the C-terminal β-strand, which is entirely absent from the VcRTx CPD. Thus, the position of the axial phosphate group and the orientation of the InsP6 in the VcRTx CPD structure are markedly different from that in TcdA CPD (Figure 2-7).

In VcRTx CPD, Lys602, 754, 766, 777, and Arg753 are conserved and important for binding InsP6 and proteolysis (Figure 2-4) (138). To test the importance of both conserved and unique InsP6-binding interactions in TcdA autoproteolysis, lysine residues 577, 602, 754, and 794 were mutated to asparagine in the g-CPD protein background and tested for InsP6 inducible cleavage (Figure 2-7C). The mutant proteins required 10-1,000 times higher concentrations of InsP6 to achieve an equivalent amount of cleavage as wild type g-CPD, confirming the significance of each of these InsP6 ligands.

The β-flap

In the structural analysis of VcRTx, Lupardus *et al.* propose the β-flap as a link between InsP6 binding and the catalytic site (138). In TcdA, the β-flap contains a network of amino acid side chains that may be involved in transmitting InsP6-induced structural changes to the active site (Figure 2-10). The active site side of the β-flap includes Val746, Ile748 and Trp763, residues that line the hydrophobic pocket thought to

be involved in binding the P1 residue of the substrate (Leu542) (Figure 2-5B,D). On the InsP6 side of the β-flap, the conserved Arg753-Lys754 pair and Lys766 directly coordinate the InsP6 (Figure 2-7B, 2-10). Lys766 is located at the beginning of α4, a structural element not found in the structure of VcRTx CPD. In addition to interacting with InsP6, the α4 helix stabilizes the β-flap with an electrostatic interaction between Glu768 and Lys762 (Figure 2-10). Arg747 also stabilizes the β-flap structure. This is through an electrostatic interaction with Glu755 and a π-cation interaction with Trp763 (Figure 2-10). While Trp763 is strictly conserved in the broad family of CPD sequences, the Arg747 and Glu757 residues are only conserved within the LCTs (Figure 2-6). In VcRTx, and other homologs from gram-negative bacteria, the residue corresponding to Arg747 is an alanine. Efforts to analyze a TcdA CPD R747A mutant failed due to protein instability, implying a role for Arg747 in the absence of InsP6. Thus, in TcdA (and likely all LCTs) the β-flap is an important structural element both in separating the InsP6-binding and substrate-binding sites and in contributing to the global stability of the fold.

Discussion

TcdA and VcRTx represent two classes of large toxins that target the eukaryotic cytoskeleton through very different mechanisms. The proteins contain a single region of homology, the CPD, which is involved in proteolytic processing of the toxins. Comparison of the atomic structures from these highly divergent domains provides an opportunity to gain insight into the selective pressures needed for function in both systems.
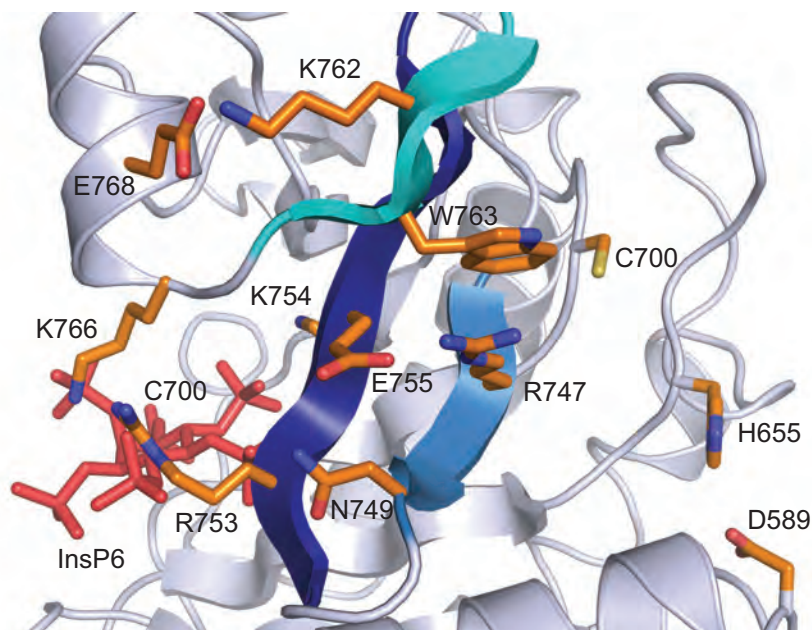
Figure 2-10. TcdA CPD β-flap. A close up of the β-flap (blue) reveals a network of interactions that link InsP6 (red) to the catalytic site (Cys700, His655 and Asp589).

The CPD from TcdA is larger than that of VcRTx and contains a C-terminal extension that was not readily apparent through sequence analysis. The observation that Lys777 and Lys794 are involved in coordination of InsP6 provides ready explanation for the lack of activity in the 510-769 construct we made based on sequence alignment with VcRTx. The altered conformation of InsP6 within the active site reflects the contributions from this C-terminal extension as well as other structural differences in the TcdA and VcRTx InsP6 binding sites. Lys766 is presented by the α4 helix, not present in VcRTx, and the loop containing Lys577 is positioned very differently in the two molecules (Figure 2-3). The capacity to evolve differences in how InsP6 (and related molecules) are bound has been noted in the family of RNA editing enzymes (164) and could be relevant in a newly identified and largely uncharacterized family of CPD homologs from bacteria and eukaryotes (165).

The VcRTx and TcdA CPD structures were obtained in the presence of InsP6. The dispersion of peaks observed in the apo- and InsP6-bound proteins by NMR

suggest that both proteins are folded but that InsP6 binding confers stability and induces significant structural change (2-9). This is not unexpected given the energetically unfavorable accumulation of positive charges that one would predict in the absence of InsP6 (Figure 2-7). We propose that the energetic gains of complementing these positive charges with InsP6 allow the molecule to overcome the energetic barrier needed to access its active conformation. An example of such an allosteric transition in a single domain protein has been documented in the phosphorylation of the NtrC response regulator (166). Recent work by Shen *et al.* has demonstrated that the TcdB CPD exists in a dynamic equilibrium between active and inactive states (167). InsP6 binding shifts the equilibrium markedly toward the active conformation. Shen, *et al.* showed by mutational analysis that the β-flap does, in fact, play an important role in the allosteric activation mechanism through coupling the InsP6-binding and catalytic sites.

A trio of aspartate, histidine, and cysteine residues has been shown to be essential in the auto-proteolytic processing of VcRTx (137), TcdB (135), and TcdA (this work). The structures of the TcdA and VcRTx CPDs, however, indicate that these residues do not adopt a conventional catalytic triad arrangement. Instead, the cysteine and histidine residues are separated by a large cavity which is likely to serve as the substrate-binding pocket (Figure 2-5B,D). In TcdA, this pocket is ~11Å long, ~7Å wide, ~7Å deep and bounded by α1, the β-flap, and the loop containing His655. One end of the cavity is hydrophobic and believed to represent a binding site for the P1 leucine residue of the substrate. While His655 is on a flexible loop and could re-orient in the presence of substrate, the structural changes needed to account for the observed importance of Asp589 would be substantial and require rearrangement of α1 and occlusion of the substrate binding cavity (Figure 2-5B).

An inactive VcRTx CPD, with the catalytic cysteine mutated to a serine, has recently been crystallized with the P1 leucine bound in the active site (pdb 3FZY (168)).

51

Even in this structure with the substrate bound, the mutated residue, C6538S, remains separated from the hydrogen bonded His-Asp pair. The substrate chain lies between them. Structures of the TcdB CPD have also been determined recently, and the active site residues are likewise separated (pdb 3PEE and 3PA8 (167,169)).

Thus, despite biochemical evidence consistent with a catalytic triad, the TcdA and VcRTx CPD crystal structures indicate that the histidine and aspartate residues are not positioned to increase the nucleophilicity of the cysteine. Rather, His655 may be acting as an oxyanion hole with Asp589 important in its capacity to orient His655. TcdA and VcRTx CPDs may employ a primitive mechanism of proteolysis wherein the need to "activate" the nucleophile is obviated by the intramolecular and/or single turnover nature of the reaction. Mutant serine proteases in which the histidine and aspartate have been removed from the catalytic triad still catalyze reactions with ~10,000 fold rate enhancements (170), explained largely by their ability to constrain substrate and stabilize the developing oxyanion (171).

In summary, we have shown that the CPD of TcdA is able to use InsP6 to activate an intra-molecular auto-proteolytic processing event. The structural analysis reveals striking similarities in the mechanisms of TcdA and VcRTx InsP6-inducible cleavage but also shows differences in the conformation of InsP6 and the organization of the β-flap. These observations should aid the elucidation of unique and conserved molecular features required for virulence in TcdA, the LCTs and the larger family of CPDs from MARTX toxins.

CHAPTER III

STRUCTURAL ORGANIZATION OF THE FUNCTIONAL DOMAINS OF
CLOSTRIDIUM DIFFICILE TOXINS A AND B

Introduction

The action of LCTs on mammalian target cells depends on a multi-step mechanism of receptor-mediated endocytosis, membrane translocation, autoproteolytic processing, and mono-glucosylation. Many of these functional activities have been ascribed to discrete regions within the primary sequence suggesting that the toxins will adopt multi-modular three-dimensional structures (Figure 1-4). LCTs are important virulence factors, but structural data on these proteins have thus far been limited to crystal structures of a few isolated domains. With the exception of an analysis of TcdB by small angle X-ray scattering (SAXS) (172), the holotoxin structures of these molecules have not been characterized.

We have imaged TcdA and TcdB using electron microscopy (EM) and shown that they share many similar structural features. We have determined a three-dimensional (3D) structure of the TcdA holotoxin at neutral pH by negative stain EM and experimentally mapped three of the four functional domains to discrete regions within the density. These data allow us to evaluate structural models of the TcdA receptor-binding domain (115,124), the TcdA autoprotease domain (173), and the TcdB glucosyltransferase domain (150) within the architecture of the holotoxin. In addition to the analysis at neutral pH, we present images of TcdA after autoprocessing and after exposure to acidic pH. A 3D structure at low pH suggests that the conformational changes required for translocation of the glucosyltransferase domain into the host

cytosol will be significant. Since members of the LCT family are similar in many aspects of sequence and function, these structures of TcdA at neutral and acidic pH provide a framework for understanding the molecular mechanism of cellular intoxication for all members of the LCT family.

Methods

Cloning

DNA corresponding to the TcdA receptor-binding domain (aa 1832-2710), TcdA protease domain (aa 543-795), TcdA delivery domain (aa 799-1859), and TcdA D1-delivery sub-domains (aa 799-1460 and 771-1460) was amplified from *C. difficile* strain 10643 genomic DNA. The TcdA domains were cloned into a modified pET27 vector such that the resulting proteins contains an N-terminal His_{10} tag separated from the protein by a 3C protease cleavage site. After 3C cleavage, only non-native residues GPGS remain. DNA corresponding to the TcdB holotoxin was cloned into a pHis1622 (MoBiTec, BMEG20) vector using the restriction sites BsrGI and KpnI. The nucleotide sequence corresponding to the TcdA GTD (aa 1-542) was cloned into pHis1622 using the BsrGI and SphI restriction sites.

Protein expression and purification

Native TcdA was obtained from the supernatant of *C. difficile* strain 10643 grown in dialysis sac culture, as described previously (3). TcdA was purified from the supernatant by multiple rounds of anion-exchange chromatography, followed by gel-filtration chromatography in 150 mM NaCl, 20 mM Tris, pH 8.0. The identity of TcdA was verified by mass spectrometry. TcdB was expressed in *Bacillus megaterium* as described previously (174) except cells were harvested ~4 h post induction. The TcdA-

GTD was expressed the same way. All of the other recombinant TcdA domains were expressed from *E. coli* as reported previously for the autoprotease domain (173), except BL21(DE3)-CodonPlus cells (Stratagene) were used. Accordingly, 35 mg chloramphenicol was added for every liter of media. At the final step in purification, proteins were exchanged into 50 mM NaCl, 50 mM potassium phosphate, pH 7.0 by gel filtration chromatography.

Autoproteolysis of TcdA

100 µg of purified TcdA in 150 mM NaCl, 20 mM Tris, pH 8.0 was mixed with 100 mM InsP6 and 100 mM DTT and incubated at 37 °C for 3 hours. Some of the protein precipitated during this process. Protein that remained soluble was run on a S200 gel-filtration column in 150 mM NaCl, 20 mM Tris, pH 8.0.

Specimen preparation and electron microscopy

Uranyl formate (0.7% mass/volume) was used for conventional negative staining as described (175). The low-pH TcdA delivery domain and TcdA holotoxin grids were prepared as above, except the grid was washed with 50 mM sodium citrate, pH 4.5 instead of water.

Images of TcdA holotoxin at neutral and low-pH were recorded using a Tecnai T12 electron microscope (FEI) equipped with a LaB_6 filament and operated at an acceleration voltage of 120 kV. Images were taken under low-dose conditions at a magnification of 67,000X using a defocus value of –1.5 mm. Images were recorded on DITABIS digital imaging plates (Pforzheim, Germany). The plates were scanned on a DITABIS micron scanner (Pforzheim, Germany), converted to mixed raster content (mrc) format, and binned by a factor of 2 yielding final images with 4.48 Å/pixel. Images of cleaved TcdA were collected under similar conditions except images were recorded

using a 2048 x 2048 pixel Gatan CCD camera. Cleaved TcdA images were also converted to mrc format, and binned by a factor of 2 resulting in final images with 3.0 Å/pixel.

Images of TcdB and the TcdA delivery domain were collected on a Tecnai F20 electron microscope equipped with a field emission electron source and operated at an acceleration voltage of 120 kV under low-dose conditions at a magnification of 100,000X and a defocus value of -1.5 mm. Images were collected using a Gatan 4Kx4K CCD camera. CCD images were converted to mrc format and binned by a factor of 4 resulting in final images with 4.26 Å/pixel.

Particle images of cleaved TcdA, TcdA delivery domain, and TcdB were selected with Boxer (176) and windowed with a 192 pixel (3.0 Å/pixel), 64 pixel (4.26 Å/pixel), and 120 pixel (4.26 Å/pixel) side lengths, respectively. Image analysis was carried out with SPIDER and the associated display program WEB (177).

For antibody labeling, the proteins were mixed in approximately 1:2 mass ratios and incubated at room temperature for 3 h before grid preparation. Particle images of TcdA bound to 15A4 were selected with Boxer (176) and windowed with a 136 pixel side length.

Random conical tilt reconstruction of negatively stained TcdA

Micrograph tilt pairs of TcdA at neutral and low pH were recorded at 60° and 0°. Particle pairs (7,396 at neutral pH and 8,319 pairs at low pH) were selected interactively from both the images of the untilted and 60° tilted sample using WEB, and windowed into 128 x 128 pixel images (4.48 Å/pixel). The untilted images were rotationally and translationally aligned and subjected to 10 cycles of multi-reference alignment and K-means classification. Particles of neutral pH TcdA were grouped into 12 classes (Figure 3-1A). From the class averages, three representative projections were chosen and used

as references for another cycle of multi-reference alignment (Figure 3-1A, marked with a '*'). TcdA particles at low pH were aligned to four references chosen from a previous alignment of ~4,000 images of TcdA at a low pH (Figure 3-3A, marked with a '*').

The larger of the resulting classes for both the neutral pH and low pH TcdA particles (4,956 and 2,503 particles, respectively) (Figure 3-1B and 3-3B, marked with a '*') were used to calculate an initial 3D reconstruction by back-projection using the in-plane rotation angles determined by rotational alignment and the pre-selected tilt angle of 60˚ implemented in the processing package SPIDER (177). The density map was improved by back-projection and angular refinement in SPIDER. 10% of the particles selected from the images of the untilted specimens in either the neutral or low pH class were included in the data set (500 and 250 particle images, respectively) and angular refinement was repeated. The Fourier shell correlation (FSC) curve corresponds to normalized cross-correlation coefficients of Fourier shells from even and odd particles within the dataset. Using a FSC = 0.5 criteria (Figure 3-1C and 3-3C), neutral and low pH structures were filtered to 25 Å and 30 Å resolutions, respectively. The 3D structure of the neutral pH TcdA structure was also filtered using the IMAGIC-V software package (178) using the THREED-SMOOTH command to diminish 'salt and pepper' noise from the map by removing single voxels that were unconnected to the main volume of the 3D density. The contouring threshold was chosen so that the volume of the each TcdA structure was continuous. The estimated molecular weight of each structure was then calculated in IMAGIC-V (178) using this value and a protein density of 0.8 Da/$Å^3$. The surface rendering of the structure was performed with the program Chimera (179).

A

B

C
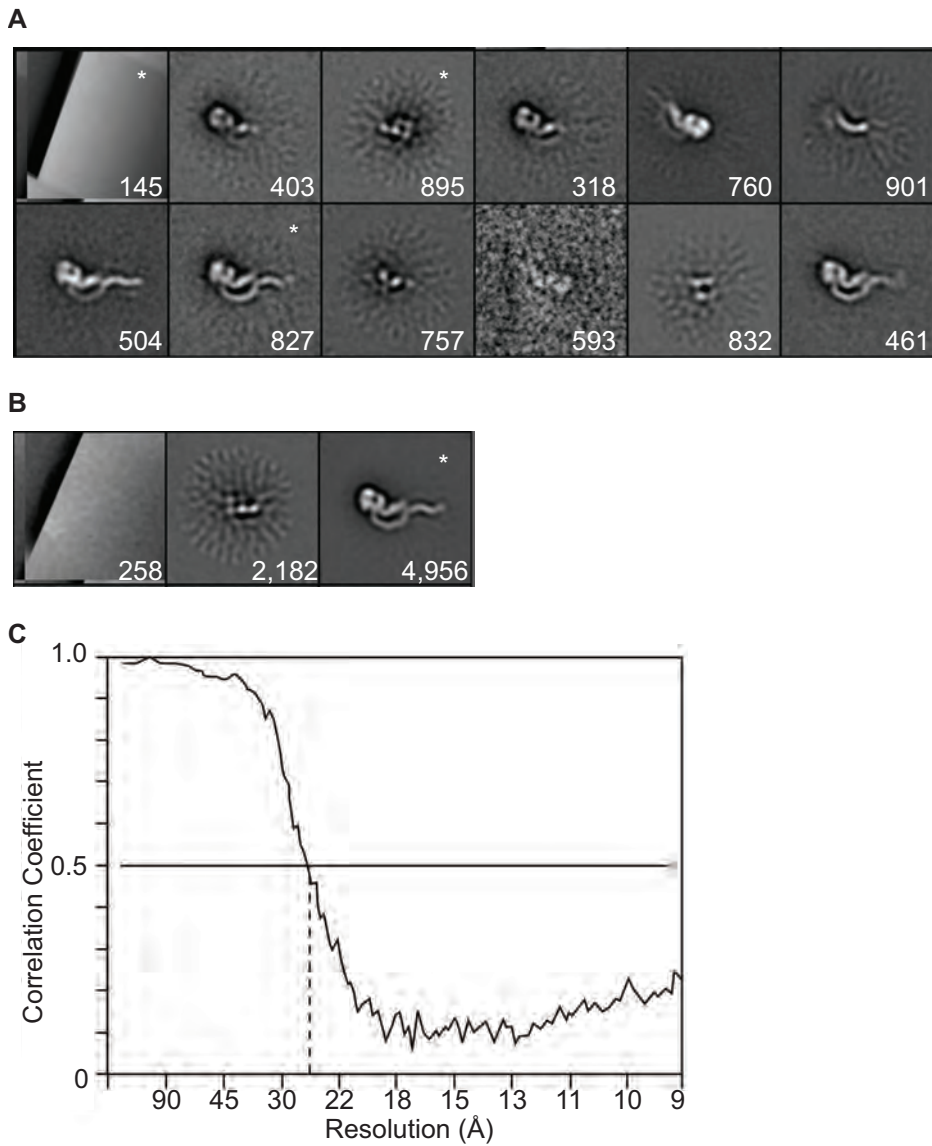
Figure 3-1. EM analysis of TcdA. (A) Average of 7,396 TcdA particles classified into 12 class averages. Boxes marked with "*" indicate averages used as references for (B). (B) Multi-reference alignment of 7,396 particles into 3 classes. The class indicated with "*" was used for the reconstruction. The numbers in each square indicate the number of particles in each class average. Each box has a side length of 57.3 nm. (C) FSC curve for neutral pH structure of TcdA.

A

B

C

279	263 *	197	205	398 *
210	185	224	22 *	150 *

D

915	620	16	582
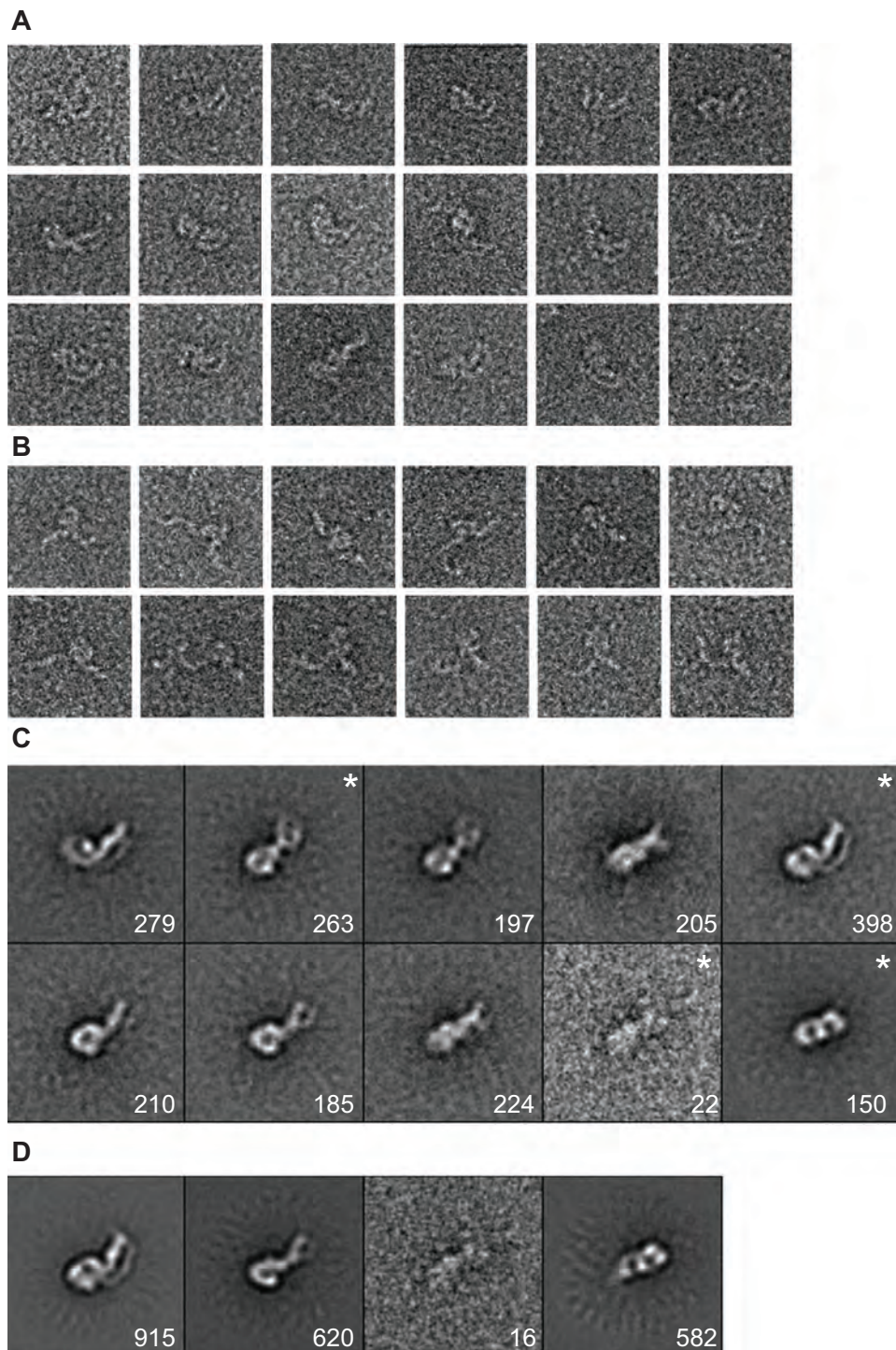
Figure 3-2. EM analysis of TcdB. (A) A sample of single particle images that resemble TcdA. (B) Images of single particles that appear to be in a more 'open' conformation. (C) Averages of 2,133 TcdB particles grouped into 10 class averages. Boxes marked with "*" indicate averages used as references for (D). (D) Multi-reference alignment of 2,133 particles into 4 classes. The numbers in each square indicate the number of particles in each class average. Each box has a side length of 51.1 nm.

