Chemistry & Biology
Article

Rational Design of Inhibitors and Activity-Based Probes Targeting *Clostridium difficile* Virulence Factor TcdB

Aaron W. Puri,[1] Patrick J. Lupardus,[2,3,6] Edgar Deu,[4] Victoria E. Albrow,[4,7] K. Christopher Garcia,[2,3] Matthew Bogyo,[1,4,5,*] and Aimee Shen[4,*]
[1]Department of Chemical and Systems Biology
[2]Department of Molecular and Cellular Physiology
[3]Howard Hughes Medical Institute
[4]Department of Pathology
[5]Department of Microbiology and Immunology
Stanford University School of Medicine, 300 Pasteur Drive, Stanford, California 94305, USA
[6]Present address: Genentech, Inc., South San Francisco, CA 94080, USA
[7]Present address: Pfizer, Inc., Sandwich, Kent 01304, UK
*Correspondence: mbogyo@stanford.edu (M.B.), ashen2@stanford.edu (A.S.)
DOI 10.1016/j.chembiol.2010.09.011

SUMMARY

Clostridium difficile **is a leading cause of nosocomial infections. The major virulence factors of this pathogen are the multi-domain toxins TcdA and TcdB. These toxins contain a cysteine protease domain (CPD) that autoproteolytically releases a cytotoxic effector domain upon binding intracellular inositol hexakisphosphate. Currently, there are no known inhibitors of this protease. Here, we describe the rational design of covalent small molecule inhibitors of TcdB CPD. We identified compounds that inactivate TcdB holotoxin function in cells and solved the structure of inhibitor-bound protease to 2.0 Å. This structure reveals the molecular basis of CPD substrate recognition and informed the synthesis of activity-based probes for this enzyme. The inhibitors presented will guide the development of therapeutics targeting *C. difficile*, and the probes will serve as tools for studying the unique activation mechanism of bacterial toxin CPDs.**

INTRODUCTION

The Gram-positive anaerobic bacterium *Clostridium difficile* is a major cause of hospital-acquired diarrhea and the severe gastrointestinal illness pseudomembranous colitis (Kelly and LaMont, 2008; Rupnik et al., 2009). Although infection rates have risen dramatically in the last decade, there is currently a lack of therapeutics to treat *C. difficile* infection (Halsey, 2008; Kelly and LaMont, 2008). This is in large part due to the organism's resistance to most classes of antibiotics. A viable strategy for combating *C. difficile* and other prominent bacterial pathogens is to target virulence factors instead of essential enzymes (Clatworthy et al., 2007; Puri and Bogyo, 2009). This method limits the selective pressure on the organism to develop resistance to treatment, extending the effective lifespan of the drug. The large glucosylating toxins TcdA and TcdB are ideal targets for this approach because they are the primary virulence factors of *C. difficile* (Genth et al., 2008; Jank and Aktories, 2008). TcdB in particular has been shown to be critical for virulence and is found in all clinical isolates (Lyras et al., 2009; Rupnik et al., 2009).

Both TcdA and TcdB cause cell death through an orchestrated sequence of events (Jank and Aktories, 2008). These multi-domain toxin proteins first enter cells by triggering receptor-mediated endocytosis (Frisch et al., 2003; Rolfe and Song, 1993); acidification of toxin-containing endosomal compartments subsequently initiates translocation of the N-terminal cytotoxic glucosyltransferase domain and presumably the cysteine protease domain (CPD) into the cytosol (Just et al., 1995; Pfeifer et al., 2003; Qa'Dan et al., 2000). The CPD is activated by the eukaryotic-specific small molecule inositol hexakisphosphate ($InsP_6$) (Egerer et al., 2007; Reineke et al., 2007). This activation catalyzes the autoproteolytic release of the toxin's cytotoxic glucosyltransferase domain from the endosomal membrane (Egerer et al., 2007; Pfeifer et al., 2003). The liberated effector domain then monoglucosylates small Rho family GTPases (Just et al., 1995), resulting in loss of cell-cell junctions and ultimately cell death (Genth et al., 2008; Gerhard et al., 2008; Qa'Dan et al., 2002).