A

248	382 *	668 *	209	705
600	318 *	152	237	495 *

B

1491	2384 *	1941	2503 *
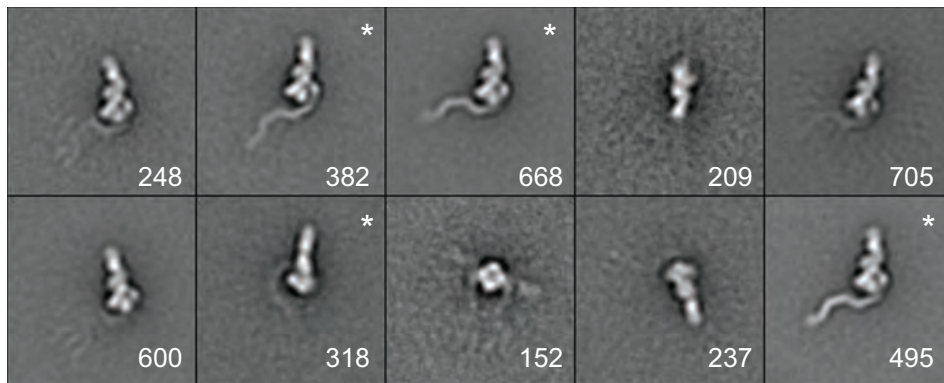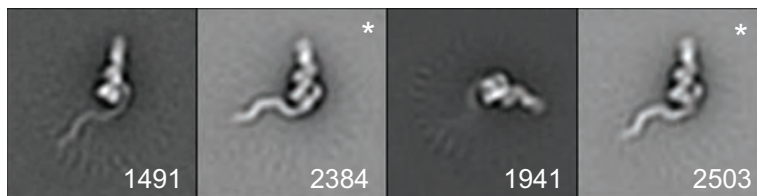
C
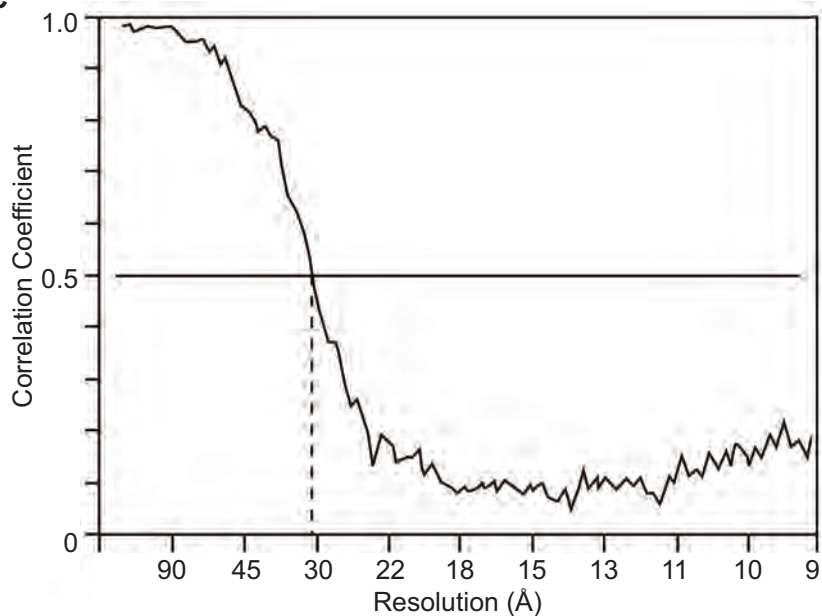
Figure 3-3. EM analysis of low pH TcdA. TcdA was adsorbed to the grid, washed with 50 mM sodium citrate, pH 4.5, stained with uranyl formate and visualized by EM. (A) Alignment of 4014 low-pH TcdA particles into 10 groups. The numbers indicate the number of particle averaged for each image. Boxes marked with "*" indicate classes used as references for (B). (B) Reference alignment of low-pH TcdA particles into 4 classes. The classes indicated with "*" were used as references for 3D structure determination. The side length of each box is 57.3 nm. (C) FSC curve for low-pH structure of TcdA.

Generation of monoclonal antibodies

Four BALB/c mice were immunized with a fragment of the TcdA delivery domain corresponding to residues 771-1460. For primary injections, 50 µg of purified antigen were emulsified in 50% PBS, 50% Freund's complete adjuvant and injected subcutaneously into the nape of the neck (50%) and intramuscularly to the gluteal muscles (50%). In subsequent booster injections, Freund's incomplete adjuvant was substituted for Freund's complete adjuvant. Four days after the final immunization, spleen cells were harvested and fused by standard methods with SP2/0 myeloma cells. Antibodies from the resulting hybridomas were screened by ELISA for interaction with the delivery domain (799-1460) and holotoxin, but not the glucosyltransferase, autoprotease, or receptor binding domains. Clones producing antibodies with the desired properties were subcloned to ensure monoclonality and cryopreserved. Monoclonal antibodies were purified from the supernatant of the selected clone by affinity chromatography on Protein-G sepharose (GE Life Sciences).

ELISA for screening monoclonal antibodies

Either recombinant protein or native holotoxin at 5 µg/mL was coated onto 96-well ELISA plates at 4 ºC overnight. The wells were washed with PBST (PBS with 0.1% Tween) at least three times after this and every subsequent step. The wells were blocked with 1% BSA in PBST for 1 h at room temperature. The antibody 15A4 was diluted to 0.3 µg/mL in PBST and added to the plates for 1 h at room temperature. HRP-conjugated goat anti-mouse-IgG antibody (Jackson) was added to the plates in PBST and incubated 1 hour. The ELISAs were developed by the addition of ABTS substrate solution (1mM ABTS (2,2' azino-di-(3-ethylbenzthiozoline sulfonic acid)) (Sigma) in 70 mM citrate-phosphate buffer, pH 4.2 w/ 0.03% H_2O_2) for 30 minutes, and absorbance was measured at 410 nm.

TcdA glucosyltransferase domain solubility assay

Purified glucosyltransferase domain was exchanged into 150 mM NaCl, 10 mM Tris, pH 8 by gel filtration chromatography and concentrated to ~0.38 mg/mL. 18 µL protein was mixed with 2 µL 1 M sodium citrate at the indicated pH. After a 30 min. incubation at room temperature, the samples were centrifuged at 14000 *g*. The absorbance of the supernatant was measured at 280 nM.

Results

Visualization of TcdA and TcdB by negative stain electron microscopy

TcdA was purified from *C. difficile* culture supernatant (Figure 3-4B) and shown to be active in a cell-rounding assay. Toxin was adsorbed to carbon-coated glow-discharged grids, and stained with uranyl formate. Negative stain electron microscopy (EM) revealed homogeneous particles with a non-symmetric shape and an elongated 'tail' (Figure 3-4C). Image pairs of grids containing negatively stained TcdA were recorded at tilt angles of 60° and 0°. A total of 7,396 particles were selected and images of the untilted specimens were classified into 12 class averages. These class averages revealed that while essentially all TcdA particles adsorbed to the carbon grid in the same orientation, there was some variation in the ability to resolve the elongated 'tail' (Figure 3-1A). From the 12 classes we selected one that represented a well-resolved TcdA particle, one that represented a TcdA particle with poorly resolved 'tail' and one poorly resolved image (Figure 3-1A, marked with a '*') and used them as references for another cycle of multi-reference alignment (Figure 3-1B). The largest of the three resulting classes (4,956 particles, Figure 3-4D and Figure 3-1B, marked with a '*') showed a TcdA particle with many clear structural features (Figure 3-4D). The TcdB holotoxin was also purified and visualized by negative stain EM (Figure 3-4E). While a more heterogeneous

field of particles was seen than with TcdA (compare Figure 3-4C and F), the classification of 2,133 particles into 10 class averages revealed a number of classes that had similar structural features as observed in the TcdA alignment (Figure 3-2C). To improve the alignment we selected four classes as references for a round of multi-reference alignment (Figure 3-1D, marked with a '*'). The largest of the four resulting classes (915 particles, Figure 3-4G and Figure 3-2D) showed a TcdB molecule with a well-resolved globular 'head' domain connected to two extended tails, structural features that are very similar to that of TcdA (Figure 3-4D). Although TcdA and TcdB are clearly structurally similar, the TcdA sample was more homogenous and a better candidate for structural characterization. For this reason further 3D analysis was done using TcdA.

A 3D reconstruction of TcdA was generated using the random conical tilt approach (180) and is presented in Figure 3-5. The FSC curve calculated from our final density map suggests a resolution of ~25 Å based on the FSC = 0.5 criterion (Figure 3-1C) (181). The face view of the 3D density map shows structural features very similar to those seen in the projection average (compare first panel in Figure 3-5 with Figure 3-4D) suggesting that the 3D reconstruction was successful. The 3D density map of TcdA adopts an elongated, asymmetric structure that is ~310 Å x ~150 Å x 90 Å in dimension. The structure contains three prominent features: a 'head' domain, a long kinked 'tail', and short inner 'tail'. The head, ~ 90 Å x 90 Å x 60 Å in size, appears to contain two globular 'pincher-like' domains that are connected by a small density at the top of the head, thus creating a small channel that is ~20 Å wide and 90 Å deep. Emanating from the head domain are two tails. A long kinked tail extends from the bottom of the larger of the two 'pincher' domains and stretches ~ 270 Å in an undulating curve. A second tail domain connects from the smaller of the two 'pincher' domains, extending ~100 Å before making contact with the longer kinked tail domain.
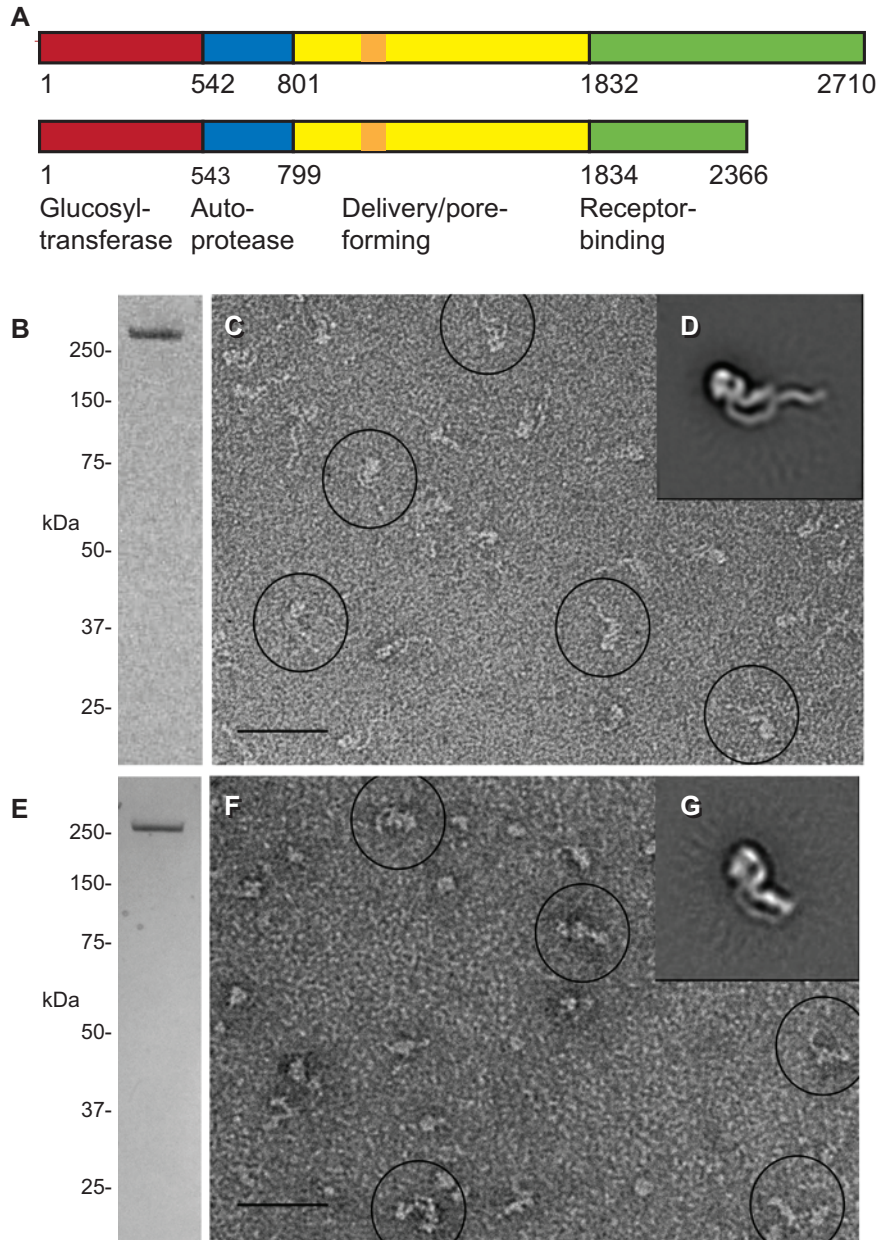
Figure 3-4. Purification and characterization of TcdA and TcdB. (A) The proposed domain organization for TcdA and TcdB is shown. The numbers are for amino acids that mark domain boundaries. (B) SDS-PAGE of purified TcdA, visualized by Coomassie staining. (C) A typical electron micrograph showing TcdA particles in negative stain. A few of the particles are circled in black. Scale bar, 500 Å (D) A representative class average of TcdA particles (4,956) selected from images of untilted specimens in negative stain. Side length of panel is 57.3 nm. (E) SDS-PAGE of TcdB, visualized by Coomassie staining. (F) Electron micrograph of TcdB in negative stain with a few particles circled in black. Scale bar, 500 Å (G) Class average of 915 TcdB particles in negative stain. Side length of panel is 51.1 nm.
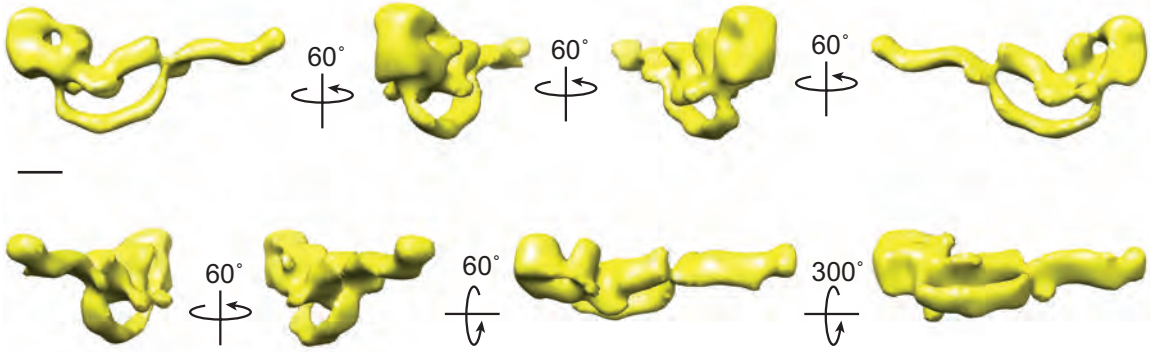
Figure 3-5. Random conical tilt reconstruction of TcdA in negative stain. 3D reconstruction of TcdA filtered to 25 Å. The structure is rotated about the vertical axis in 60° steps or about the horizontal axis by 60° or 300° steps (in reference to the top left structure) as indicated by arrows. TcdA has an elongated shape with three distinct domains: a head domain, a long kinked tail, and a shorter straight tail. Scale bar is 5 nm.

Identification of TcdA domains

To gain insight into how the functional domains of TcdA are organized in the 3D structure we performed direct domain visualization and domain subtraction experiments. The C-terminal domain of TcdA is composed of 39 sequence repeats and is responsible for binding cell surface carbohydrates (115). Crystal structures of two fragments from the C-terminus of TcdA revealed a β-solenoid fold, suggesting that the entire TcdA C-terminal binding domain would adopt an elongated serpentine structure (115,124) (Figure 3-6A). Several features of this predicted model are seen in the long tail observed by EM (Figure 3-6A,B). To confirm that the long tail is indeed the binding domain, we expressed the TcdA binding domain (amino acids 1832-2710) in *E. coli*, purified it, and subjected it to negative stain EM (Figure 3-6C,D). The kinks, corresponding to the seven long repeats, and the approximate lengths of the straight sections observed by EM are consistent with the model derived by crystallography (115) (Figure 3-6). Thus, the long kinked tail domain found in our structure represents the C-terminal binding domain of TcdA.
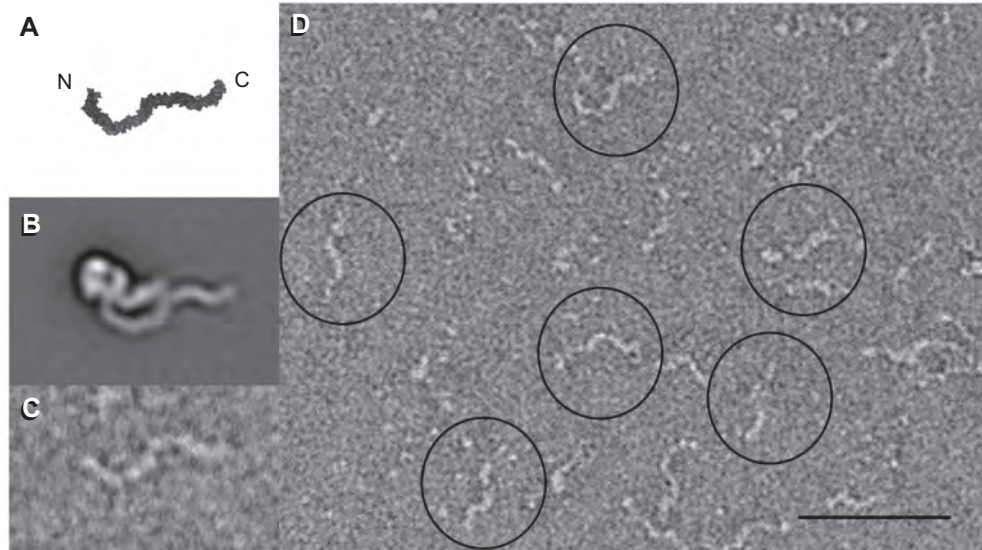
Figure 3-6. Characterization and localization of the TcdA C-terminal binding domain. (A) The predicted model of the TcdA binding domain (115). (B) Average of 4,956 TcdA particles (from Figure 3-4D). (C) Image of a single negative stained particle of the TcdA binding domain. (D) A typical electron micrograph area of the recombinantly expressed C-terminal domain of TcdA (aa 1832-2710). A few of the particles are circled in black. Scale bar, 500 Å.
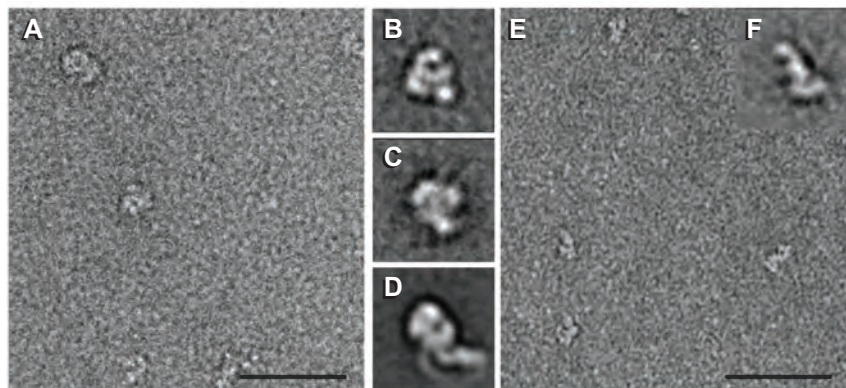
Figure 3-7. Visualization of the TcdA delivery domain by negative stain EM. (A) A typical electron micrograph showing particles of the TcdA delivery domain. The particles have a globular shape resembling the 'head' of TcdA. Scale bar, 500 Å. (B-C) Two class averages of the TcdA delivery domain. Classes contain 139 and 314 particles respectively. Side length of panel is 27.2 nm (D) Class average of the head of TcdA. Side length of panel is 27.2 nm (E) TcdA delivery domain was applied to an EM grid, washed with sodium citrate at pH 4.5, and visualized by negative stain EM. Scale bar, 500 Å. (F) Class average (405 particles) of the TcdA delivery domain following exposure to a low pH buffer. Side length of panel is 27.2 nm.

66

To identify the location of the central TcdA delivery domain, we imaged a recombinantly expressed protein corresponding to residues 799-1859. The images suggest that the protein is capable of binding the grid in a variety of orientations (Figure 3-7A), but classification of 1,523 particles into 5 class averages resulted in two classes of 139 and 314 particles (Figure 3-7B,C) that recapitulate the size, shape and pincher-like features of the holotoxin head domain (Figure 3-7D). In an alternative approach, we labeled the toxin with a monoclonal antibody specific for the delivery domain (Figure 3-8). The antibody bound at the top of the head to the larger of the two lobes providing further evidence that the head corresponds to the delivery domain.
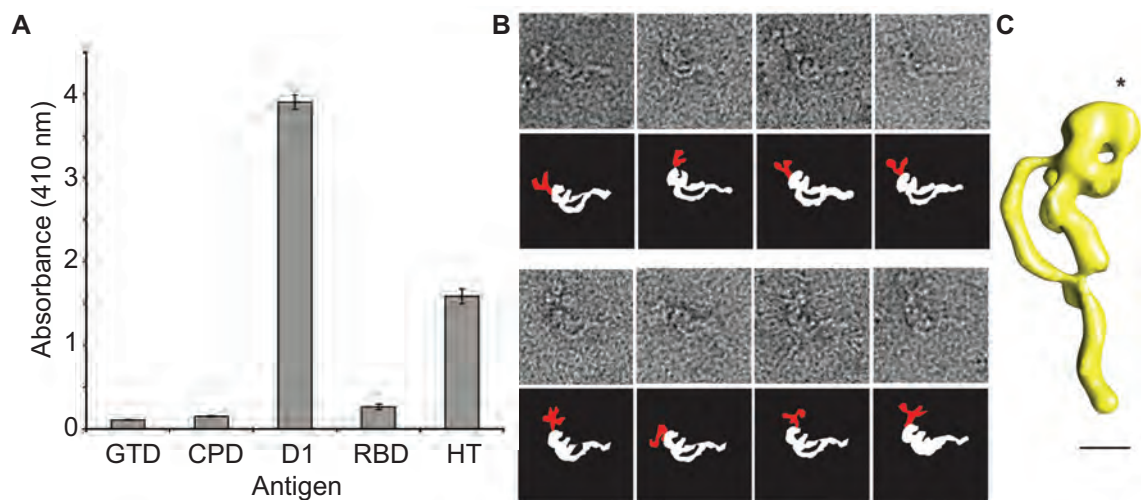
Figure 3-8. Labeling TcdA with a monoclonal antibody against the delivery domain. (A) The monoclonal antibody 15A4 was tested by ELISA for recognition of native holotoxin and the TcdA glucosyltransferase (aa 1-543), autoprotease (aa 543-795), delivery (aa 799-1460), and binding (aa 1832-2710) domains. Binding was detected using an HRP-conjugated secondary antibody. (B) TcdA was incubated with 15A4 and visualized by negative stain EM. A gallery of labeled TcdA particles is depicted. Below each image is a schematic representation of how the antibody (red) is bound to TcdA (white). Side lengths of panels are 60.9 nm. (C) Two-dimensional view of 15A4 antibody position (marked with '*') on the 3D model of TcdA. Scale bar, 5 nm.

The proposed function of this domain is to change structure in response to the low pH of the endosome and form a pore. To test whether such a structural change

would be visible by EM, we applied the protein to a grid and then washed the grid with a buffer at a pH of 4.5 before staining with uranyl formate (Figure 3-7E). The most populated class average from this analysis (Figure 3-7F) suggests that this domain is capable of changing into an extended conformation at low pH.

With the long tail identified as the receptor-binding domain, and the head identified as the delivery domain, we hypothesized that the small tail most likely represents the N-terminal glucosyltransferase domain. To test this, we induced autoproteolytic removal of the TcdA glucosyltransferase domain by incubating holotoxin with InsP6 and DTT (Figure 3-9A). The glucosyltransferase domain was removed by gel-filtration chromatography and the larger fragment was analyzed by negative stain EM (Figure 3-9). Upon examining the images of cleaved TcdA (Figure 3-9B,C), it is clear that while the long tail and globular domains are still visible, the shorter tail is missing. Thus, the shorter tail domain seen in the 3D TcdA structure is the glucosyltransferase domain.
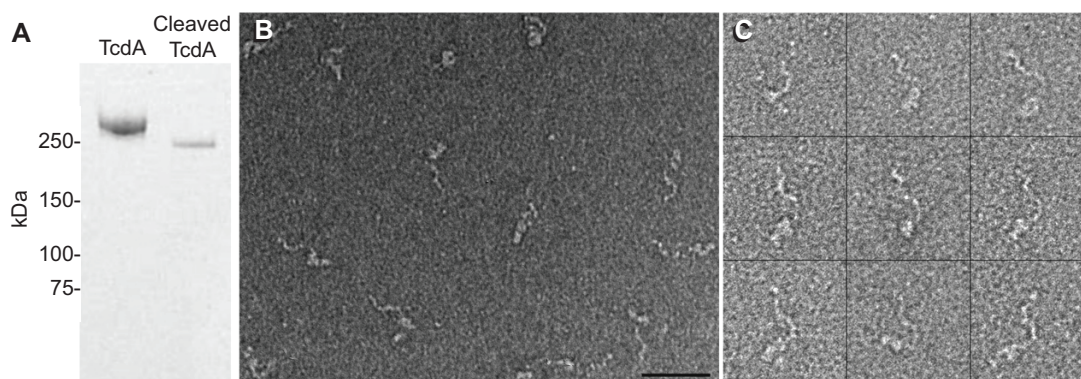
Figure 3-9. Localization of the TcdA N-terminal glucosyltransferase domain by autoproteolysis. Cleavage of TcdA was induced by the addition of InsP6 and DTT. (A) Coomassie stained SDS-PAGE of native TcdA and the larger fragment of TcdA following autoproteolytic cleavage and purification by size exclusion chromatography. (B-C) The large fragment containing residues 543-2710 was isolated, applied to a grid, stained with uranyl formate, and visualized by negative stain EM. (B) Representative image of cleaved TcdA particles in negative stain. Scale bar, 500 Å. (C) A gallery of TcdA cleaved particles. Although the head and long kinked tail domain are visible, the short tail region is missing. The side-length of each panel is 57.3 nm.
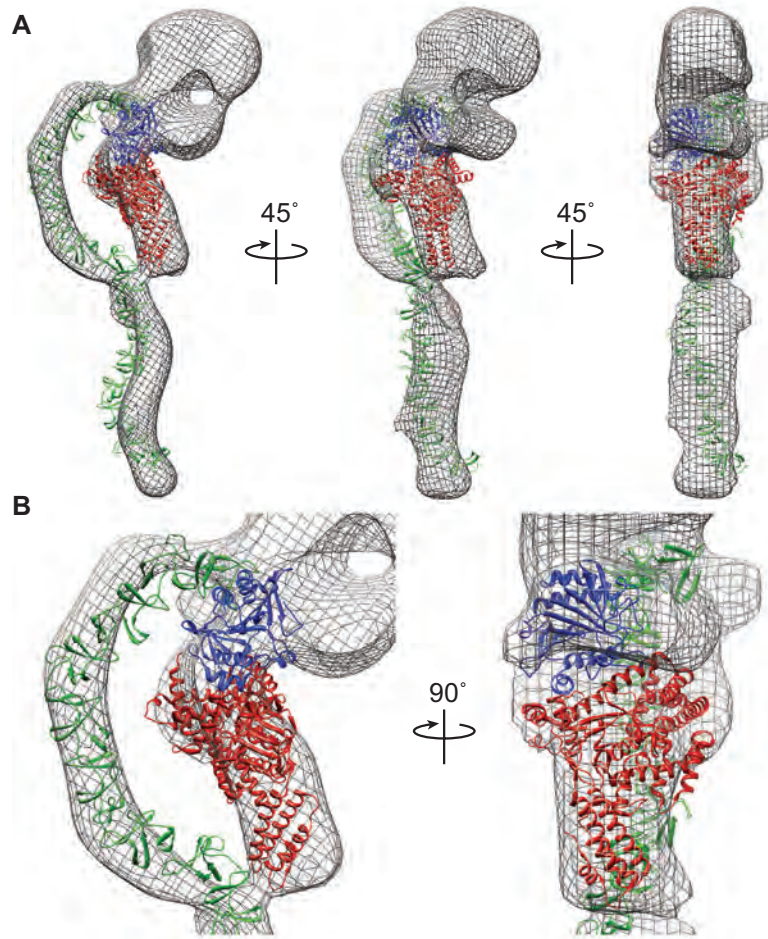
Figure 3-10. Placement of the functional domains of TcdA within EM map. (A,B) The 3D reconstruction of TcdA filtered to 25 Å is shown as mesh surface. Structures of the TcdB glucosyltransferase domain (red), the TcdA autoprotease domain (blue), and a model of the TcdA binding domain (green) were manually placed into the density. Panel B shows two views of the toxin enlarged to show the placement of the glucosyltransferase and protease domains.

A model of the binding domain of TcdA was placed into the 3D EM map (Figure 3-10). The N-terminus of the domain connects to the back of the head near the larger of the pincher domains, whereas the C-terminus extends away from the rest of the toxin. As is common for negative stained specimens, the binding domain was somewhat flattened in the EM map compared to the model. The distortion was fairly minor, however, and most of the model clearly and unambiguously fit into the EM map. The

crystal structure of the glucosyltransferase domain of TcdB was placed into the small tail in the EM map with the N-terminus distal from the head domain (Figure 3-10). The domain is 90 Å long, consistent with the ~100 Å length of the short tail, and is about 25 Å wide at the extreme N-terminus. The C-terminal side of the domain is considerable larger with a diameter of about 65 Å. The small tail in the EM map has a narrow distal end and a larger, wider region near the head that is consistent with the shape of the TcdB glucosyltransferase domain. We did not determine the location of the autoprotease domain experimentally. However, there are only four amino acids between the glucosyltransferase and autoprotease domains that are not present in either crystal structure. The crystal structure of the autoprotease domain was, therefore, positioned between the small tail and the small pincher domain and oriented so that the C-terminal residue of the glucosyltransferase structure (Leu542) could be connected with the N-terminal residue of the autoprotease structure (Gly547) using a short linker.

pH dependent conformational changes of TcdA

To understand the structural basis for pore formation within the endosome, we analyzed TcdA particles at low pH. TcdA was applied to a carbon-coated, glow-discharged grid, washed with a pH 4.5 buffer, and then stained with uranyl formate. This resulted in a clear conformational change in the toxin as shown in Figure 3-11 (compare with Figure 3-4C). To examine the structural homogeneity of TcdA in a low pH state, approximately 4,000 particles were selected from images of untilted specimens and classified into 10 groups. This analysis revealed a number of classes with structurally homogenous particles suitable for further structural analysis (Figure 3-3A). Significantly, although the binding domain in these classes looks similar to TcdA at neutral pH, the head domain clearly has undergone a major conformational rearrangement (compare Figure 3-11B with 3-4D). To more carefully address the structural changes in TcdA

70

structure at endosomal pH, additional images of TcdA in a low pH state were collected at 60° and 0°. 8,319 particle pairs were selected, and images from the untilted specimens were grouped into four classes by reference based alignment using references chosen from the original alignment of ~4,000 particles (Figure 3-3A, marked by *). Two of the resulting classes looked very similar, differing only in the orientation of the receptor-binding domain relative to the head domain (Figure 3-5B, marked by *). One of these classes was used to generate a 3D reconstruction with a resolution of ~30 Å (Figure 3-11C, and Figure 3-3C).
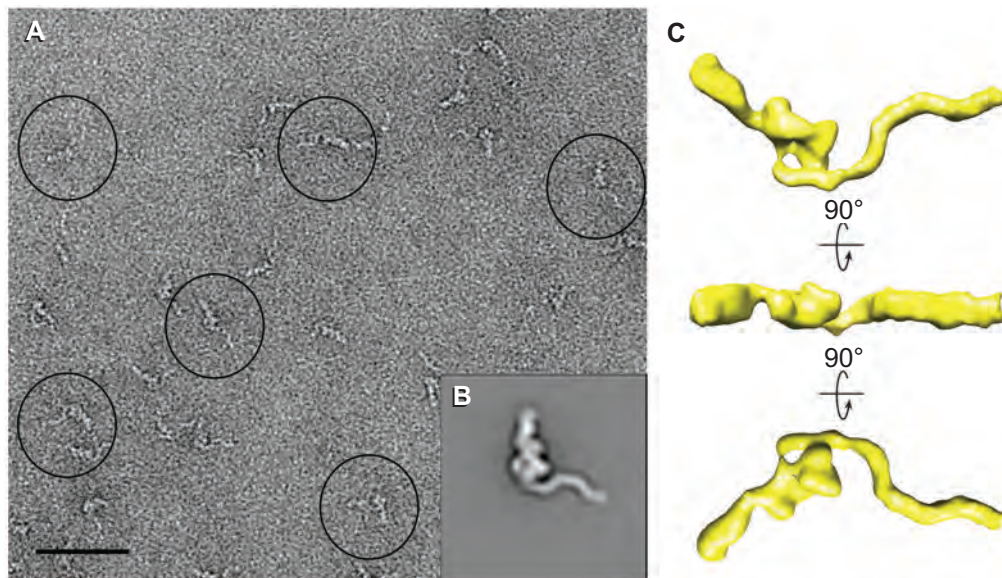
Figure 3-11. pH induced conformational changes of TcdA. TcdA was adsorbed to a carbon-coated glow-discharged grid, washed with 50 mM sodium citrate at pH 4.5, and stained with 0.7% uranyl formate. (A) A typical electron micrograph of TcdA at low pH. A few of the particles are circled in black. Scale bar, 500 Å. (B) A representative class average of negatively stained TcdA particles (1,327) at low pH. Side length of panel is 57.3 nm. (C) The 3D reconstruction of TcdA in a low pH state is shown filtered to 30 Å.

The 3D structure of TcdA at low pH shows a molecule that has undergone a major conformational change from the structure observed at neutral pH. Although the conformation of the binding domain has not changed, the low pH form is more elongated

due to an extension in the opposite direction from the binding domain. This extended 'appendage' is ~100 Å long and ~ 40 Å x ~40 Å wide at the distal end. The proximal end where the appendage connects to the rest of the toxin is narrower (~20 Å x 20 Å). The orientation of this extended appendage is relatively flexible as seen in both the class averages and 3D reconstructions (Figure 3-3B).

The low pH structure of TcdA has a volume that corresponds to a MW of ~240 kDa and is thus considerably smaller than the structure of native TcdA whose volume corresponds to a MW of ~320 kDa. Since the low pH and neutral pH forms of the toxin look identical by SDS-PAGE (Figure 3-12A), this difference may be due to greater flexibility and/or partial unfolding at low pH. When comparing the neutral and low pH structures, it appears that much of this volume difference can be attributed to the glucosyltransferase domain. To investigate whether the TcdA glucosyltransferase domain unfolds at low pH, we purified this domain and tested its solubility in a range of pH conditions. At pH < 5 the glucosyltransferase domain precipitates (Figure 3-12B), consistent with the behavior of an unfolded protein.
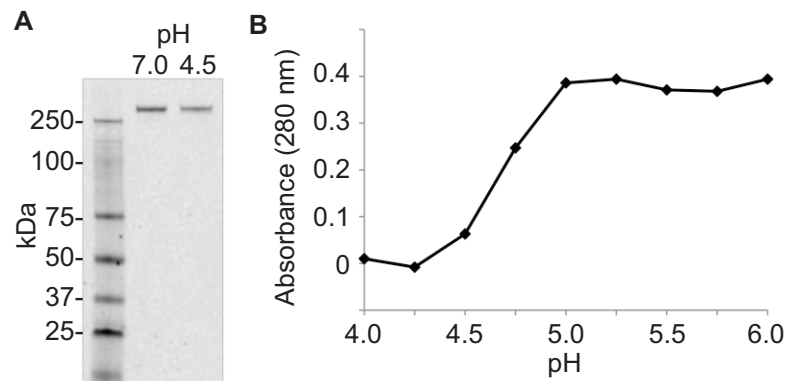
Figure 3-12. pH dependence on the solubility of TcdA glucosyltransferase domain. (A) TcdA was incubated at either pH 7 or pH 4.5 for 5 minutes. The proteins were run on a gel and visualized by SDS-PAGE with Coomassie staining. (B) Recombinantly expressed glucosyltransferase domain (0.38 mg/mL) was exposed to a range of pHs, and the absorbance of the soluble protein was measured at 280 nm.

Discussion

The homology between TcdA and TcdB (68% similar, 47% identical) and the similar modes of entry suggest that these toxins will adopt similar three-dimensional structures. Although in our studies TcdB was more structurally heterogeneous than TcdA, both toxins clearly have a bi-lobed, globular 'head' domain that directly connects to two extended tails. Due to the similarities between TcdA and TcdB at both a structural and primary sequence level, we propose that, as with TcdA, the TcdB globular head domain represents the delivery domain, while the short tail contains the glucosyltransferase domain and the serpentine tail corresponds to the receptor binding domain. Since TcdA was considerably more homogeneous in structure, we focused our domain mapping and 3D structural studies on this protein.

The structures of TcdA holotoxin provide a framework for considering the molecular events required for LCT cellular intoxication. The first event is molecular recognition of a receptor on the surface of host cells. Studies suggest that the receptors for TcdA likely differ from those used by TcdB, although both toxins can bind specifically to carbohydrate structures containing an N-acetyl-lactosamine core (119,182,183). The TcdA and TcdB binding domains consist of two types of repetitive peptide sequences: 19-22 amino acid short repeats and 31 amino acid long repeats (116). In TcdA, 32 short repeats are interspersed by 7 long repeats (115). Crystal structures of binding domain C-terminal fragments have revealed that the short repeats form β-solenoid subunits that pack together in extended rods (115,124,172). Long repeats also form β-solenoids, but are packed differently yielding kinks in the rods. Each long repeat together with an adjacent short repeat form a binding site for the saccharide receptors (124). A model of the entire TcdA binding domain generated from crystallographic data and the protein sequence suggests an extended kinked structure that agrees well with our EM

73

observations (Figure 3-6) (115). Although a narrow, elongated structure might be expected to adopt multiple conformations, our EM images indicate that the domain is fairly rigid. This rigidity is consistent with the highly conserved packing interactions and regular rotational relationships observed between the long and short repeats in three different crystal structures of TcdA fragments. This is also reflected by the homogeneity of the particles and the fact that the domain structure does not change with pH. The structure demonstrates that the binding domain extends away from the delivery domain such that multiple binding sites are accessible, consistent with a model of multivalent binding.

The principle difference in the 2D images of TcdA and TcdB is that the TcdB density corresponding to the receptor-binding domain tail is considerably shorter, consistent with it being 40% shorter in its sequence. The structure of this domain is predicted to have four structural modules (each consisting of 3-5 short repeats and 1 long repeat) as compared to seven in TcdA (115). Truncating the TcdA structural model after 4 modules results in a structure similar in length and shape to what we observe by EM. The density for the TcdB receptor binding domain was more difficult to observe in our class averages (Figure 3-2). We interpret this to mean that the TcdB receptor-binding domain is able to adopt multiple orientations with respect to the rest of the molecule. This heterogeneity is the likely explanation for why the molecular envelope for TcdB obtained by SAXS differs from what we observe (172) since *ab initio* envelope calculations from scattering data can be problematic in flexible systems where domain orientations differ between conformers (184,185).

TcdA and TcdB have been proposed to undergo pH dependent conformational changes in order to form a pore through which the glucosyltransferase domain is translocated (128). We have directly visualized pH inducible changes in TcdA and the TcdA delivery domain by exposing them to low pH on EM grids. We see significant

changes in the pincher-like head of the delivery domain that results in its extension away from the binding domain (Figure 3-13 and 3-7E,F). This might be accomplished through a decoupling of the two lobes of the head, effectively opening the pincher.
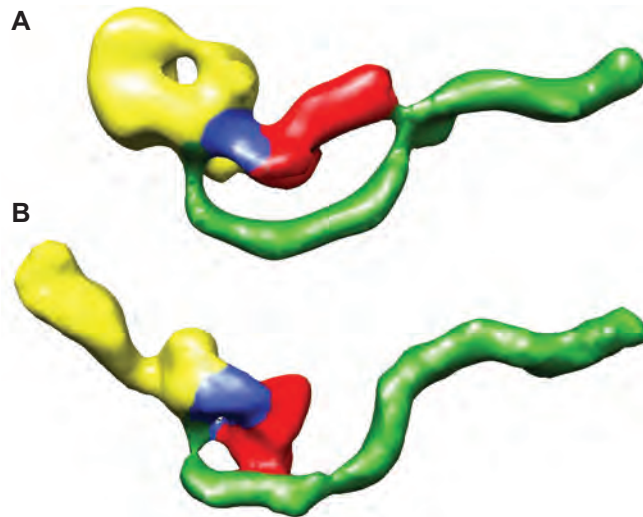
Figure 3-13. Model of the conformational changes induced at low pH. 3D reconstruction of TcdA in a (A) neutral pH and (B) low pH state. The proposed locations of the functional domains are colored as follows: glucosyltransferase (red), autoprotease (blue), delivery (yellow) and binding (green).

The two-lobed pincher-like structure of the head reveals that the TcdA delivery domain has a complex structure, and might contain two subdomains. BLAST analysis of the complete TcdA delivery domain (residues 801-1831) reveals that there are two regions with distinct homology profiles: residues ~ 801-1400 (D1) and 1401-1831 (D2) (Figure 3-14) (186). TcdA D1 is more highly conserved among LCTs (55% identity with TcdB) than D2 (33% identity with TcdB) and contains the putative membrane spanning residues, suggesting that D1 is the region that rearranges to form the elongated appendage of the low pH form. In D2, most of the BLAST homologs are distant and uncharacterized, but the region is thought to enhance the binding of TcdA to cells (182) and contains a DSG aspartate protease motif which may be involved in toxin delivery

into the cytosol (134). Further study is needed to dissect the respective roles of these delivery domains in the context of the membrane.
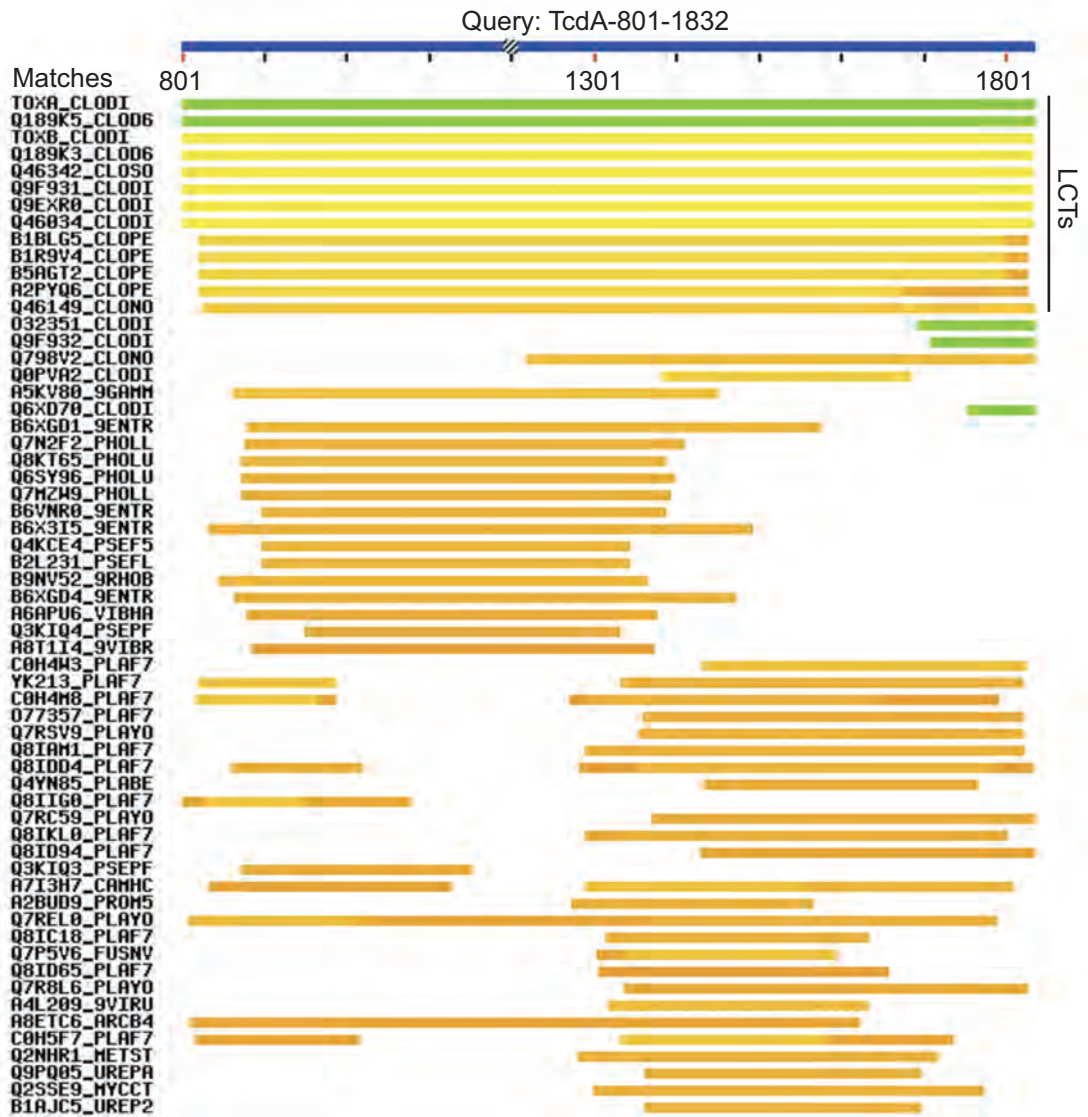
Figure 3-14. BLAST analysis of the TcdA delivery domain. Analysis of the TcdA delivery domain protein sequence (residues 801-1831) using BLAST reveals two regions of homology.

Release of the glucosyltransferase domain into the cell is mediated by the adjacent cysteine protease domain (135). The location of the autoprotease domain within the map of TcdA was not experimentally determined, but the structure is

anticipated to be located so that the C-terminus of the glucosyltransferase domain can be cleaved in the active site of the protease (173). The TcdA autoprotease domain is known to undergo significant rearrangement upon exposure to InsP6 (173). Analysis of cleaved TcdA in negative stain reveals that the particles are much more heterogenous than those of native TcdA and impeded our efforts to obtain class averages. The heterogeneity of cleaved TcdA likely results from the removal of the glucosyltransferase domain, since this domain makes contact with the binding domain in the structure determined at neutral pH. It is tempting to speculate that this contact may help "lock" the TcdA head in a non-pore forming state, a model that is further supported by our data showing the glucosyltransferase domain loses structural stability at low pH.

In our structures of TcdA, we observed a smaller volume at low pH than in the neutral pH structure. We attribute this loss of volume to partial unfolding of the glucosyltransferase domain. The enzymatic components of anthrax toxin, botulinum neurotoxin, and diphtheria toxin have also been shown to unfold at low pH (187-189). In these toxins, which also form pores at endosomal pH, unfolding is thought to be necessary for translocation of the enzymatic domains through narrow membrane pores. In spite of this mechanistic similarity, anthrax toxin, botulinum neurotoxin, and diphtheria toxin have been noted for their diversity in structure (190). The unique structural features observed in this study, namely a bi-lobed delivery domain tethered to enzymatic cargo and a serpentine receptor-binding domain, suggest that TcdA and TcdB represent yet another structural theme for bacterial protein toxins. Further work is needed to understand the commonalities and differences between the TcdA and TcdB structures and how these structures guide the complex functions of these molecules.