CPD-mediated autoprocessing of TcdB is a critical step during target cell intoxication. Genetic inactivation of the CPD has been shown to reduce the overall function of TcdB in target cells (Egerer et al., 2007). A homologous CPD also autoproteolytically regulates the Multifunctional Autoprocessing RTX (MARTX) toxins (Prochazkova et al., 2009; Sheahan et al., 2007; Shen et al., 2009), an otherwise unrelated family of toxins produced by Gram-negative bacteria (Satchell, 2007). Structural analyses of the CPD of both families of toxins have demonstrated that the protease is allosterically regulated by the small molecule $InsP_6$ (Lupardus et al., 2008; Prochazkova et al., 2009; Pruitt et al., 2009). These analyses have also revealed that the CPD is a clan CD protease whose closest known structural homolog is human caspase-7 (Lupardus et al., 2008). Despite their disparate mechanism of activation, *V. cholerae* MARTX CPD exhibits

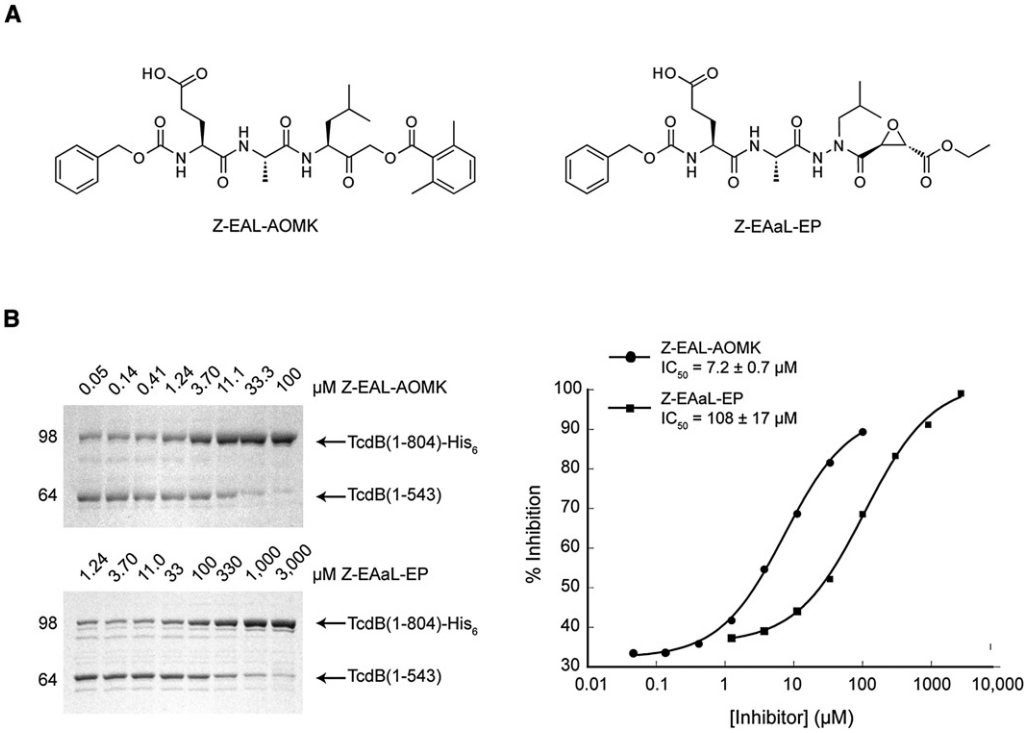


Figure 1. Chemical Inhibition of TcdB CPD

(A) $MARTX_{Vc}$ CPD inhibitors Cbz-EAL-AOMK and Cbz-EAaL-EP, which contain the AOMK and aza-Leu epoxide electrophilic warheads, respectively.

(B) Gel-based TcdB(1-804) autocleavage assay. Inhibitor concentrations were titrated, and the resulting blockade of recombinant toxin autocleavage was assessed by SDS-PAGE (left). The relative cleavage amounts were then quantified and globally fit to determine observed IC_{50} values for each compound (right). Data represent the mean of three experiments ± standard deviation. See also Figure S1.

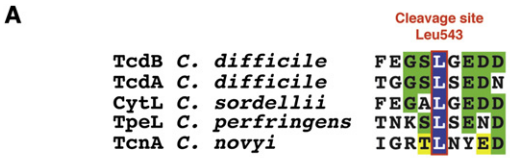
similarities in substrate recognition to the caspases (Shen et al., 2009), except that the CPD cleaves exclusively after a leucine instead of an aspartate residue. In contrast the molecular details of TcdB CPD substrate recognition remain uncharacterized.