CHAPTER IV

STRUCTURAL DETERMINANTS OF THE GLUCOSYLTRANSFERASE ACTIVITY OF *CLOSTRIDIUM DIFFICILE* TOXIN A

Introduction

Although TcdA and TcdB have long been accepted as the primary virulence factors responsible for CDAD, there are conflicting data concerning the relative importance of each of the two toxins in causing disease (101,103,104,107,108,113,114,191). Although the roles of the toxins in human disease are unclear, a number of differences in activity have been noted for TcdA and TcdB in cells and in animal models. TcdA has been shown to be a potent enterotoxin in hamster and rabbit ileal loop assays, whereas TcdB is not (103,104). Both toxins are reported to have potent enterotoxic effects on human tissues (101,111). In cell culture, both TcdA and TcdB induce cell rounding, but TcdB is ~1000 times more potent than TcdA in most cell lines (84,192).

The molecular basis for the observed differences in TcdA and TcdB cytopathicity could include differences in the binding (84), pore-forming (130), autoproteolysis (193), and glucosyltransferase (84) activities. Multiple studies have demonstrated that TcdA and TcdB have different binding activities, suggesting that the toxins have distinct receptors (84,119,182,183). Distinct binding targets almost certainly contribute to differences in potency towards various cells (84). However, in one of the few studies directly comparing the activities of TcdA and TcdB, Chaves-Olarte *et al.* have reported that the difference in glucosyltransferase activity was the major determinant contributing to the difference in cytopathic potency between TcdA and TcdB. The authors showed

that TcdB holotoxin (HT) is ~100 times more active than TcdA HT at targeting substrate *in vitro*. In the absence of GTPase, TcdB also had a higher rate of UDP-glucose hydrolysis (84).

To understand the structural basis for the differences in glucosyltransferase activities of TcdA and TcdB, we have determined crystal structures of the TcdA GTD with and without its co-substrate UDP-glucose. We find that the enzymatic core of the TcdA GTD is highly similar to that of the TcdB GTD but that the GTPase recognition surface differs significantly. In evaluating the functional importance of divergence in these surfaces, we observe that both GTDs modify RhoA, Rac1, and Cdc42 at comparable levels *in vitro,* but that TcdA can also modify members of the Ras-GTPase family: Rap1A and Rap2A. The TcdA-mediated glucosylation of Rap2A is detected in a cell intoxication model, while Rap2A glucosylation is not observed in TcdB- or mock-treated cells.

In addition, we have compared glucosyltransferase activities of the holotoxins with those of the GTDs alone. We find that in the context of the holotoxin, glucosylation of substrates is inhibited. Activity can be restored by release of the GTD through autoproteolysis. We propose a model wherein the receptor-binding domain occludes the binding of GTPase substrates and discuss the importance of cellular activation in determining the array of substrates available to the toxins once delivered into the cell.

Methods

Plasmid construction and point mutants

The nucleotide sequences encoding aa 1-2710 of TcdA (TcdA holotoxin), 1-542 of TcdA (TcdA GTD), or 1-543 of TcdB (TcdB GTD) were amplified from VPI 10463 and cloned into the *Bacillus megaterium* expression vector pC-His1622 (MoBiTec, BMEG20)

79

using the restriction sites BsrGI and SphI. The gene encoding TcdB was also cloned into BMEG20 as described previously (194). The GTPase sequences were cloned into pGEX4T (RhoA (195), Rap1A, Rap2A) or pGEX2T (Rac1 (195), Cdc42 (195)) using the sites BamHI and EcoRI to generate GST-GTPase fusions.

Protein Expression and Purification

The TcdA, TcdB, TcdA GTD, and TcdB GTD plasmids were transformed into *B. megaterium* following the manufacturer's instructions (MoBiTec). Transformed *B. megaterium* were grown in luria broth (LB) containing 10 mg/L tetracycline. 30 mL of overnight culture was used to inoculate 1 L of media and the cultures were placed at 37°C and 230 rpm. When the cultures reached OD_{600} = 0.3, expression was induced by the addition of 5 g D-xylose. After 4.5 h the cells were harvested by centrifugation and resuspended in 100 mM NaCl, 20 mM Tris, pH 8.0. Following French Press lysis, the lysates were centrifuged at 48,000 *g* for 20 min. Protein was purified from the supernatant by Ni-affinity chromatography followed by gel-filtration chromatography in 100 mM NaCl, 20 mM Tris, pH 8.

The GTP binding proteins were expressed in *E. coli* BL21 cells grown in LB containing 100 mg/L ampicillin. 10 mL of overnight culture was used to inoculate 1 L of media and the cultures were placed at 37°C and 230 rpm. When the cultures reached OD_{600} = 0.6, the temperature was changed to 21°C and expression was induced by the addition of 0.5 mM IPTG. After 16 h, the cells were harvested by centrifugation and resuspended in 100 mM NaCl, 20 mM Tris, pH 8.0. Following French Press lysis, the lysates were centrifuged at 48,000 *g* for 20 min. Protein was purified from the supernatant using glutathione-sepharose 4B (GE) followed by gel filtration chromatography. The GTPase GST tags were not removed.

Crystallization

TcdA GTD was concentrated to 16 mg/mL in 150 mM NaCl, 20 mM Tris, pH 8.0. TcdA GTD was crystallized by the sitting-drop method at 21°C with a 1:1 ratio of protein of mother liquor containing 0.2 M L-proline, 10% PEG 3350, and 0.1 M HEPES, pH 7.5. For co-crystallization of the GTD with substrate, 10 mM UDP-glucose and 2 mM $MnCl_2$ were added to the protein, and hanging drops were prepared with mother liquor containing 20% PEG 6000 and 0.1 M bicine, pH 8-9. Crystals were exchanged into the appropriate mother liquor containing either 15% glycerol (protein alone) or 20% ethylene glycol (protein plus UDP-glucose/Mn^{2+}), mounted on cryo loops, and flash cooled in liquid nitrogen.

Structure Determination and Refinement

X-ray data were collected from single crystals on NE-CAT beamline 24 ID-C at the Advanced Photon Source (Argonne, IL) at 100 K and a wavelength of 1.0094 Å. Diffraction data were indexed, integrated, and scaled using HKL2000 (155). A starting model was obtained for the TcdA GTD without UDP-glucose by molecular replacement with the TcdB GTD (pdb 2BVM) as a search model using Phenix. The model was iteratively built using Coot (157) and refined using Phenix (158) with 5 TLS groups per chain (Table 4-1). The structure with UDP-glucose bound was determined in the same way, except the apo- structure was used as the search model for molecular replacement. In the final structures, 90.0 and 91.2% of the residues were in the most favored positions in the Ramachandran plot for the bound and apo structures, respectively (calculated by PROCHECK (159)). No residues were in disallowed regions. For the apo structure the final model consists of residues 2-538, 1 manganese ion, and 233 water molecules. For the structure with UDP-glucose bound, the final model consists of residues 2-538, 1 manganese, 1 UDP-glucose, and 109 water molecules.

Table 4-1. X-ray data collection and refinement statistics for the crystal structures of the TcdA GTD

	TcdA GTD with UDP-glucose	Apo-TcdA GTD
Data collection		
Space group	$P6_5$	$P6_5$
Cell dimensions		
a, b, c (Å)[1]	141.8, 141.8, 63.4	141.9, 141.9, 66.0
α, β, γ (°)	90, 90, 120	90, 90, 120
Resolution (Å)	50-2.58 (2.67-2.58)	50-2.2 (2.28-2.2)
R_{merge}[2]	5.7 (54.6)	5.6 (53.4)
$I / \sigma I$	21.5 (2.7)	23.4 (3.1)
Completeness, %	99.8 (99.4)	100.0 (99.9)
Redundancy	6.1 (5.8)	6.2 (6.0)
Refinement		
Resolution, Å	40.93-2.58	48.30-2.20
No. reflections	22,207	37,260
R_{cryst} / R_{free}[3]	17.74/24.76	18.19/22.70
No. atoms		
Protein	4,403	4,403
Ligand/ion	37	1
Water	109	233
B-factors		
Protein	48	46
Ligand/ion	38	99
Water	37	41
Bond lengths rmsd, Å	0.007	0.007
Bond angles rmsd, °	1.087	1.004

[1] Values in parentheses are for highest-resolution shell. Each structure was determined using data from a single crystal.

[2] R_{merge} = $S_{hkl}(S_i|I_{hkl,i} - <I_{hkl}>))/S_{hkl,i}<I_{hkli}>$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices h, k and l, and $<I_{hkl}>$ is the mean intensity of that reflection.

[3] R_{cryst} = $S||F_{obs, hkl}| - |F_{calc, hkl}||/|F_{obs, hkl}|$, where $|F_{obs, hkl}|$ and $|F_{calc, hkl}|$ are the observed and calculated structure factor amplitudes. R_{free} is equivalent to R_{cryst} but calculated with reflections (5%) omitted from the refinement process.

In vitro glucosyltransferase assay

     UDP-[^{14}C]glucose (250 mCi/mmol) was obtained from PerkinElmer. GTD (6.2 nM) and GTP binding protein (2 µM) were mixed with 24 µM UDP-[^{14}C]glucose in a buffer containing 50 mM HEPES, pH 7.5, 100 mM KCl, 1 mM $MnCl_2$, 2 mM $MgCl_2$, and 0.1 mg/mL BSA. The total reaction volume was 10 µL. The reactions were incubated at

37°C for 1 h. The reaction was stopped by the addition of loading buffer and heating, and the proteins were separated by SDS-PAGE. Glucosylation of GTPase was analyzed by phosphorimaging. For graphical representation, band densitometry was measured with Image J software, and the band intensities were normalized with RhoA modified by TcdA GTD set at 100%.

In vitro autoprocessing of TcdA and TcdB

TcdA and TcdB (4 µM) were mixed with 10 mM InsP6 and 50 mM ethylenediaminetetraacetic acid (EDTA) in a buffer consisting of 100 mM NaCl, 20 mM Tris, pH 8. The samples were incubated 2 h at 37°C, and then dialyzed against 100 mM NaCl, 1mM MnCl$_2$, 20 mM Tris, pH 8. A fraction of the samples was subjected to SDS-PAGE (Figure 4-1) The proteins were diluted in the glucosyltransferase buffer and used at a final concentration of 6.2 nM.
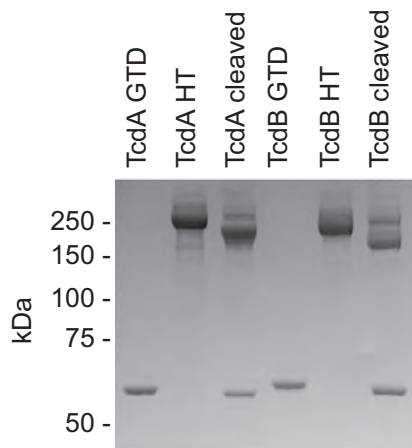
Figure 4-1. Coomassie stained gel showing the GTDs, holotoxins, and cleaved holotoxins used to assay the glucosylation of RhoA.

Generation of HeLa cells expressing FLAG-Rap2A

Cells were cultured in DMEM supplemented with 10% FBS. Flag pLJM1 Rap2A (196) was obtained from Addgene. Vesicular stomatitis virus-G-pseudotyped, human immunodeficiency virus-1 based lentiviruses were generated in human embryonic kidney 293 cells by cotransfection with pHCMV-G (197) and pHCMVΔR8.91 (198). After, 40 h the virus containing supernatant was filtered and used to transduce HeLa S3 cells in the presence of 8 μg/mL polybrene. Transductants were selected with 2 μg/mL puromycin and maintained with 0.5 μg/mL puromycin. Expression of Rap2A was confirmed by western blotting for FLAG.

Pull down of FLAG-Rap2A

Cells were plated to 10 cm dishes and grown to confluency. The cells were then treated with buffer, TcdA (10 nM), or TcdB (10 nM) in 10 mL DMEM with 10% FBS. After 3.5 h, the cells were washed three times with PBS and placed on ice. 1 mL lysis buffer (150 mM NaCl, 50 mM Tris, pH 7.5, 1% Triton X-100, 10 μL/mL protease inhibitor cocktail, Sigma P8340) was added to each plate for 30 minutes. Cells were scraped from the plates, and the lysates were clarified by centrifugation. The supernatants were incubated with 25 μL FLAG M2 resin (Sigma A2220) for 2.5 h at 4°C. The resin was washed three times with 50 mM Tris, pH 7.5, 150 mM NaCl, 1% Triton X-100. FLAG-Rap2A was eluted at 30°C for 30 minutes with 45 μL 50 mM Tris, pH 7.5, 500 mM NaCl, 0.5% CHAPS, 5 mg/mL FLAG peptide.

Mass spectrometry

Proteins were resolved by SDS-PAGE and visualized by Coomassie blue staining (SimplyBlue SafeStain, Invitrogen). Protein bands of interest were excised and cubed, equilibrated in 100mM NH_4HCO_3, reduced with DTT (1/10 volume in 100 mM

NH4HCO3, 37°C for 15 min), and alkylated with iodoacetamide (1/10 volume in 100 mM

NH_4HCO_3, 15 min). After dehydration with acetonitrile, the gel cubes were rehydrated

with 15 μL of 25 mM NH_4HCO_3 containing 0.01 μg/μL trypsin protease (Modified Trypsin-

Gold, Promega), and digestion was carried out for >2 h at 37°C. Peptides were extracted

with 60% acetonitrile, 0.1% trifluoroacetic acid, dried by vacuum centrifugation, and

reconstituted in 10 μL of 60% acetonitrile, 0.1% trifluoroacetic acid. 0.5 μL were applied

to a stainless steel target and mixed with 0.5 mL of α-cyano-4-hydroxycinnamic acid

matrix (5 mg/mL, supplemented with 1 mg/mL ammonium citrate). Matrix-assisted laser

desorption/ionization, time-of-flight (MALDI-TOF) MS and tandem TOF/TOF MS/MS

were carried out using a Voyager 4700 mass spectrometer (Applied Biosystems),

operated in positive-ion reflectron mode. Each MALDI-TOF mass spectrum was

calibrated to within 10 ppm using trypsin autolytic peptides present in the sample (*m/z*

842.50, 1045.56 and 2211.10).

Results

Structure of the TcdA GTD

TcdA GTD crystal structures were determined in the presence and absence of

UDP-glucose at 2.6 Å and 2.2 Å resolution, respectively (Figure 4-2, 4-3, and Table 4-1,

pdb 3SRZ and 3SS1). As observed in GTD structures from other LCTs, the molecule is

composed of a core GT-A fold surrounded by multiple helical projections (150). The N-

terminal projection (Figure 4-2A) is thought to act as a membrane localization domain

(MLD), targeting the GTD to the site of membrane-bound GTPases (152,153). The other

projections at the top right and top left of the GTD 'front' (Figure 4-2A) are thought to

contribute to GTPase substrate specificity (150).

85
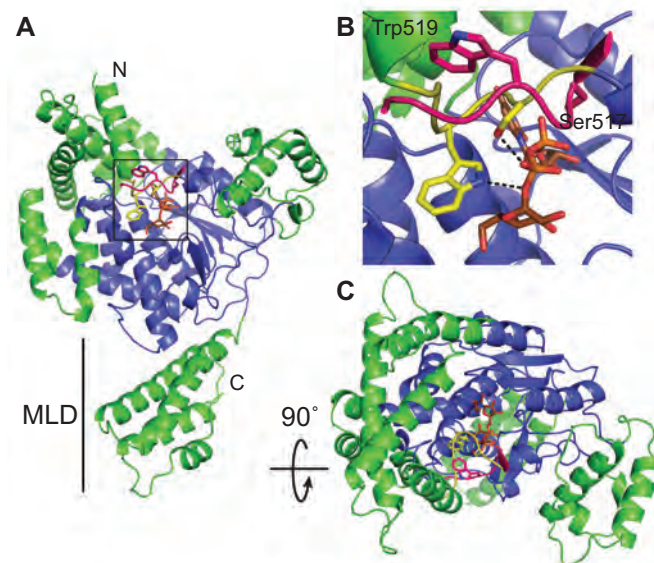
Figure 4-2. Structure of the TcdA GTD. (A-C) The UDP-glucose bound TcdA GTD is shown with the core GT-A family fold in blue and the α-helical protrusions from the fold in green. The mobile 516-522 loop is in yellow, and the loop from the apo structure is superimposed on the structure in pink. Ser517, Trp 519, and UDP-glucose are represented as sticks. Panel (B) shows a close up view of the 516-522 loop and panel (C) corresponds to a 'top' view (rotated 90˚ from the 'front' view shown in panel (A)).

Comparison of the apo- and UDP-glucose-bound structures shows a significant difference in the position of the 516-522 loop (Figure 4-2, Figure 4-3). This loop contains a conserved serine residue, Ser517 in TcdA, which forms a hydrogen bond with the β-phosphate group in UDP-glucose, and a conserved tryptophan residue, Trp519 in TcdA, which forms a hydrogen bond to the glycosidic oxygen (Figure 4-2B and Figure 4-4). In the apo- structure, the loop is moved such that Trp519 is located ~10 Å away from its position in the UDP-glucose-bound structure. A similar conformational difference has been noted in a comparison of the apo- structure of *C. novyi* Tcnα GTD and TcdB GTD bound to a hydrolyzed substrate (151). As described in the mammalian glycosyltransferases involved in carbohydrate synthesis (199), the loop acts as a 'lid' covering the bound UDP-glucose when viewed from the 'top' (Figure 4-2C).
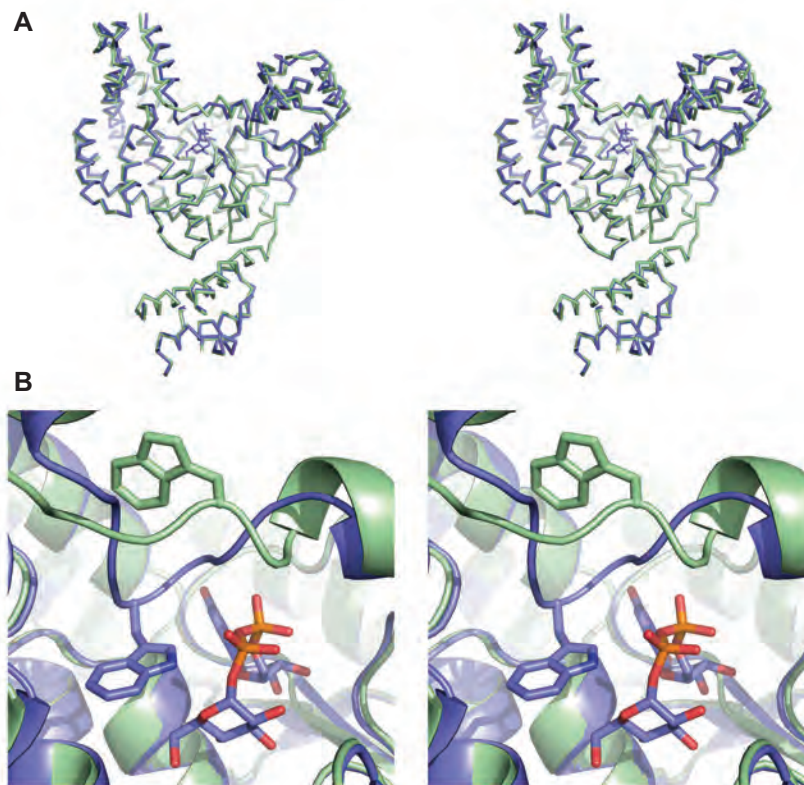
Figure 4-3. Comparison of the apo- and UDP-glucose bound- TcdA GTD structures. (A) The apo- (green) and UDP-glucose bound- (blue) structures are shown in stereo as ribbon diagrams. The two structures are very similar, with an rmsd of 1 Å. (B) A close up view of the 516-522 loop is shown with Trp519 and UDP-glucose represented as sticks. Trp519 forms hydrogen bonds with the glucose in the UDP-glucose bound structure. In the apo- structure, however, Trp519 is moved ~10 Å away.
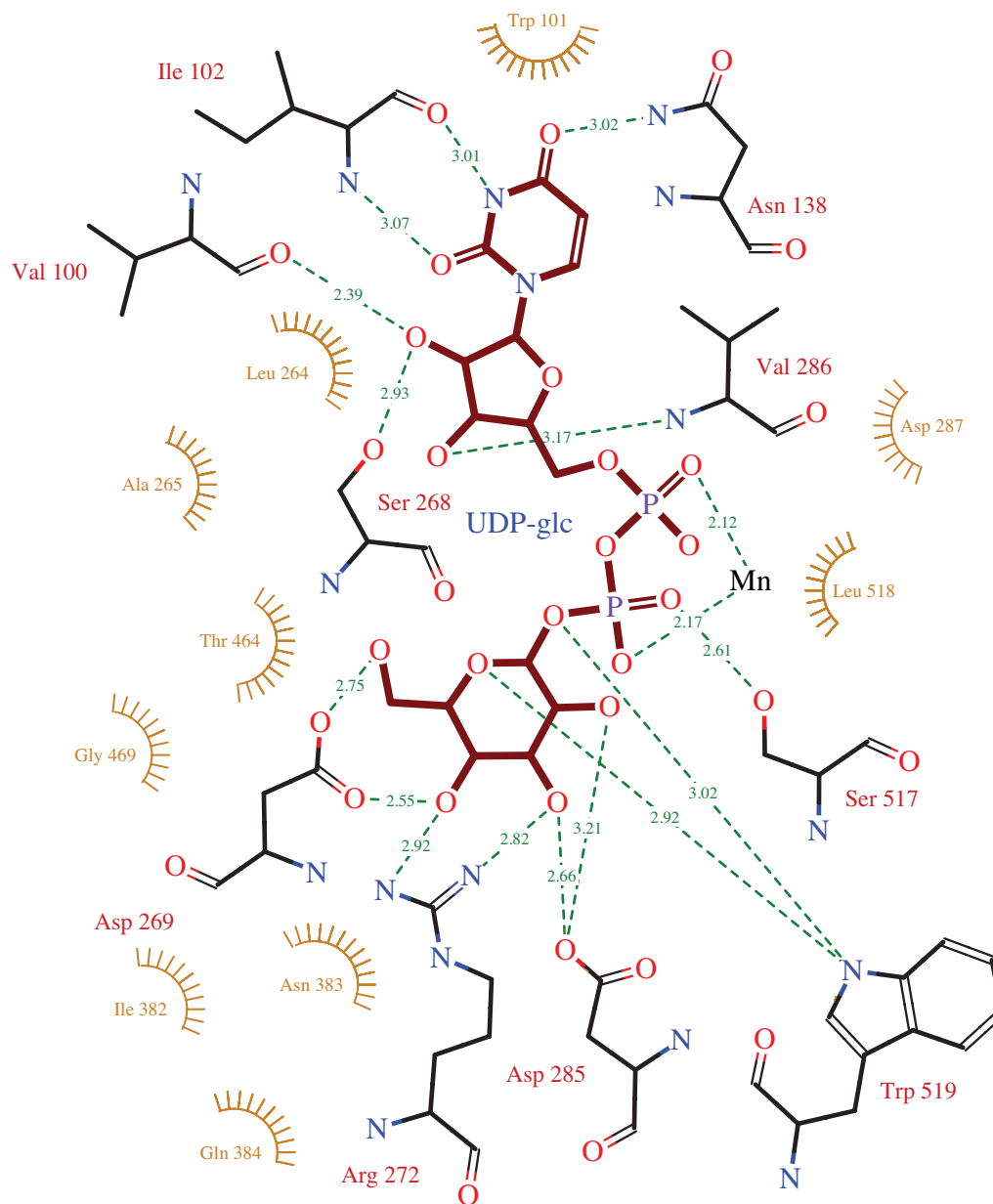
Figure 4-4. The coordination of UDP-glucose by the TcdA GTD in pdb 3SRZ. UDP-glucose is shown as red sticks. Residues making hydrogen bonding contacts are shown in black sticks. Bond distances are shown in angstroms. Hydrophobic interactions are shown in orange. The image was generated by LIGPLOT (200).