In this study we used a combination of chemical synthesis and structural analyses to probe the substrate recognition and inhibitor sensitivity of TcdB CPD. By screening a focused library of substrate-based CPD inhibitors, we identified several compounds capable of blocking holotoxin function in cell culture. We also solved the structure of TcdB CPD bound to one of these inhibitors. Combined with the structure-activity relationship (SAR) series derived from our inhibitor analyses, these results provide a foundation for the development of therapeutics targeting this important virulence factor. We further used this information to develop activity-based probes (ABPs) specific for TcdB CPD that will permit the molecular dissection of its unique allosteric activation mechanism. The information presented here may also be valuable for the study of protease domains in other bacterial toxins.

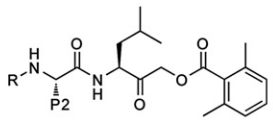
RESULTS

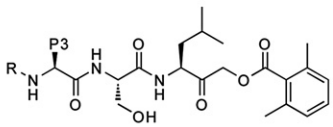
Inhibitor Design and Screening

The use of peptide-based inhibitors is an effective strategy for selectively inactivating proteases through mimicry of natural substrates (Berger et al., 2006; Kato et al., 2005; Powers et al., 2002). Given the importance of the CPD in regulating *C. difficile* glucosylating toxin function (Egerer et al., 2007; Reineke et al., 2007), we sought to identify inhibitors of the TcdB CPD protease. We first tested whether inhibitors specific for a related CPD found in *V. cholerae* MARTX ($MARTX_{Vc}$) toxin (Shen et al., 2009) could also inhibit TcdB CPD function (Figure 1). These inhibitors contain tripeptide sequences coupled to either an aza-epoxide or acyloxymethyl ketone (AOMK)-reactive electrophile (Figure 1A; see Figure S1 available online). Inhibitor potency against TcdB CPD was determined using a gel-based autocleavage assay, in which inhibitor concentration was varied in the presence of the activator $InsP_6$. The autocleavage substrate TcdB(1-804) used in this assay consists of TcdB's N-terminal 804 amino acids and contains the glucosyltransferase and CPDs and the natural autoprocessing site. The aza-epoxide $MARTX_{Vc}$ CPD inhibitors were only weakly inhibitory, with both Cbz-LLaL-EP and the related Cbz-EAaL-EP exhibiting observed IC_{50}'s greater than 100 μM (Figure 1B). In contrast the Cbz-EAL-AOMK inhibitor was significantly more potent, exhibiting an observed IC_{50} of 7.2 ± 0.7 μM. Because the primary difference between the Cbz-EAaL-EP and Cbz-EAL-AOMK inhibitors is the electrophilic reactive group, we reasoned that the AOMK electrophile is more optimal for TcdB CPD inhibition. Therefore, we synthesized a focused library of covalent AOMK inhibitors based on the natural substrate cleavage sequence of the *C. difficile* TcdB CPD (Figure 2A). These inhibitors consist of



B





R	P2	IC_{50} (μM)
Cbz	Gly	2.21 ± 0.10
	Ala	1.64 ± 0.08
	Leu	> 100
	Ser	1.55 ± 0.26
	Glu	33.92 ± 4.54
	Orn	1.60 ± 0.07
	Lys	1.06 ± 0.07
	Arg	1.26 ± 0.11
Ac	Ser	4.24 ± 0.42
	Lys	1.85 ± 0.15
Hpa	Ser	0.71 ± 0.05
	Lys	1.03 ± 0.06

R	P3	IC_{50} (μM)
Cbz	Gly	4.46 ± 0.47
Ac	Gly	1.78 ± 0.16
	Leu	4.95 ± 0.37
	Phe	7.12 ± 0.51
	hPhe	3.32 ± 0.26
	Bip	12.84 ± 2.87
	2-Nal	3.96 ± 0.21
	Tyr	7.47 ± 0.10
	Glu	12.72 ± 4.07
	Lys	2.98 ± 0.24
	Arg	1.31 ± 0.07

Figure 2. TcdB CPD Rational Inhibitor Design and Screening

(A) Conserved substrate autocleavage site of *C. difficile* TcdB and related bacterial toxins. The toxin CPD cleaves after the highlighted leucine residue.