One difference between the TcdA and TcdB GTD structures is that the UDP-glucose is intact in the TcdA-GTD structure, whereas in the TcdB-GTD structure it is hydrolyzed (Figure 4-5). While consistent with previously published studies showing that TcdB has a higher rate of UDP-glucose hydrolysis than TcdA (84,201), the molecular explanation for this difference is not apparent from the structure. Because TcdA was previously reported to have a much lower glucosyltransferase activity than TcdB (84), we expected to see differences in the positions of the catalytic residues. Yet, other than the difference in the hydrolysis of the UDP-glucose, the enzymatic core is surprisingly similar between TcdA and TcdB. The residues and waters involved in UDP-binding and catalysis are highly conserved and the binding of UDP-glucose is nearly identical (Figure 4-5B).

Although the core structures of the TcdA and TcdB GTDs are conserved (Figure 4-5A), the surface residues of these two enzymes are highly divergent. This is particularly notable on the 'front' GTPase-binding surface adjacent to the UDP-glucose site (154). Amino acid changes in this region result in a significant change in the electrostatic properties of the surface and suggest that the TcdA and TcdB GTDs could have different substrate specificities within the cell (Figure 4-5C-D). In addition, there are significant differences in the electrostatic potential properties of the MLD. The TcdB MLD is markedly more charged than that of TcdA. The 'front' surface (as shown in Figure 4-5C-D, left panel) is dominated by a highly basic patch, while the opposite face is almost entirely acidic. TcdA has a smaller basic area on the front of the MLD, and lacks the negatively charged patch on the back. These differences could further differentiate the activities of the TcdA and TcdB GTDs in cells.
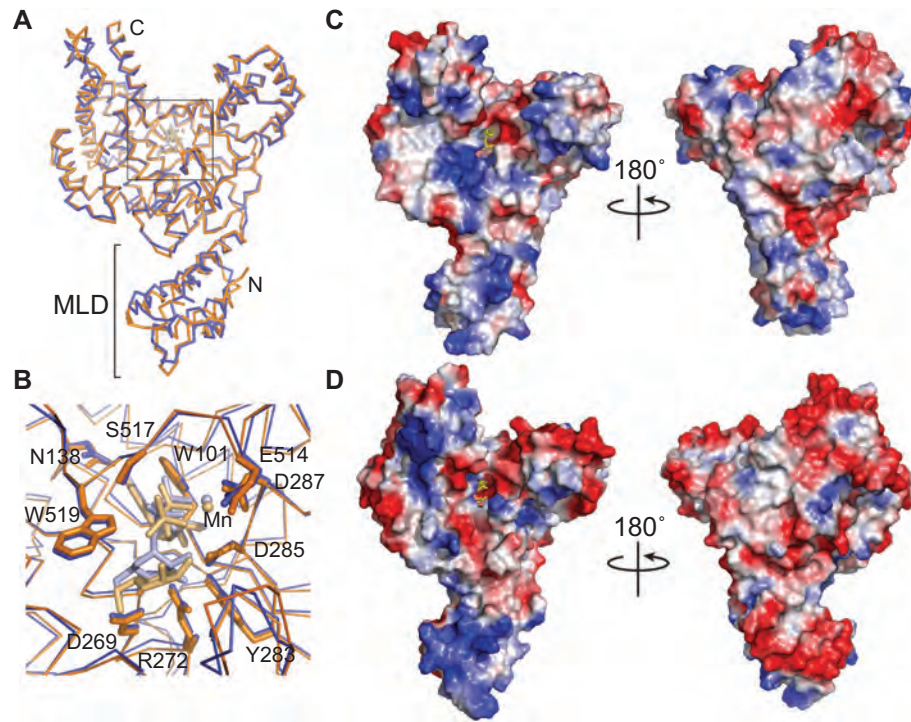
Figure 4-5. Comparison of the TcdA and TcdB GTDs. (A) The aligned structures of the TcdA (blue) and TcdB (orange) GTDs are shown as backbone traces. (B) A close up view of the catalytic site with UDP-glucose coordinating residues shown as sticks. UDP-glucose (light blue), UDP (light orange), and glucose (light orange) are also shown as sticks. Manganese ions are shown as small spheres. Although the cores of the GTDs are conserved, the surfaces are highly divergent (C-D). The electrostatic surface potentials of the (C) TcdA GTD and the (D) TcdB GTDs are shown with positively charged surfaces colored blue and negatively charged surfaces in red.

Glucosyltransferase activity of the TcdA and TcdB GTDs

The striking differences between the TcdA and TcdB surfaces led us to perform a side-by-side comparison of TcdA and TcdB GTD activity against a panel of Rho and Ras family GTPase substrates. TcdA and TcdB GTD were incubated with purified GTPases in the presence of radiolabeled UDP-glucose. Transfer of glucose was detected by SDS-PAGE followed by phosphorimaging. Surprisingly, we found that TcdA and TcdB glucosylated comparable amounts of RhoA, Rac1, and Cdc42 (Figure 4-6). Consistent with a previous report (84), we observed that TcdA could glucosylate Rap1A and Rap2A, whereas TcdB did not.
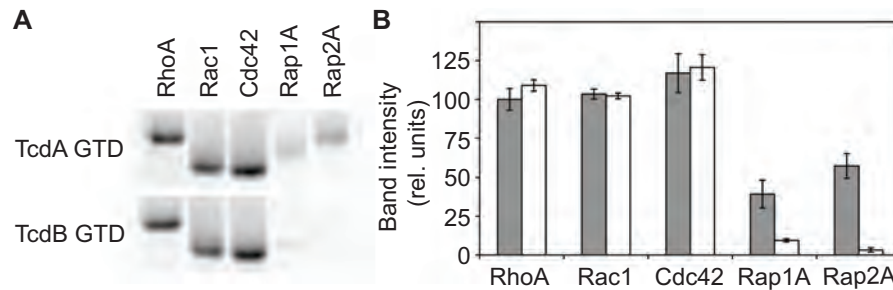
Figure 4-6. Glucosyltransferase activity of the TcdA and TcdB GTDs. (A) TcdA and TcdB GTDs (6.2 nM) were tested for their ability to glucosylate a panel of GTPases (2 μM). (B) The band intensities for the representative experiment shown in (A) were quantified by densitometry. The means ± s. d. from four independent replicates are shown with TcdA in gray bars and TcdB in white bars. The data are scaled with the average value for RhoA modified by TcdA GTD set at 100%.

Enhanced glucosyltransferase activity following autoprocessing

The highly conserved enzymatic core observed in the structure of the TcdA GTD provides no evidence that it is a deficient enzyme as compared to TcdB. Consistent with this view, our *in vitro* studies indicate that the GTDs of TcdA and TcdB modify similar amounts of RhoA, Rac1, and Cdc42. Previous studies on the glucosyltransferase activity of TcdA have used only TcdA holotoxin (HT). In cells, however, the GTD is released and this isolated domain is thought to traffic to the membrane where it acts on host GTPases (132,152). We wondered whether using GTD versus HT might result in a difference in activity. Therefore, we tested the ability of GTD or HT for both TcdA and TcdB to modify RhoA (Figure 4-7). As with the isolated GTDs, we observed similar levels of substrate glucosylation when the TcdA and TcdB HTs were compared to each other. However, for both TcdA and TcdB, the HT modified less substrate than the free GTD. Therefore, in the context of the HT, the glucosyltransferase activity of TcdA and TcdB is somehow inhibited. Consistent with this observation, we show that releasing the GTD by initiating *in vitro* autoprocessing results in enhanced substrate modification, comparable to that of the GTD alone (Figure 4-7, Figure 4-1).
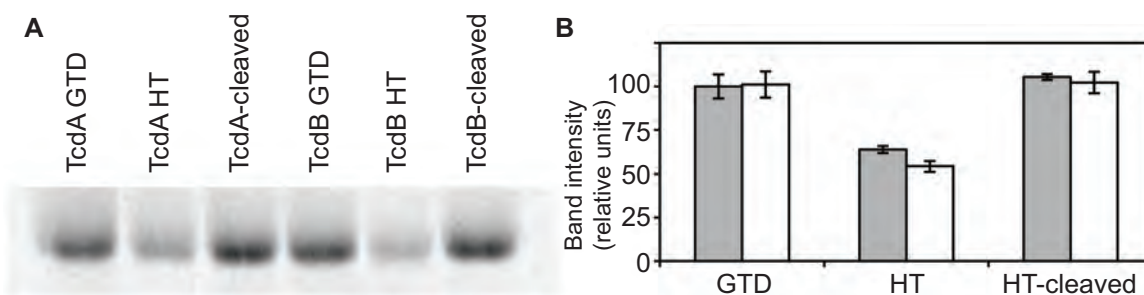
Figure 4-7. Enhanced glucosyltransferase activity following release from the holotoxin. (A) Recombinant RhoA (2 µM) was incubated with UDP-[^{14}C]-glucose (24 µM) and the isolated GTD, holotoxin, or cleaved toxin of TcdA or TcdB (6.2 nM of each). The proteins were resolved by SDS-PAGE, and the gels were analyzed by phosphorimaging. (B) The band intensities were measured from three replicate experiments. TcdA is shown in gray bars and TcdB in white (n=3, mean ± s. d). The data are scaled with the average value for RhoA modified by TcdA GTD set at 100%.

Structure of the TcdA GTD in the context of the holotoxin

Because the HT has a lower activity than the GTD (Figure 4-7), we sought to understand the structural determinants of the glucosyltransferase activity in the context of the holotoxin. Our lab has recently determined the structure of TcdA HT at 25 Å resolution using electron microscopy and random conical tilt (Figure 4-8A) (194). The structure contains a large bi-lobed head with two extensions. We have shown that the larger of these extensions contains the receptor-binding domain, and the smaller one contains the GTD (194). Figure 4-8 shows the EM structure of TcdA with a model of the TcdA receptor-binding domain (green) and the structure of the TcdA GTD (blue) placed into the density. The GTD fits into the map with the N-terminal MLD pointing away from the head and fitting into a narrower region of density. The other helical projections of the GTD fit into density 'flaps' that project toward the receptor-binding domain and provide confidence that the general orientation of the GTD is correct.
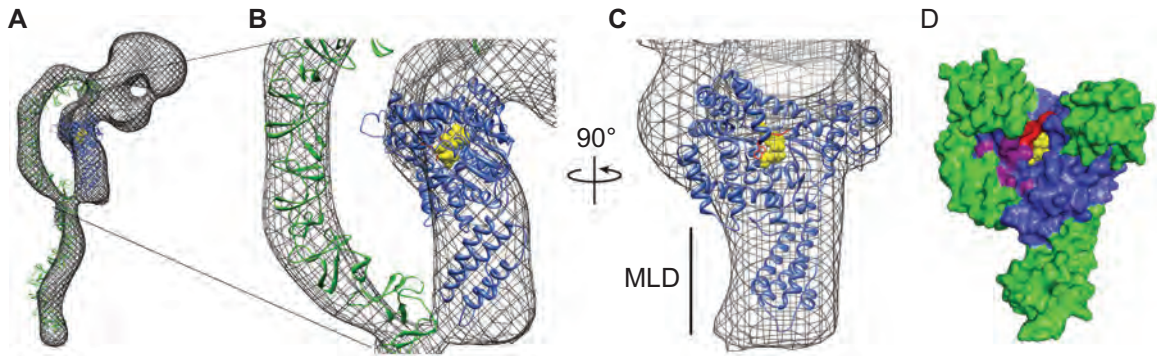
Figure 4-8. Structure of the TcdA GTD alone and in the context of the TcdA holotoxin structure. (A-C) The 3D reconstruction of TcdA (194) filtered to 25 Å is shown as a mesh surface with the crystal structure of the TcdA GTD (blue) and a model of the TcdA receptor-binding domain (115) (green) placed into the density. UDP-glucose is shown as yellow spheres. In panel (C), the model of the binding domain and the corresponding map density are removed. (D) Surface of the TcdA GTD shown in the same orientation as in panel (C). The core GT-A fold is shown in blue with the additional α-helical regions in green (as in Figure 4-2). The 516-522 loop is colored red, and UDP-glucose is represented as yellow spheres. Amino acids Lys448, Gln454, Glu460, Arg462, and Gly471 are shown in purple. The corresponding residues in TcdB have been shown to be involved in substrate binding (154).

Oriented in this way (Figure 4-8), the 'front' surface involved in GTPase-binding faces the receptor-binding domain. We predict that the presence of the receptor-binding domain sterically inhibits binding of GTPases in the context of the holotoxin. In addition, the model suggests that the 'top' surface containing the mobile 516-522 loop will be occluded by the autoprocessing domain in the context of the holotoxin. This occlusion could affect the position of the 516-522 loop involved in UDP glucose binding, hydrolysis, and transfer.

Glucosylation of Rap2A in cells

The increased activity of TcdA and TcdB following autoprocessing highlights the importance of the cellular context for cellular activity. Since our *in vitro* assay indicated that Rap proteins were modified at lower levels than Rho-family proteins, we wanted to test whether TcdA-mediated Rap modification could occur in the context of cellular

intoxication. We generated a HeLa cell line that stably expresses FLAG-tagged Rap2A by lentiviral transduction. The cells were either mock-treated or treated with 10 nM TcdA or TcdB. After 3.5 h, the cells were harvested, and the Rap2A was pulled down and subjected to SDS-PAGE. The bands were extracted from the gel, digested with trypsin, and analyzed by MALDI-TOF/TOF mass spectrometry. A peptide covering Rap2A amino acids 32-41 (m/z 1318.60 Da) was observed for Rap pulled down from mock-treated, TcdA-treated, and TcdB-treated cells (Figure 4-9A-C). A related peptide (m/z 1480.65) was observed only for pull downs from TcdA-treated cells, consistent with glucosylation (+162 shift) of Thr35 (Figure 4-9B). TOF/TOF fragmentation spectra of both species were consistent with the predicted sequence, and mapped the site of glucosylation to Thr35 (Figure 4-9D-E). These results closely resembled data obtained from *in vitro* glucosylation of Rap2A (Figure 4-10). Therefore, TcdA is not only capable of glucosylating Rap2A *in vitro*, but can also glucosylate it in cells.
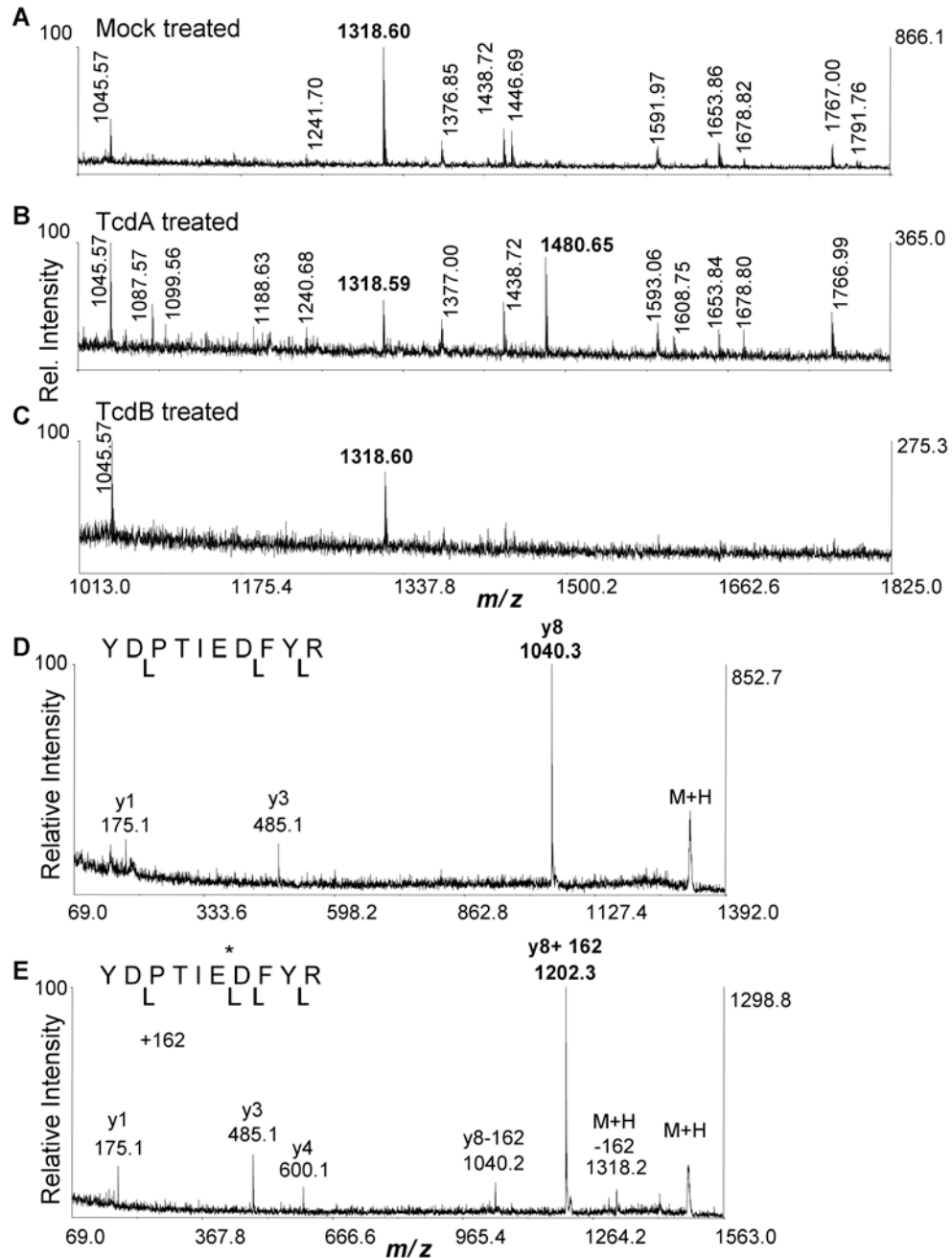
Figure 4-9. Glucosylation of Rap2A in cells treated with TcdA. (A-E) MALDI-TOF/TOF mass spectrometry indicating the glucosylation of the peptide YDPTIEDFYR. (A-C) A portion of the MALDI-TOF peptide mass map (*m/z* 1013-1825) is shown to highlight the diagnostic singly-charged peptide ions at *m/z* 1318.60 and 1480.65 (labeled in bold) that represent the peptide in the native and glucosylated state, respectively. The *m/z* 1480.65 is only found in the TcdA-treated samples (B). MALDI-TOF/TOF fragmentation spectra are shown for the *m/z* 1318.60 peptide (D) and for the glucosylated *m/z* 1480.65 form (E). Labeled y-ions are denoted by cleavage brackets below the sequence. The y8 fragment ion that contains Thr35 is diagnostic for the modification, which adds 162 Da. Ions are also observed that are consistent with neutral loss of glucose from the y8 and M+H ions.
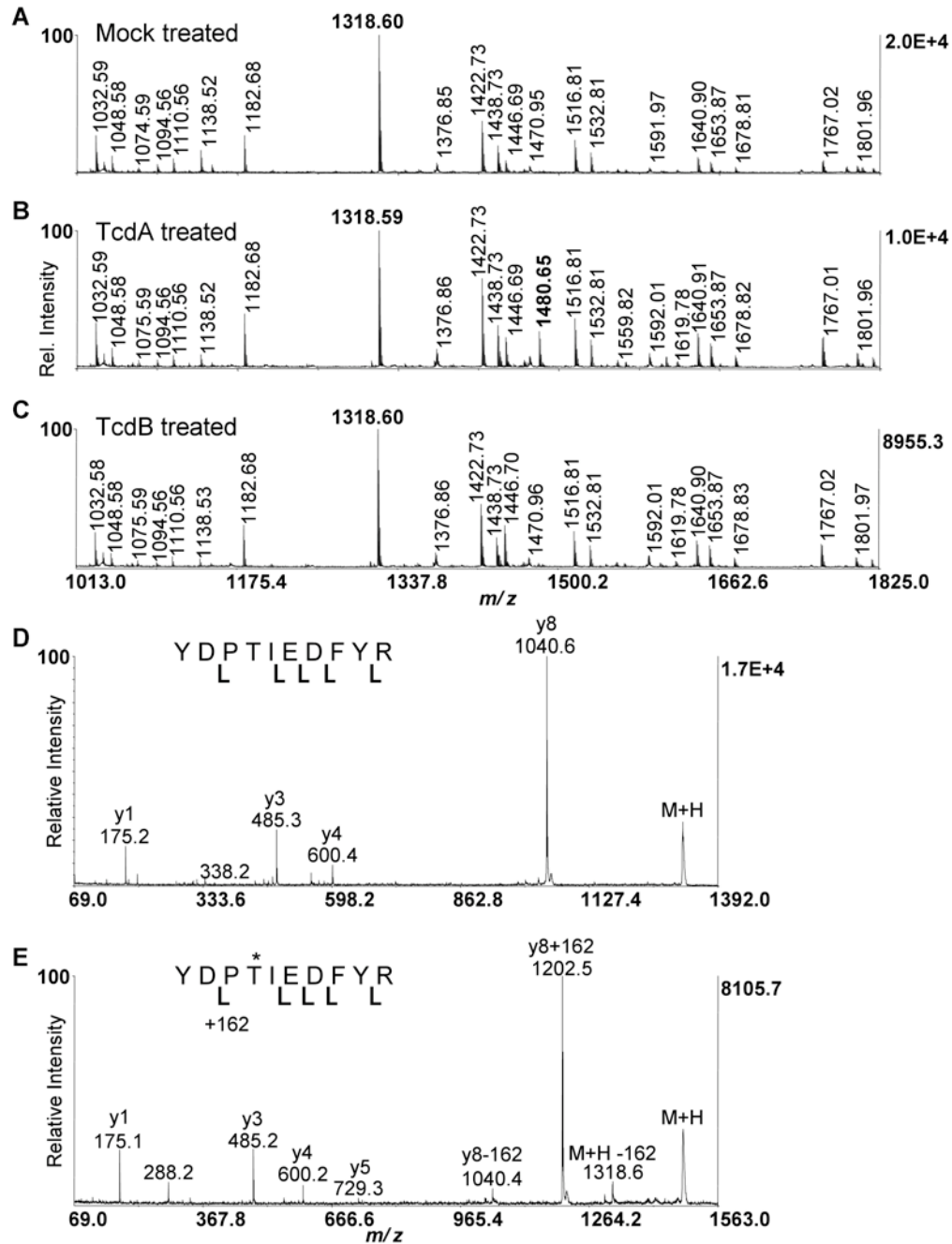
Figure 4-10. Detection of Rap2A glucosylation *in vitro* by mass spectrometry. Rap2A (3.4 µM) was incubated with 1.4 µM TcdA-GTD, 1.4 µM TcdB-GTD, or buffer in the presence of 100 µM unlabeled UDP-glucose for 2 h at 37°C. The samples were analyzed by mass spectrometry. (A-E) Glucosylation of the Thr35 containing peptide YDPTIEDFYR by TcdA GTD is demonstrated by MALDI-TOF/TOF mass spectrometry. (A-C) A portion of the *m/z* range from 1013-1825 is shown to highlight the diagnostic peptides at *m/z* 1318.60 and 1480.65 (labeled in bold) that represent the peptide in the native and glucosylated state, respectively. The *m/z* 1480.65 that is only found in the TcdA-treated samples (B). MALDI-TOF/TOF fragmentation spectra are shown for the *m/z* 1318.60 peptide (D) and for the glucosylated *m/z* 1480.65 form (E). Labeled y-ions are denoted by cleavage brackets below the sequence.

96

Discussion

We initiated this study with the goal of understanding differences in TcdA and TcdB glucosyltransferase activity and substrate specificity. The glucosylation of RhoA, Rac1, and Cdc42 by TcdB has been well-characterized (77,150,154,202), however, much less is known about TcdA and the interaction of TcdA with these substrates. One of the few studies directly comparing the enzymatic activities of TcdA and TcdB reports that TcdB is ~100x more active than TcdA at glucosylating RhoA, Rac1, and Cdc42 *in vitro* and has a higher rate of UDP-glucose hydrolysis in the absence of substrate (84). These studies were done with the holotoxins and not with the isolated domains that are released into the cell.

Here we present crystal structures of the TcdA GTD with and without its co-substrate UDP-glucose at 2.6 Å and 2.2 Å, respectively. Because TcdA was reported to have a much lower glucosyltransferase activity than TcdB (84), we expected to see differences in the residues involved in UDP-binding and catalysis. Yet, the structure reveals that these residues are highly conserved, not only in identity, but also in their position within the GTD (Figure 4-5B).