(B) Focused screen of capped di- and tripeptide covalent TcdB CPD inhibitors. Observed IC_{50} values were determined using the autocleavage assay for covalent AOMK inhibitors with diverse P2 (left) and P3 (right) residues. These compounds were N terminally capped with Cbz, Ac, or Hpa groups. The dipeptide inhibitor Hpa-SL-AOMK was found to be the most potent compound. Data represent the mean of three experiments ± standard deviation. See also Figure S1.

dipeptide or tripeptide sequences coupled to the dimethylbenzoic acid AOMK, an electrophilic group that has been described as optimal for targeting the structurally related clan CD proteases, the caspases (Kato et al., 2005; Thornberry et al., 1994). For all inhibitors the P1 position (the residue N terminal to the scissile bond) was held constant as leucine because this is the primary substrate specificity determinant of bacterial CPDs (Egerer et al., 2007; Prochazkova et al., 2009; Shen et al., 2009). The P2 and P3 positions, as well as the N-terminal capping group, were varied in the library.

TcdB CPD Inhibitor SAR Profile

We first probed TcdB CPD P2 specificity using a diverse set of carboxybenzyl (Cbz)-capped dipeptide compounds (Figure 2B). Acidic and branched aliphatic amino acids were poorly tolerated in the P2 position, with calculated IC_{50}'s for Cbz-EL-AOMK and Cbz-LL-AOMK of 30 and 100 μM, respectively. Inhibitors containing smaller residues such as alanine and serine were more potent, with IC_{50}'s for Cbz-AL-AOMK and Cbz-SL-AOMK of 1.64 ± 0.08 μM and 1.55 ± 0.26 μM, respectively. Unexpectedly, compounds with basic residues in the P2 position were also potent, with an IC_{50} of 1.06 ± 0.07 μM observed for Cbz-KL-AOMK. We examined whether the enhanced potency of Cbz-KL-AOMK was specific to this residue or whether other amino acids with a positive charge could recapitulate this effect. Consistent with the latter interpretation, compounds with arginine and ornithine in the P2 position were also potent inhibitors (Figure 2B).

We next assessed the contribution of the Cbz cap to inhibitor recognition by synthesizing acetyl (Ac) and hydroxyphenyl acetyl (Hpa)-capped analogs of Cbz-KL-AOMK and Cbz-SL-AOMK. Although the smaller Ac group decreased potency compared to the Cbz cap for both compounds, the Hpa cap increased potency, resulting in an observed IC_{50} of 0.71 ± 0.05 μM for Hpa-SL-AOMK (Figure 2B). These results indicate that hydrophobic bulk in the dipeptide cap/ P3 binding position contributes to TcdB CPD inhibitor potency. Notably, the most potent compound (Hpa-SL-AOMK) contains amino acids found in the natural TcdB CPD substrate (Figure 2A).

Based on these results, we surveyed the P3 specificity of the protease domain using the same technique. We synthesized a focused library of Ac-capped tripeptide AOMK inhibitors containing the natural leucine and serine residues fixed in the P1 and P2 positions, respectively. The P3 position was varied using a diverse set of amino acids (Figure 2B). Additionally, given the favorable contribution of aromatic bulk in the dipeptide cap site, three nonnatural amino acids with an aromatic side chain were included in the P3 position (Figure 2; Figure S1B). Ac-capped tripeptide inhibitors containing small or basic residues in the P3 position were the most potent, whereas an acidic residue in the P3 position was poorly tolerated (Ac-ESL-AOMK) (Figure 2B). More diversity was tolerated in the P3 position of inhibitors containing the natural SL cleavage sequence in the P2/P1 positions. Bulky residues such as leucine were tolerated, with an IC_{50} of 4.95 ± 0.37 μM observed for Ac-LSL-AOMK. Homophenylalanine (hPhe) was the most potent of the inhibitors containing aromatic P3 residues, with an observed IC_{50} of 3.32 ± 0.26 μM. This is possibly because its methylene extension at the β-carbon position affords more flexibility in assuming productive interactions with the protease domain (Figure S1B). Conversely, aromatic bulk in the P4 position decreased inhibitor potency because the Cbz-capped GSL-AOMK inhibitor was ~3-fold less potent than the analogous inhibitor carrying the smaller Ac cap (4.46 ± 0.47 μM versus 1.78 ± 0.16 μM). However,

Table 1. Data Collection and Refinement Statistics for the TcdB CPD in Complex with InsP6 and Ac-GSL-AOMK

	TcdB CPD + Ac-GSL-AOMK
Data collection	
Space group	C 2
Cell dimensions	
α, β, χ (Å)	128.3, 45.5, 87.2
α, β, γ (°)	90, 103.5, 90
Wavelength (Å)	1.0
Resolution (Å)	50–2.0 (2.11–2.0)
R_{merge}	0.116 (0.534)
I / s*I*	9.2 (2.9)
Completeness (%)	100.0 (100.0)
Redundancy	6.7 (5.8)
Refinement	
Resolution (Å)	50.0–2.0
Number of reflections (total/test)	31,654/1,692
R_{work}/R_{free}	18.8/23.3
Number of atoms	
Protein	3,886
InsP6	72
Calcium	1
Sodium	2
Inhibitor	44
Water	286
B factors	
Protein	34.6
InsP6	29.1
Calcium	29.0
Sodium	34.8
Inhibitor	43.4
Water	41.0
Rmsd	
Bond lengths (Å)	0.010
Bond angles (°)	1.304

Highest resolution shell is in parentheses.

none of the compounds in the P3 library was more potent than the most potent dipeptide inhibitor, Hpa-SL-AOMK.

Crystal Structure of Inhibitor-Bound TcdB CPD

In order to rationalize the results of the SAR analyses and to gain structural insight into substrate recognition by TcdB CPD, we co-crystallized and solved the structure of $InsP_6$-bound CPD covalently inhibited with the Ac-GSL-AOMK inhibitor at 2.0 Å (Table 1). Although this inhibitor was not the most potent compound identified in the screen, it reflects the natural substrate cleavage site and exhibited improved solubility over dipeptide compounds carrying aromatic caps. The overall structure of inhibitor-bound, activated TcdB CPD (Figures 3A and 3B) is similar to the previously solved $InsP_6$-bound structure of TcdA CPD, which shares ~60% identity with TcdB CPD (rsmd of ~1 Å) (Pruitt et al., 2009).

As with most proteases, the substrate-binding pocket of TcdB CPD can be subdivided into multiple subsites. The catalytic residues are positioned between the S_1 and S_1' subsites; the numbering of the subsites reflects the corresponding substrate residue recognized, with the prime subsites interacting with substrate residues C terminal to the scissile bond. The most striking feature of the inhibitor structure is the insertion of the P1 leucine within a deep hydrophobic S1 pocket (Figure 3C). Eight residues, seven of which are nonpolar, are within van der Waals (4.4 Å) bonding distance of the P1 leucine side chain. Ile589 and Ala593 are contributed by helix 1 (Figure S2); Ile651 and Gly652 are contributed by strand D; Leu696 and Gly697 are contributed by strand E; and Val744 is contributed by strand G1 of the β-flap, a functional region that is involved in both $InsP_6$ recognition and substrate binding (Lupardus et al., 2008; Prochazkova et al., 2009; Pruitt et al., 2009). Ile746 from the G1 strand of the β-flap forms the distal side of the P1 pocket away from the active site, yet falls just outside van der Waals bonding distance in our structure. It likely also contributes to P1 recognition in vivo because the inhibitor may be pulled in slightly toward the catalytic cysteine due to the covalent nature of the modification. Because most of these residues are also conserved in the related $MARTX_{Vc}$ CPD, which binds to and cleaves after a P1 leucine (Prochazkova et al., 2009; Shen et al., 2009), bacterial CPDs would appear to share a common mechanism for recognition of this primary substrate specificity determinant.

Whereas many residues in the S1 subsite participate in recognition of the P1 Leu, residues in the other subsites make minimal interactions with the inhibitor (Figure 3C). In the S2 subsite the main chain carbonyl of Val744 forms a hydrogen bond with the P2 serine backbone amide of the inhibitor, whereas on the prime side the main chains of Gly654 and Cys698, which form the "oxyanion hole," hydrogen bond with the carbonyl formed after inhibitor reaction (Figure 3C). The P2 serine side chain points toward Glu743; in contrast the P3 glycine is oriented such that the P3 side chain would be exposed to solvent. Thus, the inhibitor structure of TcdB CPD reveals the mechanistic basis for substrate recognition.