The conservation of the core of the GTDs led us to reinvestigate their activity towards a panel of GTPase substrates. Using experimental conditions similar to those of previous reports, we found that TcdA and TcdB were equally active at glucosylating RhoA, Rac1, and Cdc42 (Figure 4-6). These results contradict the previous report that TcdB is 100 times more potent as an enzyme (84). We hypothesized that this difference might be because previous studies used HT, whereas we used isolated GTD. Upon testing the ability of GTD and HT to modify RhoA, we found that the glucosyltransferase activity is inhibited in the context of the HT. The activity of the HT can be restored through the initiation of autoprocessing, an InsP6-induced proteolysis event that

97

releases the GTD from the rest of the toxin (134,135). Our data indicate that TcdA and TcdB, in each of their three structural states, modify comparable amounts of RhoA, Rac1, and Cdc42. We now know that TcdB undergoes autoprocessing much more readily than TcdA (193), and, in our hands, TcdB is also more sensitive to degradation. A gel documenting the cleavage state of our reagents is included as Figure 4-1. It is possible that in the previous study TcdB was more active than TcdA because the GTD was released by proteolysis, whereas TcdA remained locked in a less active conformation. Our findings suggest that the ~1000 fold difference in cytopathic potency between TcdA and TcdB is not due to differences in the intrinsic glucosyltransferase activity, but is more likely due to differences in binding and/or delivery of the GTD into the target cell.

Although the TcdA and TcdB GTDs were equally active in our assays towards RhoA, Rac1, and Cdc42, TcdA was able to glucosylate the additional substrates Rap1A and Rap2A. To test whether this family of substrates was modified within cells treated with holotoxin, we generated a stable HeLa cell line expressing FLAG tagged Rap2A. Following treatment with TcdA HT, we lysed the cells and pulled down Rap2A. The protein was digested with trypsin protease and analyzed by tandem mass spectrometry. We were able to confidently identify a glucosylated peptide (aa 32-41) consistent with modification of Thr35. Thus, TcdA is capable of modifying Ras family substrates not only *in vitro*, but also in cells. Modified Rap2A peptides were not identified in mock-treated samples or samples treated with TcdB.

While TcdB and C. novyi α-toxin (Tcnα) target only Rho family substrates, TcdA and C. sordellii lethal toxin (TcsL) can glucosylate at least some members of both the Rho and Ras families. When the structures are compared, TcdA and TcsL have some striking similarities on the putative substrate-binding surface, most notably a large positively charged pocket adjacent to the UDP-glucose binding pocket (Figure 4-11). In

TcdB and Tcnα, however, an acidic patch replaces this basic pocket. It is important to note that these characteristics are not merely due to relatedness of the proteins. In fact, the GTDs of TcdB and TcsL are more closely related (75% identity) than the GTDs of TcdA and TcsL (47%).

TcdA	TcdB	TcsL	Tcnα
Rho, Rac, Cdc42, Rap	Rho, Rac, Cdc42	Rac, Ras Ral, Rap	Rho, Rac, Cdc42
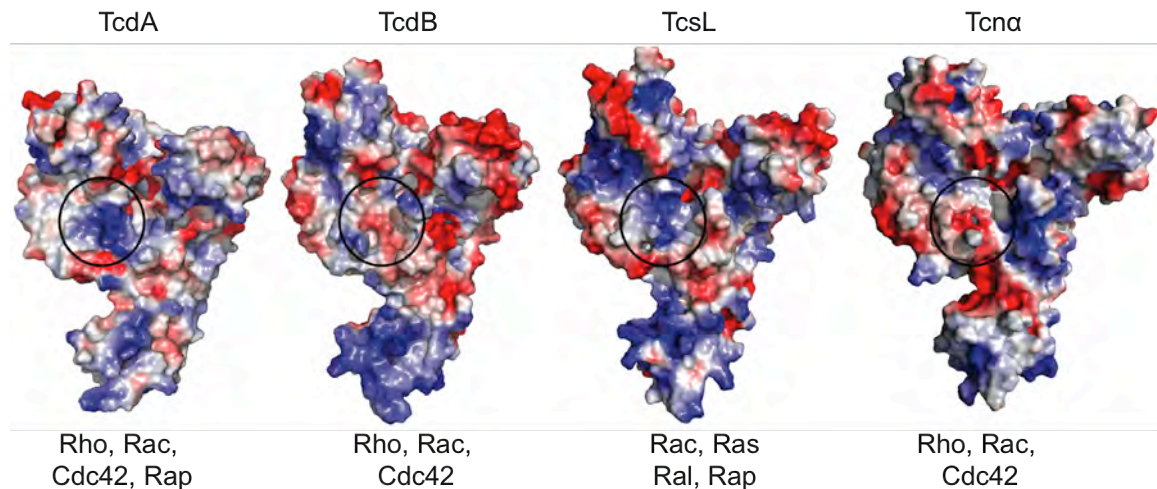
Figure 4-11. The electrostatic surface potential is shown for TcdA, TcdB, TcsL, and Tcnα. Postively charged surfaces are shown in blue. Negatively charged surfaces are shown in red. The circled region marks the region corresponding to the basic pocket in TcdA.

The ability of TcdA to modify Ras family substrates has important implications for its role in pathogenesis. Rap proteins are known to be important in the regulation of cell-cell junctions (52). Thus, inactivation of Rap may be important for disrupting the integrity of the intestinal epithelium. Other Ras family members are involved in complex pathways regulating cell proliferation and survival (52). It is interesting to note that all TcdB sequences from the pathogenic TcdA⁻TcdB⁺ strains characterized thus far have mutations within the GTD that allow it to modify Ras family substrates including Rap (53,54). It is tempting to speculate that modification of Ras substrates by TcdA is a key process in CDAD, and that the mutated TcdB in TcdA⁻TcdB⁺ strains allows it fulfill the

role of TcdA in its absence. A more complete understanding of the glucosyltransferase activities of these two toxins and downstream effects of glucosylation of specific substrates will be essential in understanding the molecular mechanisms important in CDAD.

CHAPTER V

CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

As the primary virulence factors of *C. difficile*, TcdA and TcdB are ideal therapeutic targets for treatment of CDAD. While a detailed understanding of toxin action is needed, a model of the molecular mechanism of cellular intoxication is far from complete. Structural information can greatly inform the mechanistic details of protein function. When we initiated this study, the only structural data for TcdA and TcdB were X-ray crystal structures of the TcdB GTD and a fragment of the TcdA RBD (Figure 5-1). In my thesis work, I have determined several new and important structures including high-resolution structures of the TcdA autoprocessing and glucosyltransferase domains and low resolution structures of the TcdA holotoxin.

In Chapter II, I have presented the 1.6 Å X-ray crystal structure of the TcdA CPD, which is responsible for InsP6-induced auto-processing (173). InsP6 is bound in a highly basic pocket that is separated from an unusual active site by a β-flap structure. Functional studies confirmed an intra-molecular mechanism of cleavage and highlighted specific residues required for InsP6-induced TcdA processing. Analysis of the structural and functional data in the context of sequences from similar and diverse origins revealed a C-terminal extension and a π-cation interaction within the β-flap that appear to be unique among the large clostridial cytotoxins.

In Chapter III, I described the structures of the TcdA and TcdB holotoxins (194). I imaged the toxins using negative stain electron microscopy and showed that these molecules are similar in structure. A three-dimensional structure of TcdA was

determined at 24 Å by random conical tilt, and the organization of the functional domains was mapped within the structure. One interesting feature of the structure is an interaction between the N-terminal part of the GTD and the RBD. This structural arrangement appears to occlude glucosylation of the substrates in the HT. Moreover, in *in vitro* assays, I have observed that autoprocessing is inhibited in the HT relative to the autoprocessing that occurs with the CPD alone. Interdomain contacts such as the GTD-RBD interaction may be involved in this inhibition of enzymatic activity until the appropriate time for activation.
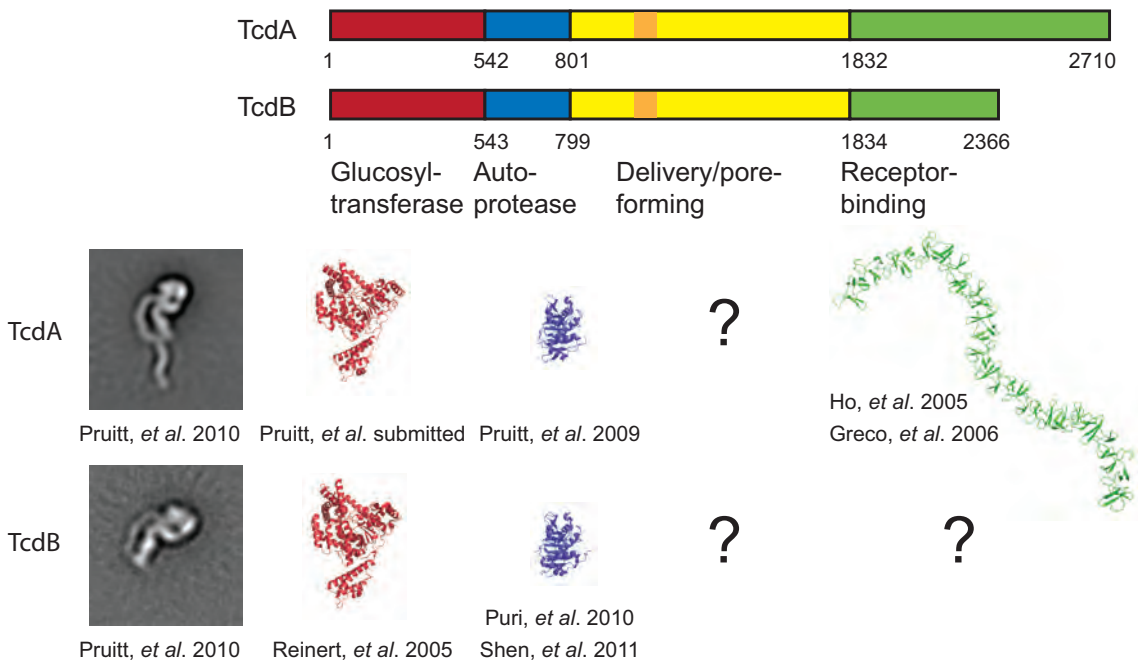
Figure 5-1. Summary of the structural information available for TcdA and TcdB.

The structure of the TcdA HT is important, in part, because it provides a platform to investigate the conformational changes involved in intoxication. One of the expected changes is pH-inducible pore formation. So far, no method has been developed for imaging the LCTs in a pore state. However, we have obtained a structure of TcdA on a

grid following exposure to the endosomal pH. This structure reveals a significant structural change in the delivery domain. These rearrangements may reflect changes involved in pH-inducible pore-formation. In addition, we see an apparent unfolding of a large part of the GTD, a step likely necessary for its translocation through a pore. These structures of TcdA at neutral and acidic pH provide a framework for understanding the complex functions of this class of toxins.

Chapter IV describes the crystal structures of the TcdA GTD in the presence and absence of the co-substrate UDP-glucose. While the enzymatic core is similar to that of TcdB, the proposed GTPase-binding surface differs significantly. I showed that TcdA is comparable to TcdB in its modification of Rho-family substrates and that, unlike TcdB, TcdA is also capable of modifying Ras-family GTPases both *in vitro* and in cells. The glucosyltransferase activities of both toxins are reduced in the context of the holotoxin but can be restored with autoproteolytic activation and GTD release. Based on these structures, I propose a model wherein the receptor-binding domain occludes the binding of GTPase substrates. These studies highlight the importance of cellular activation in determining the array of substrates available to the toxins once delivered into the cell.

Collectively these studies have provided considerable insight into the structure of these toxins, the structural response to acidic pH, the mechanisms of autoprocessing, and the enzymatic activity of TcdA. Still, many questions remain regarding toxin function. What are the receptors for TcdA and TcdB, and what cells do they target *in vivo*? Does D2 have a role in binding to host cells? What is the nature of the pore structure of the toxins? What structural rearrangements are necessary for the toxins to access this pore state? How is the GTD translocated through this pore? How is the autoprotease activity occluded in the HT, and when/how does it be come activated? How is the GTD localized to the plasma membrane, and to which parts of the membrane is it targeted? How does it recognize its substrates? What are the effects of modification of different subsets of

GTPases? Clearly, a great deal of work needs to be done to understand the action of these two toxins. Below, I discuss our planned and ongoing efforts to address a few of these questions.

Future directions

Determine a subnanometer resolution structure of TcdA

We have been able to generate the first structure of the TcdA holotoxin and learn a great deal about its organization by EM using negative stained particles. The use of negative stain, however, has a number of drawbacks such as flattening and potential distortion of the particles. Perhaps more importantly for our studies, it limits the resolution at which the structures can be determined. Using cryo-EM, we could significantly improve the structural resolution and visualize the atomic details of how the TcdA holotoxin is organized. This method traps particles in a layer of vitrified ice and is routinely used to generate 3D structures at subnanometer resolution. As a preliminary experiment, our collaborators, Yoshimasa Takizawa and Melanie Ohi, have obtained images of TcdA in vitreous ice (Figure 5-2). The particles appear to be in random orientations, and although the contrast is lower than what is seen in negative stain, we can clearly distinguish the 'head' and 'tail' domains in the individual particles. Using cryo-EM, we hope to determine the structure of TcdA at < 10 Å resolution.

Cryo-EM is amenable to a number of gold-particle labeling techniques which are not possible in negative stain. By imaging toxins with labeled cysteines, we can more accurately determine the domain organization of the toxin. Cryo-EM will allow us, not only to obtain higher resolution structure of the toxin and more accurately place the domains, but also to assess the structural changes that occur in response to environmental cues such as low pH, InsP6, and reductant.
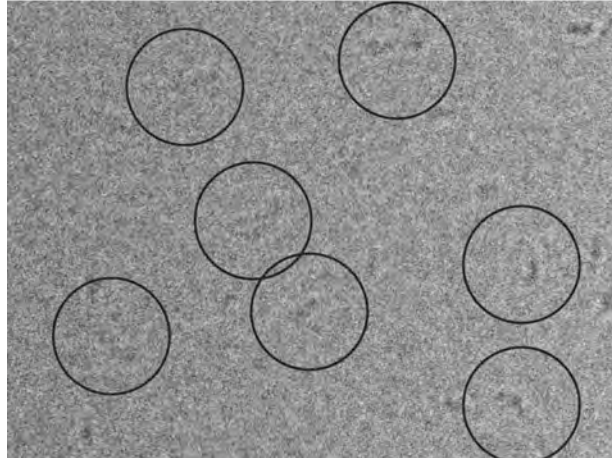
Figure 5-2. TcdA particles in vitreous ice, imaged by transmission EM. Grid preparation and image collection were done by Yoshimasa Takizawa.

As a complementary approach, we can attempt to crystallize TcdA or TcdB holotoxin. When we initiated this study, we could only purify the toxins from *C. difficile* culture supernatants. This did not yield high enough amounts for large-scale crystallization trials. However, we can now express both toxins recombinantly with reasonable yields. We would not be able to assess changes in the toxin upon exposure to environmental cues using crystallography, as we could with EM. Nevertheless, a crystal structure would likely yield higher resolution information, revealing much more about the interdomain contacts and the parts of the toxins for which we have no high-resolution structural information (D1 and D2).

Define the molecular structure of the delivery domain

An important gap in our current understanding of TcdA and TcdB structure is highlighted by the lack of structural models for the central region (Figure 5-1). This region has been dubbed a delivery domain based on its putative role in pore formation and translocation of the GTD. Our observation that this region exists as a bi-lobed structure has led us to consider a hypothesis wherein the central region is composed of

discrete structural and functional domains, D1 and D2 (203). A hydrophobic sequence located within D1 has been hypothesized to span the membrane during pore formation, and very recent studies from Genisyuerek, *et al.* support the idea that this region contains the region essential for forming pores. (131,204). To aid in determining the mechanism of pore-formation, I have attempted to crystallize the TcdA delivery domains, D1 and D2. I have cloned, expressed, and purified amino acids 799-1460 (D1) and 799-1859 (D1-D2) in sufficient amounts for crystallization trials (Figure 5-3A).

Preliminary experiments with the full delivery domain have yielded clusters of microcrystals (Figure 5-3B). Single crystals cannot be isolated from the clusters; nor would they be large enough to use for structure determination. However, this does show that the recombinant D1-D2 domain is structured and capable of crystallization. We are hopeful that further crystallization trials with the protein, or the corresponding protein from TcdB, will yield diffraction-quality crystals.
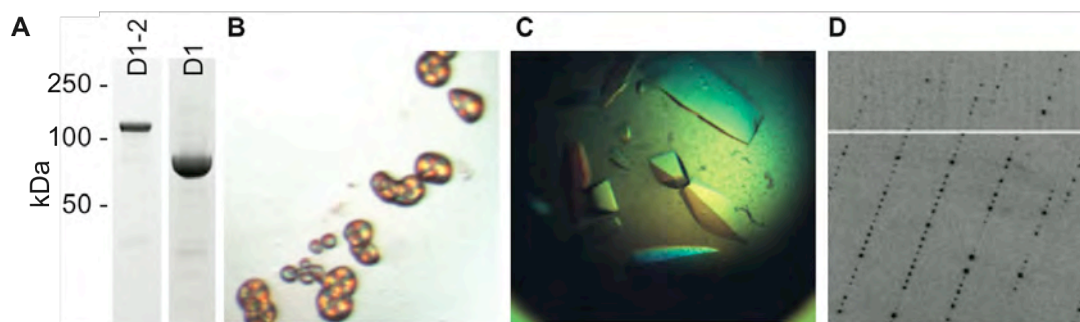
Figure 5-3. Crystallization of the TcdA delivery domain. (A) The TcdA delivery domain (D1-2, residues 771-1859) and TcdA D1 (residues 799-1460) can be expressed recombinantly and purified with high yield. The purified proteins are shown by SDS-PAGE with Coomassie blue staining. (B) Microcrystals have been generated for a ~125 kDa protein corresponding to TcdA D1-2. (C) A construct corresponding to TcdA D1 forms large, reproducible crystals. (D) The TcdA D1 crystals diffract to 7 Å.

Large, reproducible crystals have been obtained with the TcdA-D1 protein (Figure 5-3C). The crystals of TcdA-D1 diffract to 7 Å (Figure 5-3D), and a complete

dataset has been obtained (99.8% complete overall, 96.7% complete in the highest bin). The data are isotropic and display no obvious crystallographic defects as analyzed by Phenix Xtriage (158). We could potentially determine the structure of D1 at 7 Å resolution using these crystals. However, it would be preferable to generate higher quality crystals. Thus, we will try modifying the current crystallization conditions or identifying alternate crystal forms.

Determine the structural features of the TcdA pore

The EM structures at neutral and low pH suggest significant structural change in the delivery domain, but the sequences associated with membrane insertion and pore formation are unknown. Genisyuerek, *et al.* have recently reported that the minimal pore-forming region is contained within residues 830-990. Characterizing the pore state will require a method of getting a homogenous sample of the sample in the pore state in a lipid system. We have made preliminary attempts to prepare samples with the toxin inserted in liposomes. Further studies are necessary to determine whether or not the toxin is actually in a pore state in this system. To test for functional pores we will monitor the release of potassium, chloride, or fluorescent probes such as calcein from preloaded liposomes.

To identify the transmembrane sequence in the pore state we plan to use a liposome-based protease protection assay. Toxin inserted proteoliposomes will be digested with protease, washed, and subjected to mass spectrometry analysis. The identification of extended peptides would suggest that they were protected within the proteoliposome. This approach has recently been used within our lab to identify regions of the *Clostridium botulinum* neurotoxin that are associated with the membrane at low pH (205). Mutational analysis of the residues within the protected region will be used in liposome-based calcein release assays to confirm their importance in pore-formation.

A challenging, but more revealing, objective is to obtain a structure of TcdA or TcdB in the pore state. To do this, we plan to work with our collaborator, Melanie Ohi, to obtain images of the TcdA pore in a lipid system. 2D electron crystallography has been used to determine the structure of many membrane proteins. We propose to generate 2D crystals of TcdA in its pore state in planar lipid bilayers. To aid in this process, I have identified detergents that allow the solubilization of TcdA under low pH conditions. TcdA in detergent micelles will be mixed with lipids, and the detergent will be dialyzed out. Under the right conditions, the TcdA pores will be concentrated within the remaining lipid bilayers and form regular 2D arrays which can be used for structure determination. If we are unable to obtain 2D crystals, we will image single TcdA pores in lipid bilayers, monolayers, or liposomes. These methods may not yield as high of resolution information as 2D crystallography, but will minimally answer the important question of the oligomeric state of the pore.

Elucidate the structural determinants of GTD-GTPases binding and specificity

We plan to use the structure of the TcdA and TcdB GTD to elucidate the structural determinants that account for the difference in activity towards Rap2A. Jank, *et al.* have identified 5 TcdB residues that are involved in binding to GTPases: Glu449, Arg455, Asp461, Lys463, and Glu472 (154). These residues are located adjacent the UDP-glucose binding pocket (Figure 1-7B). Interestingly, these residues are not conserved between TcdA and TcdB. Divergence within these residues contribute to striking differences in electrostatic surface potential on the surface of the GTDs that is known to be involved in GTPase binding (Figure 4-5C,D). For example, Glu449 and Glu472 of TcdB contribute to an acidic patch just below and to the left of the UDP-glucose binding pocket (Figure 4-11, circled). By contrast, in TcdA these residues are Lys448 and Gly471, and they are found in a large basic pocket. I hypothesized that this

108

basic pocket that is in TcdA but not TcdB may be involved in recognition of Rap.

The structures of GTDs from homologous LCTs support this hypothesis. TcdB and Tcnα target only Rho family substrates, whereas TcdA and TcsL can glucosylate members of both the Rho and Ras families (Figure 4-11). When the structures are compared, TcdA and TcsL both have a positively charged pocket (Figure 4-11). TcdB and Tcnα, however, have an acidic patch. The sequences of the TcdB variants which modify Ras family substrates suggest that they will also have a basic pocket similar to TcdA and TcsL. I hypothesize that these differences contribute strongly to the different substrate specificities of the GTDs.
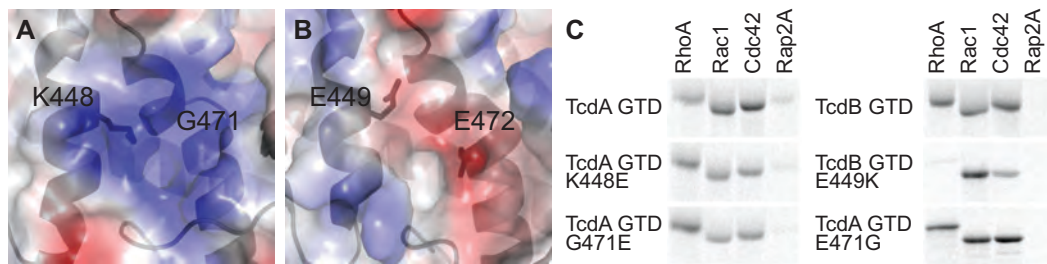
Figure 5-4. Residues of TcdA and TcdB that may contribute to substrate specificity. (A) In the TcdA GTD K448 and G471 are found in a large basic pocket. (B) The corresponding residues of TcdB, Glu449 and Glu472, contribute to an acidic patch. (C) TcdA and TcdB GTD mutants were tested for their ability to modify RhoA, Rac1, Cdc42, and Rap2A. Mutation of TcdA K448 or G471 to glutamate reduces glucosylation of Rac, Cdc42, and Rap2A. Likewise, mutation of TcdB E449 lysine reduces the potency towards its substrates.

In my initial attempts to map residues involved in substrate specificity, I have mutated TcdA Lys448 and Gly471 to glutamate as in the corresponding protein from TcdB. These mutant GTDs have a decrease in their ability to modify Rac1, Cdc42, and Rap2A (Figure 5-4). The corresponding mutation, Glu449Lys in TcdB also decreases modification of substrates, particularly Rho. These experiments suggest that these residues are involved in binding substrates for TcdA, as has already been shown for

TcdB. I expected that mutation of TcdA residues to the identity of the corresponding residues in TcdB would specifically affect Rap modification, but instead they affected the interaction with multiple substrates. These results indicate that the differences in substrate binding between TcdA and TcdB are likely to be complex. I have generated a number of other mutants and combinations of mutants and chimeras to continue trying to define regions involved in Rap-specific substrate binding.