Inhibitor Docking Findings

To gain additional insight into the inhibitor-specificity requirements of the active site, we used docking simulations to replace the Ac-GSL-AOMK in the TcdB CPD crystal structure with other compounds from the focused library. The structure of the modeled inhibitor and all side chains within 4.5 Å were energy minimized using the default parameter of the Molecular Operating Environment (MOE) software. The docking simulations suggest an explanation for why compounds containing basic amino acids such as lysine or arginine in the P2 position were particularly potent. In these analyses the positively charged P2 residue forms an electrostatic interaction with the acidic Glu743 residue, which helps to form the S2 pocket (Figure 4). This negatively charged subsite likely also explains why the Cbz-EL-AOMK was not as potent because this inhibitor cannot form the same favorable electrostatic interactions during binding (Figure 4). Instead, the acidic glutamate of the inhibitor is predicted to interact with Arg745; however, this interaction may be unfavorable given that Arg745 likely stabilizes the activated

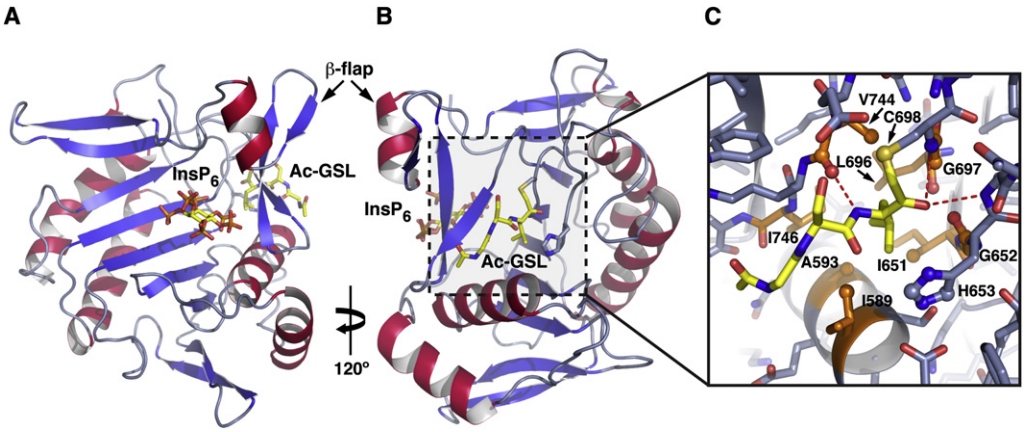


Figure 3. Structure of Activated TcdB CPD Bound to Ac-GSL-AOMK Inhibitor

(A) Ribbon structure of TcdB CPD in complex with $InsP_6$ viewed from above the $InsP_6$ binding site. $InsP_6$ is shown as a stick model.

(B) A view of the structure rotated ~120° to show inhibitor bound in the active site. $InsP_6$ and the inhibitor are shown as stick models. The β-flap hairpin that separates the $InsP_6$ binding and active sites is indicated.

(C) Close-up view of the substrate-binding pocket. Hydrophobic residues in the S1-binding pocket are shown as orange sticks, and the inhibitor is shown as yellow sticks. Side chains that interact with the P1 leucine are shown; hydrogen bonds are indicated by dotted lines. See also Figure S2.

conformation of the CPD through a π-cation interaction with the highly conserved Trp761 (Pruitt et al., 2009).

The docking studies also rationalize the increased potency of the dipeptide inhibitors capped with bulky Cbz or Hpa groups. These caps are predicted to fit into a hydrophobic groove formed between Ile746 and Ile589 such that the hydroxyl of the Hpa cap serves as the donor in a hydrogen bond with Ile746 (Figure 4). This may help favorably orient the inhibitors during binding and subsequent reaction with the catalytic Cys698.

TcdB CPD Inhibitor Blocks Full-Length Toxin Function

In order to determine if the inhibitors identified in vitro were functional in cells, we assessed their ability to block the intoxication of primary human foreskin fibroblasts (HFFs) by TcdB holotoxin. Pretreating cells with the most potent inhibitors prevented the cytopathic effects induced by recombinant TcdB holotoxin (Figure 5A). Cell rounding was quantified for select compounds by counting the number of rounded cells visible per field upon inhibitor titration (Figure 5B). The results corroborate the potency rankings observed in the initial autocleavage screen. Two of the most potent compounds, Hpa-SL-AOMK and Hpa-KL-AOMK, prevented cell rounding. Hpa-SL-AOMK completely inhibited toxin function by 100 μM, with an observed IC_{50} of approximately 20 μM, whereas the Hpa-KL-AOMK was slightly less potent. The difference in potency between Hpa-SL-AOMK and Hpa-KL-AOMK may be due to the increased cell permeability of the P2 serine relative to the positively charged lysine. Cell permeability of the inhibitors is likely to determine their potency because CPD inhibition can only occur after toxin entry into cells, when the CPD can become activated by $InsP_6$. In contrast the negative controls Cbz-EL-AOMK and the epoxide-based pan-cathepsin inhibitor JPM-OEt (Bogyo et al., 2000; Greenbaum et al., 2000) both failed to inhibit toxin function just as they poorly inhibit CPD activity (Figure 2B; Figure S1). Surprisingly, Cbz-KL-AOMK and to a lesser extent Cbz-SL-AOMK were found to be cytotoxic (Figure S3). Given that no toxicity was observed for the Hpa-capped analogs, these findings implicate the Cbz cap in affecting cell viability.

We confirmed that inhibition of the TcdB CPD directly prevented holotoxin effector domain activity by monitoring glucosylation of its cellular target Rac1, a Rho GTPase (Genth et al., 2006; Yang et al., 2008). Addition of wild-type toxin to cells resulted in complete glucosylation of Rac1, whereas pretreatment with either Hpa-SL-AOMK or Hpa-KL-AOMK was protective (Figure 5C). Importantly, inhibition of cathepsins using JPM-OEt had no effect on toxin function; JPM-OEt was used as a control because P1 Leu AOMKs have previously been shown to weakly cross-react with the cathepsins (Kato et al., 2005).

Probe Design to Monitor Toxin Activation

The activation of the CPD by the eukaryotic-specific small molecule $InsP_6$ is a critical step in regulating the function and trafficking of *C. difficile* glucosylating toxins (Egerer and Satchell, 2010; Shen, 2010). In order to facilitate more detailed studies of this important process, we created fluorescently tagged and biotin-tagged probe versions of the SL-AOMK inhibitor (AWP19 and AWP15, respectively) to visualize toxin activation (Figure 6A). Both AWP19 and AWP15 covalently label the recombinantly produced TcdB(1-804) autocleavage substrate in response to small molecule activation of the protease by $InsP_6$ (Figure 6B). This substrate represents the toxin region predicted to translocate into the cytosol during intoxication. Notably, the cleavage product TcdB(544-804) formed after proteolytic cleavage was more active than the full-length substrate, as shown by the increase in ABP labeling relative to the amount of protein present (Figures 6B and 6C). This observation suggests that the isolated CPD TcdB(544-804) is either more accessible or more reactive with the probe than the full-length substrate. As expected, the ABPs failed to label the catalytically inactive C698A mutant of TcdB(1-804) in the presence or absence of $InsP_6$.