In addition, we will attempt to co-crystallize the TcdA or TcdB GTD in complex with a target GTPase. We have the reagents necessary to prepare large amounts of the purified proteins necessary to begin crystallization trials. A co-crystal structure of a GTD and GTPase would reveal all of the residues involved in the GTD-substrate interaction and facilitate the elucidation of the determinants of substrate binding and specificity for all of the LCTs.

APPENDIX

LIST OF PUBLICATIONS

Pruitt, R. N., Chagot, B., Cover, M., Chazin, W. J., Spiller, B., and Lacy, D. B. (2009) Structure-function analysis of inositol hexakisphosphate-induced autoprocessing in *Clostridium difficile* toxin A. *J Biol Chem* **284**, 21934-21940

Pruitt, R. N., Chambers, M. G., Ng, K. K., Ohi, M. D., and Lacy, D. B. (2010) Structural organization of the functional domains of *Clostridium difficile* toxins A and B. *PNAS* **107**, 13467-13472

Pruitt, R. N., Chumbler, N. C., Farrow, M. F., Seeback, S. A., Friedman, D. B., Spiller, B. W., Lacy, D. B., Structural determinants of the *Clostridium difficile* glucosyltransferase activity. (Manuscript submitted).

BIBLIOGRAPHY

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

2. Snyder, L. S. (1937) Further studies on Bacillus difficilis (Hall and O'Toole). *J Infect Dis* **60**, 223-231.

3. Lyerly, D. M., Krivan, H. C., and Wilkins, T. D. (1988) Clostridium difficile: its disease and toxins. *Clin Microbiol Rev* **1**, 1-18.

4. Kelly, C. P., Pothoulakis, C., and LaMont, J. T. (1994) Clostridium difficile colitis. *N Engl J Med* **330**, 257-262.

5. Finney, J. M. T. (1893) Gastroenterostomy for cicatrizing ulcer of the pylorus. *Bull. Johns Hopkins Hosp.* **4**, 53-55.

6. Kabins, S. A., and Spira, T. J. (1975) Outbreak of clindamycin-associated colitis. *Ann Intern Med* **83**, 830-831.

7. Cohen, L. E., McNeill, C. J., and Wells, R. F. (1973) Clindamycin-associated colitis. *JAMA* **223**, 1379-1380.

8. Sayedy, L., Kothari, D., and Richards, R. J. (2010) Toxic megacolon associated Clostridium difficile colitis. *World J Gastrointest Endosc* **2**, 293-297.

9. Cloud, J., and Kelly, C. P. (2007) Update on Clostridium difficile associated disease. *Curr Opin Gastroenterol* **23**, 4-9.

10. Calfee, D. P. (2008) Clostridium difficile: a reemerging pathogen. *Geriatrics* **63**, 10-21.

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

12. McDonald, L. C., Owings, M., and Jernigan, D. B. (2006) Clostridium difficile infection in patients discharged from US short-stay hospitals, 1996-2003. *Emerg Infect Dis* **12**, 409-415.

13. Redelings, M. D., Sorvillo, F., and Mascola, L. (2007) Increase in Clostridium difficile-related mortality rates, United States, 1999-2004. *Emerg Infect Dis* **13**, 1417-1419.

14. Ozaki, E., Kato, H., Kita, H., Karasawa, T., Maegawa, T., Koino, Y., Matsumoto, K., Takada, T., Nomoto, K., Tanaka, R., and Nakamura, S. (2004) Clostridium difficile colonization in healthy adults: transient colonization and correlation with enterococcal colonization. *J Med Microbiol* **53**, 167-172.

15.   Bartlett, J. G. (1994) Clostridium difficile: history of its role as an enteric pathogen and the current state of knowledge about the organism. *Clin Infect Dis* **18 Suppl 4**, S265-272.

16.   Simor, A. E., Bradley, S. F., Strausbaugh, L. J., Crossley, K., and Nicolle, L. E. (2002) Clostridium difficile in long-term-care facilities for the elderly. *Infect Control Hosp Epidemiol* **23**, 696-703.

17.   Sorg, J. A., and Sonenshein, A. L. (2008) Bile salts and glycine as cogerminants for Clostridium difficile spores. *J Bacteriol* **190**, 2505-2512.

18.   Malamou-Ladas, H., O'Farrell, S., Nash, J. Q., and Tabaqchali, S. (1983) Isolation of Clostridium difficile from patients and the environment of hospital wards. *J Clin Pathol* **36**, 88-92.

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

20.   Fawley, W. N., Parnell, P., Verity, P., Freeman, J., and Wilcox, M. H. (2005) Molecular epidemiology of endemic Clostridium difficile infection and the significance of subtypes of the United Kingdom epidemic strain (PCR ribotype 1). *J Clin Microbiol* **43**, 2685-2696.

21.   McFarland, L. V., Mulligan, M. E., Kwok, R. Y., and Stamm, W. E. (1989) Nosocomial acquisition of Clostridium difficile infection. *N Engl J Med* **320**, 204-210.

22.   Nerandzic, M. M., and Donskey, C. J. Triggering germination represents a novel strategy to enhance killing of Clostridium difficile spores. *PLoS One* **5**, e12285.

23.   Gerding, D. N., Muto, C. A., and Owens, R. C., Jr. (2008) Measures to control and prevent Clostridium difficile infection. *Clin Infect Dis* **46 Suppl 1**, S43-49.

24.   Vonberg, R. P., Kuijper, E. J., Wilcox, M. H., Barbut, F., Tull, P., Gastmeier, P., van den Broek, P. J., Colville, A., Coignard, B., Daha, T., Debast, S., Duerden, B. I., van den Hof, S., van der Kooi, T., Maarleveld, H. J., Nagy, E., Notermans, D. W., O'Driscoll, J., Patel, B., Stone, S., and Wiuff, C. (2008) Infection control measures to limit the spread of Clostridium difficile. *Clin Microbiol Infect* **14 Suppl 5**, 2-20.

25.   Barbut, F., and Petit, J. C. (2001) Epidemiology of Clostridium difficile-associated infections. *Clin Microbiol Infect* **7**, 405-410.

26.   Huang, H., Weintraub, A., Fang, H., and Nord, C. E. (2009) Antimicrobial resistance in Clostridium difficile. *Int J Antimicrob Agents* **34**, 516-522.

27.   Sebaihia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., Stabler, R., Thomson, N. R., Roberts, A. P., Cerdeno-Tarraga, A. M., Wang, H., Holden, M. T., Wright, A., Churcher, C., Quail, M. A., Baker, S., Bason, N., Brooks, K., Chillingworth, T., Cronin, A., Davis, P., Dowd, L., Fraser, A., Feltwell, T., Hance,

Z., Holroyd, S., Jagels, K., Moule, S., Mungall, K., Price, C., Rabbinowitsch, E., Sharp, S., Simmonds, M., Stevens, K., Unwin, L., Whithead, S., Dupuy, B., Dougan, G., Barrell, B., and Parkhill, J. (2006) The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. *Nat Genet* **38**, 779-786.

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

29.   Walters, B. A., Roberts, R., Stafford, R., and Seneviratne, E. (1983) Relapse of antibiotic associated colitis: endogenous persistence of Clostridium difficile during vancomycin therapy. *Gut* **24**, 206-212.

30.   Sohn, S., Climo, M., Diekema, D., Fraser, V., Herwaldt, L., Marino, S., Noskin, G., Perl, T., Song, X., Tokars, J., Warren, D., Wong, E., Yokoe, D. S., Zembower, T., and Sepkowitz, K. A. (2005) Varying rates of Clostridium difficile-associated diarrhea at prevention epicenter hospitals. *Infect Control Hosp Epidemiol* **26**, 676-679.

31.   McDonald, L. C., Coignard, B., Dubberke, E., Song, X., Horan, T., and Kutty, P. K. (2007) Recommendations for surveillance of Clostridium difficile-associated disease. *Infect Control Hosp Epidemiol* **28**, 140-145.

32.   Vonberg, R. P., Reichardt, C., Behnke, M., Schwab, F., Zindler, S., and Gastmeier, P. (2008) Costs of nosocomial Clostridium difficile-associated diarrhoea. *J Hosp Infect* **70**, 15-20.

33.   Dubberke, E. R., Reske, K. A., Olsen, M. A., McDonald, L. C., and Fraser, V. J. (2008) Short- and long-term attributable costs of Clostridium difficile-associated disease in nonsurgical inpatients. *Clin Infect Dis* **46**, 497-504.

34.   Brazier, J. S. (2008) Clostridium difficile: from obscurity to superbug. *Br J Biomed Sci* **65**, 39-44.

35.   Kelly, C. P., and LaMont, J. T. (2008) Clostridium difficile--more difficult than ever. *N Engl J Med* **359**, 1932-1940.

36.   Muto, C. A., Pokrywka, M., Shutt, K., Mendelsohn, A. B., Nouri, K., Posey, K., Roberts, T., Croyle, K., Krystofiak, S., Patel-Brown, S., Pasculle, A. W., Paterson, D. L., Saul, M., and Harrison, L. H. (2005) A large outbreak of Clostridium difficile-associated disease with an unexpected proportion of deaths and colectomies at a teaching hospital following increased fluoroquinolone use. *Infect Control Hosp Epidemiol* **26**, 273-280.

37.   Loo, V. G., Poirier, L., Miller, M. A., Oughton, M., Libman, M. D., Michaud, S., Bourgault, A. M., Nguyen, T., Frenette, C., Kelly, M., Vibien, A., Brassard, P., Fenn, S., Dewar, K., Hudson, T. J., Horn, R., Rene, P., Monczak, Y., and Dascal,

A. (2005) A predominantly clonal multi-institutional outbreak of Clostridium difficile-associated diarrhea with high morbidity and mortality. *N Engl J Med* **353**, 2442-2449.

38.   Warny, M., Pepin, J., Fang, A., Killgore, G., Thompson, A., Brazier, J., Frost, E., and McDonald, L. C. (2005) Toxin production by an emerging strain of Clostridium difficile associated with outbreaks of severe disease in North America and Europe. *Lancet* **366**, 1079-1084.

39.   Stabler, R. A., Dawson, L. F., Phua, L. T., and Wren, B. W. (2008) Comparative analysis of BI/NAP1/027 hypervirulent strains reveals novel toxin B-encoding gene (tcdB) sequences. *J Med Microbiol* **57**, 771-775.

40.   Lanis, J. M., Barua, S., and Ballard, J. D. (2010) Variations in TcdB activity and the hypervirulence of emerging strains of Clostridium difficile. *PLoS Pathog* **6**

41.   McDonald, L. C., Killgore, G. E., Thompson, A., Owens, R. C., Jr., Kazakova, S. V., Sambol, S. P., Johnson, S., and Gerding, D. N. (2005) An epidemic, toxin gene-variant strain of Clostridium difficile. *N Engl J Med* **353**, 2433-2441.

42.   Merrigan, M., Venugopal, A., Mallozzi, M., Roxas, B., Viswanathan, V. K., Johnson, S., Gerding, D. N., and Vedantam, G. (2003) Human hypervirulent Clostridium difficile strains exhibit increased sporulation as well as robust toxin production. *J Bacteriol* **192**, 4904-4911.

43.   Akerlund, T., Persson, I., Unemo, M., Noren, T., Svenungsson, B., Wullt, M., and Burman, L. G. (2008) Increased sporulation rate of epidemic Clostridium difficile Type 027/NAP1. *J Clin Microbiol* **46**, 1530-1533.

44.   Gerding, D. N., Muto, C. A., and Owens, R. C., Jr. (2008) Treatment of Clostridium difficile infection. *Clin Infect Dis* **46 Suppl 1**, S32-42.

45.   Fekety, R. (1997) Guidelines for the diagnosis and management of Clostridium difficile-associated diarrhea and colitis. American College of Gastroenterology, Practice Parameters Committee. *Am J Gastroenterol* **92**, 739-750.

46.   McFarland, L. V. (2005) Alternative treatments for Clostridium difficile disease: what really works? *J Med Microbiol* **54**, 101-111.

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

48.   Nelson, R. (2007) Antibiotic treatment for Clostridium difficile-associated diarrhea in adults. *Cochrane Database Syst Rev*, CD004610.

49.   McFarland, L. V. (2006) Meta-analysis of probiotics for the prevention of antibiotic associated diarrhea and the treatment of Clostridium difficile disease. *Am J Gastroenterol* **101**, 812-822.

50. Segarra-Newnham, M. (2007) Probiotics for Clostridium difficile-associated diarrhea: focus on Lactobacillus rhamnosus GG and Saccharomyces boulardii. *Ann Pharmacother* **41**, 1212-1221.

51. Kaslow, D. C., and Shiver, J. W. (2011) Clostridium difficile and methicillin-resistant Staphylococcus aureus: emerging concepts in vaccine development. *Annu Rev Med* **62**, 201-215.

52. Keessen, E. C., Gaastra, W., and Lipman, L. J. (2011) Clostridium difficile infection in humans and animals, differences and similarities. *Vet Microbiol*

53. Songer, J. G., and Anderson, M. A. (2006) Clostridium difficile: an important pathogen of food animals. *Anaerobe* **12**, 1-4.

54. Waligora, A. J., Hennequin, C., Mullany, P., Bourlioux, P., Collignon, A., and Karjalainen, T. (2001) Characterization of a cell surface protein of Clostridium difficile with adhesive properties. *Infect Immun* **69**, 2144-2153.

55. Savariau-Lacomme, M. P., Lebarbier, C., Karjalainen, T., Collignon, A., and Janoir, C. (2003) Transcription and analysis of polymorphism in a cluster of genes encoding surface-associated proteins of Clostridium difficile. *J Bacteriol* **185**, 4461-4470.

56. Popoff, M. R., Rubin, E. J., Gill, D. M., and Boquet, P. (1988) Actin-specific ADP-ribosyltransferase produced by a Clostridium difficile strain. *Infect Immun* **56**, 2299-2306.

57. Perelle, S., Gibert, M., Bourlioux, P., Corthier, G., and Popoff, M. R. (1997) Production of a complete binary toxin (actin-specific ADP-ribosyltransferase) by Clostridium difficile CD196. *Infect Immun* **65**, 1402-1407.

58. Borriello, S. P., Davies, H. A., Kamiya, S., Reed, P. J., and Seddon, S. (1990) Virulence factors of Clostridium difficile. *Rev Infect Dis* **12 Suppl 2**, S185-191.

59. Delmee, M., Avesani, V., Delferriere, N., and Burtonboy, G. (1990) Characterization of flagella of Clostridium difficile and their role in serogrouping reactions. *J Clin Microbiol* **28**, 2210-2214.

60. Stabler, R. A., Gerding, D. N., Songer, J. G., Drudy, D., Brazier, J. S., Trinh, H. T., Witney, A. A., Hinds, J., and Wren, B. W. (2006) Comparative phylogenomics of Clostridium difficile reveals clade specificity and microevolution of hypervirulent strains. *J Bacteriol* **188**, 7297-7305.

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

62. Kelly, C. P., Pothoulakis, C., Vavva, F., Castagliuolo, I., Bostwick, E. F., O'Keane, J. C., Keates, S., and LaMont, J. T. (1996) Anti-Clostridium difficile bovine immunoglobulin concentrate inhibits cytotoxicity and enterotoxicity of C. difficile toxins. *Antimicrob Agents Chemother* **40**, 373-379.

63. Voth, D. E., and Ballard, J. D. (2005) Clostridium difficile toxins: mechanism of action and role in disease. *Clin Microbiol Rev* **18**, 247-263.

64. Rupnik, M., Grabnar, M., and Geric, B. (2003) Binary toxin producing Clostridium difficile strains. *Anaerobe* **9**, 289-294.

65. Geric, B., Carman, R. J., Rupnik, M., Genheimer, C. W., Sambol, S. P., Lyerly, D. M., Gerding, D. N., and Johnson, S. (2006) Binary toxin-producing, large clostridial toxin-negative Clostridium difficile strains are enterotoxic but do not cause disease in hamsters. *J Infect Dis* **193**, 1143-1150.

66. Barbut, F., Decre, D., Lalande, V., Burghoffer, B., Noussair, L., Gigandon, A., Espinasse, F., Raskine, L., Robert, J., Mangeol, A., Branger, C., and Petit, J. C. (2005) Clinical features of Clostridium difficile-associated diarrhoea due to binary toxin (actin-specific ADP-ribosyltransferase)-producing strains. *J Med Microbiol* **54**, 181-185.

67. Ball, D. W., Van Tassell, R. L., Roberts, M. D., Hahn, P. E., Lyerly, D. M., and Wilkins, T. D. (1993) Purification and characterization of alpha-toxin produced by Clostridium novyi type A. *Infect Immun* **61**, 2912-2918.

68. Amimoto, K., Noro, T., Oishi, E., and Shimizu, M. (2007) A novel toxin homologous to large clostridial cytotoxins found in culture supernatant of Clostridium perfringens type C. *Microbiology* **153**, 1198-1206.

69. Chaves-Olarte, E., Low, P., Freer, E., Norlin, T., Weidmann, M., von Eichel-Streiber, C., and Thelestam, M. (1999) A novel cytotoxin from Clostridium difficile serogroup F is a functional hybrid between two other large clostridial cytotoxins. *J Biol Chem* **274**, 11046-11052.

70. Martinez, R. D., and Wilkins, T. D. (1988) Purification and characterization of Clostridium sordellii hemorrhagic toxin and cross-reactivity with Clostridium difficile toxin A (enterotoxin). *Infect Immun* **56**, 1215-1221.

71. Dupuy, B., Govind, R., Antunes, A., and Matamouros, S. (2008) Clostridium difficile toxin synthesis is negatively regulated by TcdC. *J Med Microbiol* **57**, 685-689.

72. Tan, K. S., Wee, B. Y., and Song, K. P. (2001) Evidence for holin function of tcdE gene in the pathogenicity of Clostridium difficile. *J Med Microbiol* **50**, 613-619.

73. Hundsberger, T., Braun, V., Weidmann, M., Leukel, P., Sauerborn, M., and von Eichel-Streiber, C. (1997) Transcription analysis of the genes tcdA-E of the pathogenicity locus of Clostridium difficile. *Eur J Biochem* **244**, 735-742.

74. Jank, T., Giesemann, T., and Aktories, K. (2007) Rho-glucosylating Clostridium difficile toxins A and B: new insights into structure and function. *Glycobiology* **17**, 15R-22R.

75. Bishop, A. L., and Hall, A. (2000) Rho GTPases and their effector proteins. *Biochem J* **348 Pt 2**, 241-255.

76. Aktories, K., and Barbieri, J. T. (2005) Bacterial cytotoxins: targeting eukaryotic switches. *Nat Rev Microbiol* **3**, 397-410.

77. Just, I., Selzer, J., Wilm, M., von Eichel-Streiber, C., Mann, M., and Aktories, K. (1995) Glucosylation of Rho proteins by Clostridium difficile toxin B. *Nature* **375**, 500-503.

78. Genth, H., Aktories, K., and Just, I. (1999) Monoglucosylation of RhoA at threonine 37 blocks cytosol-membrane cycling. *J Biol Chem* **274**, 29050-29056.

79. Aktories, K., and Just, I. (1995) Monoglucosylation of low-molecular-mass GTP-binding Rho proteins by clostridial cytotoxins. *Trends Cell Biol* **5**, 441-443.

80. Ihara, K., Muraguchi, S., Kato, M., Shimizu, T., Shirakawa, M., Kuroda, S., Kaibuchi, K., and Hakoshima, T. (1998) Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue. *J Biol Chem* **273**, 9656-9666.

81. Sehr, P., Joseph, G., Genth, H., Just, I., Pick, E., and Aktories, K. (1998) Glucosylation and ADP ribosylation of rho proteins: effects on nucleotide binding, GTPase activity, and effector coupling. *Biochemistry* **37**, 5296-5304.

82. Herrmann, C., Ahmadian, M. R., Hofmann, F., and Just, I. (1998) Functional consequences of monoglucosylation of Ha-Ras at effector domain amino acid threonine 35. *J Biol Chem* **273**, 16134-16139.

83. Donta, S. T., Sullivan, N., and Wilkins, T. D. (1982) Differential effects of Clostridium difficile toxins on tissue-cultured cells. *J Clin Microbiol* **15**, 1157-1158.

84. Chaves-Olarte, E., Weidmann, M., Eichel-Streiber, C., and Thelestam, M. (1997) Toxins A and B from Clostridium difficile differ with respect to enzymatic potencies, cellular substrate specificities, and surface binding to cultured cells. *J Clin Invest* **100**, 1734-1741.

85. Aktories, K. (1997) Bacterial toxins that target Rho proteins. *J Clin Invest* **99**, 827-829.

86. Tucker, K. D., Carrig, P. E., and Wilkins, T. D. (1990) Toxin A of Clostridium difficile is a potent cytotoxin. *J Clin Microbiol* **28**, 869-871.

87. Qa'Dan, M., Ramsey, M., Daniel, J., Spyres, L. M., Safiejko-Mroczka, B., Ortiz-Leduc, W., and Ballard, J. D. (2002) Clostridium difficile toxin B activates dual caspase-dependent and caspase-independent apoptosis in intoxicated cells. *Cell Microbiol* **4**, 425-434.

88. Matarrese, P., Falzano, L., Fabbri, A., Gambardella, L., Frank, C., Geny, B., Popoff, M. R., Malorni, W., and Fiorentini, C. (2007) Clostridium difficile toxin B

causes apoptosis in epithelial cells by thrilling mitochondria. Involvement of ATP-sensitive mitochondrial potassium channels. *J Biol Chem* **282**, 9029-9041.

89.   Warny, M., and Kelly, C. P. (1999) Monocytic cell necrosis is mediated by potassium depletion and caspase-like proteases. *Am J Physiol* **276**, C717-724.

90.   Hippenstiel, S., Schmeck, B., N'Guessan, P. D., Seybold, J., Krull, M., Preissner, K., Eichel-Streiber, C. V., and Suttorp, N. (2002) Rho protein inactivation induced apoptosis of cultured human endothelial cells. *Am J Physiol Lung Cell Mol Physiol* **283**, L830-838.

91.   Nottrott, S., Schoentaube, J., Genth, H., Just, I., and Gerhard, R. (2007) Clostridium difficile toxin A-induced apoptosis is p53-independent but depends on glucosylation of Rho GTPases. *Apoptosis* **12**, 1443-1453.

92.   Kim, H., Kokkotou, E., Na, X., Rhee, S. H., Moyer, M. P., Pothoulakis, C., and Lamont, J. T. (2005) Clostridium difficile toxin A-induced colonocyte apoptosis involves p53-dependent p21(WAF1/CIP1) induction via p38 mitogen-activated protein kinase. *Gastroenterology* **129**, 1875-1888.

93.   Nusrat, A., Giry, M., Turner, J. R., Colgan, S. P., Parkos, C. A., Carnes, D., Lemichez, E., Boquet, P., and Madara, J. L. (1995) Rho protein regulates tight junctions and perijunctional actin organization in polarized epithelia. *Proc Natl Acad Sci U S A* **92**, 10629-10633.

94.   Feltis, B. A., Wiesner, S. M., Kim, A. S., Erlandsen, S. L., Lyerly, D. L., Wilkins, T. D., and Wells, C. L. (2000) Clostridium difficile toxins A and B can alter epithelial permeability and promote bacterial paracellular migration through HT-29 enterocytes. *Shock* **14**, 629-634.

95.   Johal, S. S., Solomon, K., Dodson, S., Borriello, S. P., and Mahida, Y. R. (2004) Differential effects of varying concentrations of clostridium difficile toxin A on epithelial barrier function and expression of cytokines. *J Infect Dis* **189**, 2110-2119.

96.   Castagliuolo, I., Keates, A. C., Wang, C. C., Pasha, A., Valenick, L., Kelly, C. P., Nikulasson, S. T., LaMont, J. T., and Pothoulakis, C. (1998) Clostridium difficile toxin A stimulates macrophage-inflammatory protein-2 production in rat intestinal epithelial cells. *J Immunol* **160**, 6039-6045.

97.   Castagliuolo, I., Riegler, M., Pasha, A., Nikulasson, S., Lu, B., Gerard, C., Gerard, N. P., and Pothoulakis, C. (1998) Neurokinin-1 (NK-1) receptor is required in Clostridium difficile- induced enteritis. *J Clin Invest* **101**, 1547-1550.

98.   Warny, M., Keates, A. C., Keates, S., Castagliuolo, I., Zacks, J. K., Aboudola, S., Qamar, A., Pothoulakis, C., LaMont, J. T., and Kelly, C. P. (2000) p38 MAP kinase activation by Clostridium difficile toxin A mediates monocyte necrosis, IL-8 production, and enteritis. *J Clin Invest* **105**, 1147-1156.

99. Voth, D. E., Qa'Dan, M., Hamm, E. E., Pelfrey, J. M., and Ballard, J. D. (2004) Clostridium sordellii lethal toxin is maintained in a multimeric protein complex. *Infect Immun* **72**, 3366-3372.

100. Flegel, W. A., Muller, F., Daubener, W., Fischer, H. G., Hadding, U., and Northoff, H. (1991) Cytokine response by human monocytes to Clostridium difficile toxin A and toxin B. *Infect Immun* **59**, 3659-3666.

101. Savidge, T. C., Pan, W. H., Newman, P., O'Brien, M., Anton, P. M., and Pothoulakis, C. (2003) Clostridium difficile toxin B is an inflammatory enterotoxin in human intestine. *Gastroenterology* **125**, 413-420.

102. Jiang, Z. D., DuPont, H. L., Garey, K., Price, M., Graham, G., Okhuysen, P., Dao-Tran, T., and LaRocco, M. (2006) A common polymorphism in the interleukin 8 gene promoter is associated with Clostridium difficile diarrhea. *Am J Gastroenterol* **101**, 1112-1116.

103. Mitchell, T. J., Ketley, J. M., Haslam, S. C., Stephen, J., Burdon, D. W., Candy, D. C., and Daniel, R. (1986) Effect of toxin A and B of Clostridium difficile on rabbit ileum and colon. *Gut* **27**, 78-85.

104. Lyerly, D. M., Saum, K. E., MacDonald, D. K., and Wilkins, T. D. (1985) Effects of Clostridium difficile toxins given intragastrically to animals. *Infect Immun* **47**, 349-352.

105. Lyerly, D. M., Lockwood, D. E., Richardson, S. H., and Wilkins, T. D. (1982) Biological activities of toxins A and B of Clostridium difficile. *Infect Immun* **35**, 1147-1150.

106. Kyne, L., Warny, M., Qamar, A., and Kelly, C. P. (2001) Association between antibody response to toxin A and protection against recurrent Clostridium difficile diarrhoea. *Lancet* **357**, 189-193.

107. Giannasca, P. J., and Warny, M. (2004) Active and passive immunization against Clostridium difficile diarrhea and colitis. *Vaccine* **22**, 848-856.

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

109. Warny, M., Vaerman, J. P., Avesani, V., and Delmee, M. (1994) Human antibody response to Clostridium difficile toxin A in relation to clinical course of infection. *Infect Immun* **62**, 384-389.

110. Sambol, S. P., Merrigan, M. M., Lyerly, D., Gerding, D. N., and Johnson, S. (2000) Toxin gene analysis of a variant strain of Clostridium difficile that causes human clinical disease. *Infect Immun* **68**, 5480-5487.

111. Riegler, M., Sedivy, R., Pothoulakis, C., Hamilton, G., Zacherl, J., Bischof, G., Cosentini, E., Feil, W., Schiessel, R., LaMont, J. T., and et al. (1995) Clostridium difficile toxin B is more potent than toxin A in damaging human colonic epithelium in vitro. *J Clin Invest* **95**, 2004-2011.

112. O'Connor, J. R., Lyras, D., Farrow, K. A., Adams, V., Powell, D. R., Hinds, J., Cheung, J. K., and Rood, J. I. (2006) Construction and analysis of chromosomal Clostridium difficile mutants. *Mol Microbiol* **61**, 1335-1351.

113. Lyras, D., O'Connor, J. R., Howarth, P. M., Sambol, S. P., Carter, G. P., Phumoonna, T., Poon, R., Adams, V., Vedantam, G., Johnson, S., Gerding, D. N., and Rood, J. I. (2009) Toxin B is essential for virulence of Clostridium difficile. *Nature* **458**, 1176-1179.

114. Kuehne, S. A., Cartman, S. T., Heap, J. T., Kelly, M. L., Cockayne, A., and Minton, N. P. (2010) The role of toxin A and toxin B in Clostridium difficile infection. *Nature* **467**, 711-713.

115. Ho, J. G., Greco, A., Rupnik, M., and Ng, K. K. (2005) Crystal structure of receptor-binding C-terminal repeats from Clostridium difficile toxin A. *Proc Natl Acad Sci U S A* **102**, 18373-18378.

116. von Eichel-Streiber, C., and Sauerborn, M. (1990) Clostridium difficile toxin A carries a C-terminal repetitive structure homologous to the carbohydrate binding region of streptococcal glycosyltransferases. *Gene* **96**, 107-113.

117. von Eichel-Streiber, C., Sauerborn, M., and Kuramitsu, H. K. (1992) Evidence for a modular structure of the homologous repetitive C-terminal carbohydrate-binding sites of Clostridium difficile toxins and Streptococcus mutans glucosyltransferases. *J Bacteriol* **174**, 6707-6710.

118. von Eichel-Streiber, C., Laufenberg-Feldmann, R., Sartingen, S., Schulze, J., and Sauerborn, M. (1992) Comparative sequence analysis of the Clostridium difficile toxins A and B. *Mol Gen Genet* **233**, 260-268.

119. Krivan, H. C., Clark, G. F., Smith, D. F., and Wilkins, T. D. (1986) Cell surface binding site for Clostridium difficile enterotoxin: evidence for a glycoconjugate containing the sequence Gal alpha 1-3Gal beta 1-4GlcNAc. *Infect Immun* **53**, 573-581.

120. Tucker, K. D., and Wilkins, T. D. (1991) Toxin A of Clostridium difficile binds to the human carbohydrate antigens I, X, and Y. *Infect Immun* **59**, 73-78.

121. Teneberg, S., Lonnroth, I., Torres Lopez, J. F., Galili, U., Halvarsson, M. O., Angstrom, J., and Karlsson, K. A. (1996) Molecular mimicry in the recognition of glycosphingolipids by Gal alpha 3 Gal beta 4 GlcNAc beta-binding Clostridium difficile toxin A, human natural anti alpha-galactosyl IgG and the monoclonal antibody Gal-13: characterization of a binding-active human glycosphingolipid, non-identical with the animal receptor. *Glycobiology* **6**, 599-609.

122. Pothoulakis, C., Gilbert, R. J., Cladaras, C., Castagliuolo, I., Semenza, G., Hitti, Y., Montcrief, J. S., Linevsky, J., Kelly, C. P., Nikulasson, S., Desai, H. P., Wilkins, T. D., and LaMont, J. T. (1996) Rabbit sucrase-isomaltase contains a functional intestinal receptor for Clostridium difficile toxin A. *J Clin Invest* **98**, 641-649.

123. Na, X., Kim, H., Moyer, M. P., Pothoulakis, C., and LaMont, J. T. (2008) gp96 is a human colonocyte plasma membrane binding protein for Clostridium difficile toxin A. *Infect Immun* **76**, 2862-2871.

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

125. Barroso, L. A., Moncrief, J. S., Lyerly, D. M., and Wilkins, T. D. (1994) Mutagenesis of the Clostridium difficile toxin B gene and effect on cytotoxic activity. *Microb Pathog* **16**, 297-303.

126. Olling, A., Goy, S., Hoffmann, F., Tatge, H., Just, I., and Gerhard, R. (2011) The Repetitive Oligopeptide Sequences Modulate Cytopathic Potency but Are Not Crucial for Cellular Uptake of Clostridium difficile Toxin A. *PLoS One* **6**, e17623.

127. Papatheodorou, P., Zamboglou, C., Genisyuerek, S., Guttenberg, G., and Aktories, K. (2010) Clostridial glucosylating toxins enter cells via clathrin-mediated endocytosis. *PLoS One* **5**, e10673.

128. Qa'Dan, M., Spyres, L. M., and Ballard, J. D. (2000) pH-induced conformational changes in Clostridium difficile toxin B. *Infect Immun* **68**, 2470-2474.

129. Barth, H., Pfeifer, G., Hofmann, F., Maier, E., Benz, R., and Aktories, K. (2001) Low pH-induced formation of ion channels by clostridium difficile toxin B in target cells. *J Biol Chem* **276**, 10670-10676.

130. Giesemann, T., Jank, T., Gerhard, R., Maier, E., Just, I., Benz, R., and Aktories, K. (2006) Cholesterol-dependent pore formation of Clostridium difficile toxin A. *J Biol Chem* **281**, 10808-10815.

131. Genisyuerek, S., Papatheodorou, P., Guttenberg, G., Schubert, R., Benz, R., and Aktories, K. (2011) Structural determinants for membrane insertion, pore formation and translocation of Clostridium difficile toxin B. *Mol Microbiol* **79**, 1643-1654.

132. Pfeifer, G., Schirmer, J., Leemhuis, J., Busch, C., Meyer, D. K., Aktories, K., and Barth, H. (2003) Cellular uptake of Clostridium difficile toxin B. Translocation of the N-terminal catalytic domain into the cytosol of eukaryotic cells. *J Biol Chem* **278**, 44535-44541.

133. Rupnik, M., Pabst, S., Rupnik, M., von Eichel-Streiber, C., Urlaub, H., and Soling, H. D. (2005) Characterization of the cleavage site and function of resulting cleavage fragments after limited proteolysis of Clostridium difficile toxin B (TcdB) by host cells. *Microbiology* **151**, 199-208.

134. Reineke, J., Tenzer, S., Rupnik, M., Koschinski, A., Hasselmayer, O., Schrattenholz, A., Schild, H., and von Eichel-Streiber, C. (2007) Autocatalytic cleavage of Clostridium difficile toxin B. *Nature* **446**, 415-419.

135. Egerer, M., Giesemann, T., Jank, T., Satchell, K. J., and Aktories, K. (2007) Auto-catalytic cleavage of Clostridium difficile toxins A and B depends on cysteine protease activity. *J Biol Chem* **282**, 25314-25321.

136. Sheahan, K. L., Cordero, C. L., and Satchell, K. J. (2007) Autoprocessing of the Vibrio cholerae RTX toxin by the cysteine protease domain. *Embo J* **26**, 2552-2561.

137. Prochazkova, K., and Satchell, K. J. (2008) Structure-function analysis of inositol hexakisphosphate-induced autoprocessing of the Vibrio cholerae multifunctional autoprocessing RTX toxin. *J Biol Chem* **283**, 23656-23664.

138. Lupardus, P. J., Shen, A., Bogyo, M., and Garcia, K. C. (2008) Small molecule-induced allosteric activation of the Vibrio cholerae RTX cysteine protease domain. *Science* **322**, 265-268.

139. Egerer, M., Giesemann, T., Herrmann, C., and Aktories, K. (2009) Autocatalytic processing of Clostridium difficile toxin B. Binding of inositol hexakisphosphate. *J Biol Chem* **284**, 3389-3395.

140. Hofmann, F., Busch, C., Prepens, U., Just, I., and Aktories, K. (1997) Localization of the glucosyltransferase activity of Clostridium difficile toxin B to the N-terminal part of the holotoxin. *J Biol Chem* **272**, 11074-11078.

141. Just, I., Wilm, M., Selzer, J., Rex, G., von Eichel-Streiber, C., Mann, M., and Aktories, K. (1995) The enterotoxin from Clostridium difficile (ToxA) monoglucosylates the Rho proteins. *J Biol Chem* **270**, 13932-13936.

142. Selzer, J., Hofmann, F., Rex, G., Wilm, M., Mann, M., Just, I., and Aktories, K. (1996) Clostridium novyi alpha-toxin-catalyzed incorporation of GlcNAc into Rho subfamily proteins. *J Biol Chem* **271**, 25173-25177.

143. Nagahama, M., Ohkubo, A., Oda, M., Kobayashi, K., Amimoto, K., Miyamoto, K., and Sakurai, J. (2011) Clostridium perfringens TpeL glycosylates the Rac and Ras subfamily proteins. *Infect Immun* **79**, 905-910.

144. Genth, H., Dreger, S. C., Huelsenbeck, J., and Just, I. (2008) Clostridium difficile toxins: more than mere inhibitors of Rho proteins. *Int J Biochem Cell Biol* **40**, 592-597.

145. Genth, H., Hofmann, F., Selzer, J., Rex, G., Aktories, K., and Just, I. (1996) Difference in protein substrate specificity between hemorrhagic toxin and lethal toxin from Clostridium sordellii. *Biochem Biophys Res Commun* **229**, 370-374.

146. Raaijmakers, J. H., and Bos, J. L. (2009) Specificity in Ras and Rap signaling. *J Biol Chem* **284**, 10995-10999.

147. Huelsenbeck, J., Dreger, S., Gerhard, R., Barth, H., Just, I., and Genth, H. (2007) Difference in the cytotoxic effects of toxin B from Clostridium difficile strain VPI 10463 and toxin B from variant Clostridium difficile strain 1470. *Infect Immun* **75**, 801-809.

148. Mehlig, M., Moos, M., Braun, V., Kalt, B., Mahony, D. E., and von Eichel-Streiber, C. (2001) Variant toxin B and a functional toxin A produced by Clostridium difficile C34. *FEMS Microbiol Lett* **198**, 171-176.

149. Giry, M., Popoff, M. R., von Eichel-Streiber, C., and Boquet, P. (1995) Transient expression of RhoA, -B, and -C GTPases in HeLa cells potentiates resistance to Clostridium difficile toxins A and B but not to Clostridium sordellii lethal toxin. *Infect Immun* **63**, 4063-4071.

150. Reinert, D. J., Jank, T., Aktories, K., and Schulz, G. E. (2005) Structural basis for the function of Clostridium difficile toxin B. *J Mol Biol* **351**, 973-981.

151. Ziegler, M. O., Jank, T., Aktories, K., and Schulz, G. E. (2008) Conformational changes and reaction of clostridial glycosylating toxins. *J Mol Biol* **377**, 1346-1356.

152. Geissler, B., Tungekar, R., and Satchell, K. J. (2010) Identification of a conserved membrane localization domain within numerous large bacterial protein toxins. *Proc Natl Acad Sci U S A* **107**, 5581-5586.

153. Mesmin, B., Robbe, K., Geny, B., Luton, F., Brandolin, G., Popoff, M. R., and Antonny, B. (2004) A phosphatidylserine-binding site in the cytosolic fragment of Clostridium sordellii lethal toxin facilitates glucosylation of membrane-bound Rac and is required for cytotoxicity. *J Biol Chem* **279**, 49876-49882.

154. Jank, T., Giesemann, T., and Aktories, K. (2007) Clostridium difficile glucosyltransferase toxin B-essential amino acids for substrate binding. *J Biol Chem* **282**, 35222-35231.

155. Minor, Z. O. a. W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. in *Methods in Enzymology* (C.W. Carter, J. a. R. M. S. ed., Academic Press, New York

156. Vonrhein, C., Blanc, E., Roversi, P., and Bricogne, G. (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* **364**, 215-230.

157. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.

158. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**, 1948-1954.

159. Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345-364.

160. Hardy, J. A., Lam, J., Nguyen, J. T., O'Brien, T., and Wells, J. A. (2004) Discovery of an allosteric site in the caspases. *Proc Natl Acad Sci U S A* **101**, 12461-12466.

161. Holm, L., and Park, J. (2000) DaliLite workbench for protein structure comparison. *Bioinformatics* **16**, 566-567.

162. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-308.

163. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680.

164. Macbeth, M. R., Schubert, H. L., Vandemark, A. P., Lingam, A. T., Hill, C. P., and Bass, B. L. (2005) Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science* **309**, 1534-1539.

165. Pei, J., Lupardus, P. J., Garcia, K. C., and Grishin, N. V. (2009) CPDadh: a new peptidase family homologous to the cysteine protease domain in bacterial MARTX toxins. *Protein Sci* **18**, 856-862.

166. Volkman, B. F., Lipson, D., Wemmer, D. E., and Kern, D. (2001) Two-state allosteric behavior in a single-domain signaling protein. *Science* **291**, 2429-2433.

167. Shen, A., Lupardus, P. J., Gersch, M. M., Puri, A. W., Albrow, V. E., Garcia, K. C., and Bogyo, M. (2011) Defining an allosteric circuit in the cysteine protease domain of Clostridium difficile toxins. *Nat Struct Mol Biol* **18**, 364-371.

168. Prochazkova, K., Shuvalova, L. A., Minasov, G., Voburka, Z., Anderson, W. F., and Satchell, K. J. (2009) Structural and molecular mechanism for autoprocessing of MARTX toxin of Vibrio cholerae at multiple sites. *J Biol Chem* **284**, 26557-26568.

169. Puri, A. W., Lupardus, P. J., Deu, E., Albrow, V. E., Garcia, K. C., Bogyo, M., and Shen, A. (2010) Rational design of inhibitors and activity-based probes targeting Clostridium difficile virulence factor TcdB. *Chem Biol* **17**, 1201-1211.

170. Carter, P., and Wells, J. A. (1988) Dissecting the catalytic triad of a serine protease. *Nature* **332**, 564-568.

171. Hilvert, D. (2000) Critical analysis of antibody catalysis. *Annu Rev Biochem* **69**, 751-793.

172. Albesa-Jove, D., Bertrand, T., Carpenter, E. P., Swain, G. V., Lim, J., Zhang, J., Haire, L. F., Vasisht, N., Braun, V., Lange, A., von Eichel-Streiber, C., Svergun, D. I., Fairweather, N. F., and Brown, K. A. (2010) Four Distinct Structural Domains in Clostridium difficile Toxin B Visualized Using SAXS. *J Mol Biol* **396**, 1260-1270.

173. Pruitt, R. N., Chagot, B., Cover, M., Chazin, W. J., Spiller, B., and Lacy, D. B. (2009) Structure-function analysis of inositol hexakisphosphate-induced autoprocessing in Clostridium difficile toxin A. *J Biol Chem* **284**, 21934-21940.

174. Yang, G., Zhou, B., Wang, J., He, X., Sun, X., Nie, W., Tzipori, S., and Feng, H. (2008) Expression of recombinant Clostridium difficile toxin A and B in Bacillus megaterium. *BMC Microbiol* **8**, 192.

175. Ohi, M., Li, Y., Cheng, Y., and Walz, T. (2004) Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. *Biol Proced Online* **6**, 23-34.

176. Ludtke, S. J., Baldwin, P. R., and Chiu, W. (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128**, 82-97.

177. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**, 190-199.

178. van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R., and Schatz, M. (1996) A new generation of the IMAGIC image processing system. *J Struct Biol* **116**, 17-24.

179. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612.

180. Radermacher, M., Wagenknecht, T., Verschoor, A., and Frank, J. (1987) Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of Escherichia coli. *J Microsc* **146**, 113-136.

181. Bottcher, B., Wynne, S. A., and Crowther, R. A. (1997) Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature* **386**, 88-91.

182. Frisch, C., Gerhard, R., Aktories, K., Hofmann, F., and Just, I. (2003) The complete receptor-binding domain of Clostridium difficile toxin A is required for endocytosis. *Biochem Biophys Res Commun* **300**, 706-711.

183. Dingle, T., Wee, S., Mulvey, G. L., Greco, A., Kitova, E. N., Sun, J., Lin, S., Klassen, J. S., Palcic, M. M., Ng, K. K., and Armstrong, G. D. (2008) Functional properties of the carboxy-terminal host cell-binding domains of the two toxins, TcdA and TcdB, expressed by Clostridium difficile. *Glycobiology* **18**, 698-706.

184. Putnam, C. D., Hammel, M., Hura, G. L., and Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* **40**, 191-285.

185. Pretto, D. I., Tsutakawa, S., Brosey, C. A., Castillo, A., Chagot, M. E., Smith, J. A., Tainer, J. A., and Chazin, W. J. (2010) Structural dynamics and single-stranded DNA binding activity of the three N-terminal domains of the large subunit of replication protein A from small angle X-ray scattering. *Biochemistry* **49**, 2880-2889.

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

187. Zhao, J. M., and London, E. (1988) Conformation and model membrane interactions of diphtheria toxin fragment A. *J Biol Chem* **263**, 15369-15377.

188. Koriazova, L. K., and Montal, M. (2003) Translocation of botulinum neurotoxin light chain protease through the heavy chain channel. *Nat Struct Biol* **10**, 13-18.

189. Wesche, J., Elliott, J. L., Falnes, P. O., Olsnes, S., and Collier, R. J. (1998) Characterization of membrane translocation by anthrax protective antigen. *Biochemistry* **37**, 15737-15746.

190. Lacy, D. B., and Stevens, R. C. (1998) Unraveling the structures and modes of action of bacterial toxins. *Curr Opin Struct Biol* **8**, 778-784.

191. Lyerly, D. M., Barroso, L. A., Wilkins, T. D., Depitre, C., and Corthier, G. (1992) Characterization of a toxin A-negative, toxin B-positive strain of Clostridium difficile. *Infect Immun* **60**, 4633-4639.

192. Sullivan, N. M., Pellett, S., and Wilkins, T. D. (1982) Purification and characterization of toxins A and B of Clostridium difficile. *Infect Immun* **35**, 1032-1040.

193. Kreimeyer, I., Euler, F., Marckscheffel, A., Tatge, H., Pich, A., Olling, A., Schwarz, J., Just, I., and Gerhard, R. (2011) Autoproteolytic cleavage mediates cytotoxicity of Clostridium difficile toxin A. *Naunyn Schmiedebergs Arch Pharmacol* **383**, 253-262.

194. Pruitt, R. N., Chambers, M. G., Ng, K. K., Ohi, M. D., and Lacy, D. B. (2010) Structural organization of the functional domains of Clostridium difficile toxins A and B. *Proc Natl Acad Sci U S A* **107**, 13467-13472.

195. Worby, C. A., Mattoo, S., Kruger, R. P., Corbeil, L. B., Koller, A., Mendez, J. C., Zekarias, B., Lazar, C., and Dixon, J. E. (2009) The fic domain: regulation of cell signaling by adenylylation. *Mol Cell* **34**, 93-103.

196. Sancak, Y., Peterson, T. R., Shaul, Y. D., Lindquist, R. A., Thoreen, C. C., Bar-Peled, L., and Sabatini, D. M. (2008) The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* **320**, 1496-1501.

197. Burns, J. C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J. K. (1993) Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors:

concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc Natl Acad Sci U S A* **90**, 8033-8037.

198. Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L., and Trono, D. (1997) Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat Biotechnol* **15**, 871-875.

199. Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem Sci* **30**, 53-62.

200. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-134.

201. Ciesla, W. P., Jr., and Bobak, D. A. (1998) Clostridium difficile toxins A and B are cation-dependent UDP-glucose hydrolases with differing catalytic activities. *J Biol Chem* **273**, 16021-16026.

202. Spyres, L. M., Daniel, J., Hensley, A., Qa'Dan, M., Ortiz-Leduc, W., and Ballard, J. D. (2003) Mutational analysis of the enzymatic domain of Clostridium difficile toxin B reveals novel inhibitors of the wild-type toxin. *Infect Immun* **71**, 3294-3301.

203. Pruitt, R. N., Chambers, M. G., Ng, K. K., Ohi, M. D., and Lacy, D. B. (2010) Structural organization of the functional domains of Clostridium difficile toxins A and B. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 13467-13472.

204. Dove, C. H., Wang, S. Z., Price, S. B., Phelps, C. J., Lyerly, D. M., Wilkins, T. D., and Johnson, J. L. (1990) Molecular characterization of the Clostridium difficile toxin A gene. *Infect Immun* **58**, 480-488.

205. Mushrush, D. J., Koteiche, H. A., Sammons, M. A., Link, A. J., McHaourab, H. S., and Lacy, D. B. (2011) Studies of the Mechanistic Details of the pH-dependent Association of Botulinum Neurotoxin with Membranes. *J Biol Chem* **286**, 27011-27018.