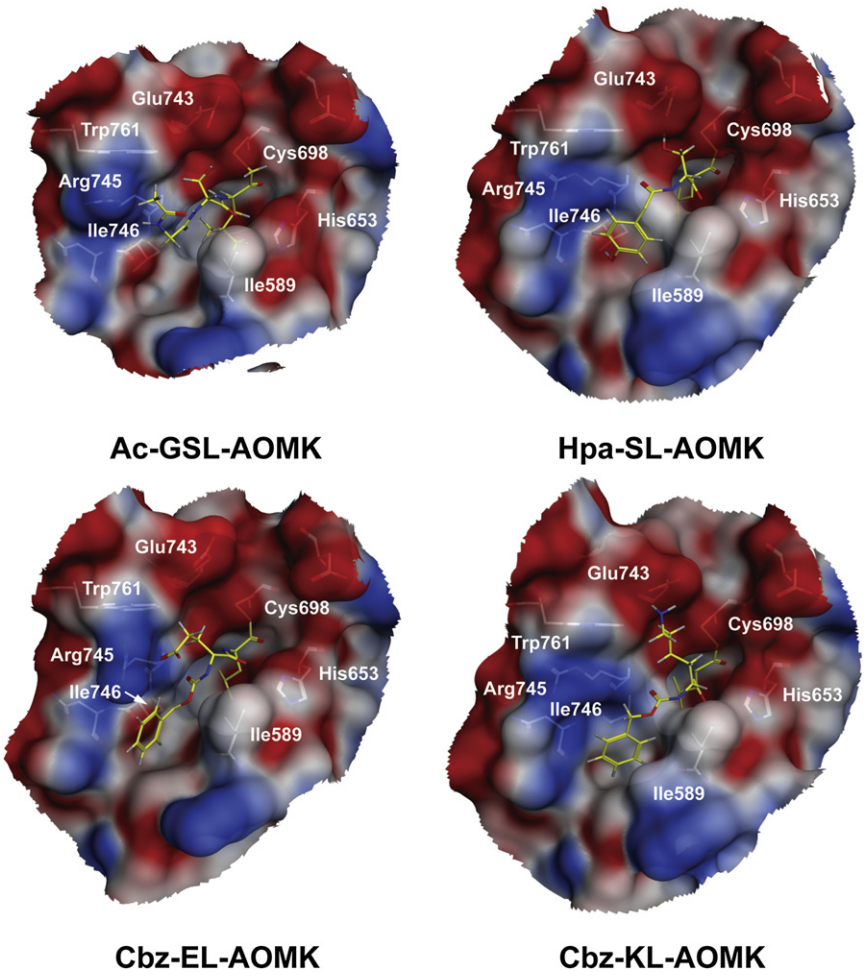


Figure 4. Molecular-Docking Analyses of TcdB CPD with AOMK Inhibitors

Close-up view of the substrate-binding pocket shown with electrostatic surface potential. Blue denotes positively charged surface; red denotes negatively charged surface. Inhibitors are shown as yellow sticks, and relevant residues are indicated. The Ac-GSL-AOMK image is derived from the crystal structure of the inhibitor-bound enzyme. Binding of the Cbz-EL-AOMK is predicted to bind TcdB CPD differently from the other inhibitors.

We also determined that the ABPs could label functional recombinant holotoxin, which was either produced in *E. coli* or in the *Bacillus megaterium* expression system (Pruitt et al., 2010; Yang et al., 2008) (Figure 6C). The *B. megaterium* system produces a His_6-tagged TcdB holotoxin that is more pure than the native toxin purified from *C. difficile* culture supernatants (Genth et al., 2006; Yang et al., 2008). Both probes could sensitively detect active CPD within the toxin because the cleaved TcdB(544-2366)-His_6 fragment could be labeled by the probe even though it is not detectable by either Coomassie or anti-His_6 antibody-conjugated horseradish peroxidase. Titration of probe labeling confirmed the potency of the probes, with AWP19 labeling active holotoxin at probe concentrations below 50 nM (Figure S4). This level of sensitivity will be valuable for tracking CPD activation during toxin trafficking because it permits the labeling of ensembles of toxin while minimizing the risk of completely inhibiting toxin function.

Monitoring toxin trafficking is further enabled by the cell permeability of AWP19's near-infrared fluorescent cyanine 5 tag, which allows the probe to be used in intact cells. In order to check for off-targets of this probe, we incubated both primary HFFs and the RAW macrophage cell line with AWP19. In both cell types the only off-target the probe labeled was cathepsin B, as confirmed by comparison with a pan-cathepsin ABP (Figure S5) (M.G. Paulick and M.B., unpublished data). Similarly, pretreatment of both cell types with Cbz-SL-AOMK prevented cathepsin B labeling by the pan-cathepsin probe, but not cathepsin L or X labeling. Thus, cathepsin B would appear to be the primary off-target TcdB CPD inhibitors. Nevertheless, cathepsin B inhibition was not sufficient to reduce TcdB toxin function because the pan-cathepsin inhibitor JPM had no effect on toxin-induced cell rounding or Rac1 glucosylation (Figure 5). This result strongly suggests that the observed reduction in TcdB-glucosylating activity in target cells upon treatment with CPD inhibitors is due to inhibition of CPD function.

DISCUSSION

The rising rate of *C. difficile* infections necessitates the development of new classes of therapeutics to combat this pathogen. Because of its natural antibiotic resistance, there has been increased focus on targeting the glucosylating toxins TcdA and TcdB for direct therapeutic intervention because they are the primary mediators of *C. difficile* pathogenesis (Halsey, 2008; Kelly and LaMont, 2008; Rupnik et al., 2009). In this study, to our knowledge, we present the first validation that the TcdB CPD is a druggable target. Although inhibition of this protease active site is difficult because the small molecule is competing with an intramolecular autoproteolytic event, our findings are encouraging for the development of competitive inhibitors for the TcdB CPD. The most potent compound in our library is the 499 Da-capped dipeptide inhibitor Hpa-SL-AOMK (Figure 2), which is within the size constraints generally accepted for therapeutics (Lipinski, 2000). In addition the minimal interaction between the protease and inhibitor peptide backbone (Figure 4) suggests that inhibitors with non-peptidic scaffolds can be developed to bypass the pharmacokinetic shortfalls of peptidic compounds.

Our rational approach to probing the inhibitor sensitivity of the CPD active site using structural analysis and a focused library of substrate-based compounds yielded multiple inhibitors capable of blocking holotoxin function (Figure 5). These analyses

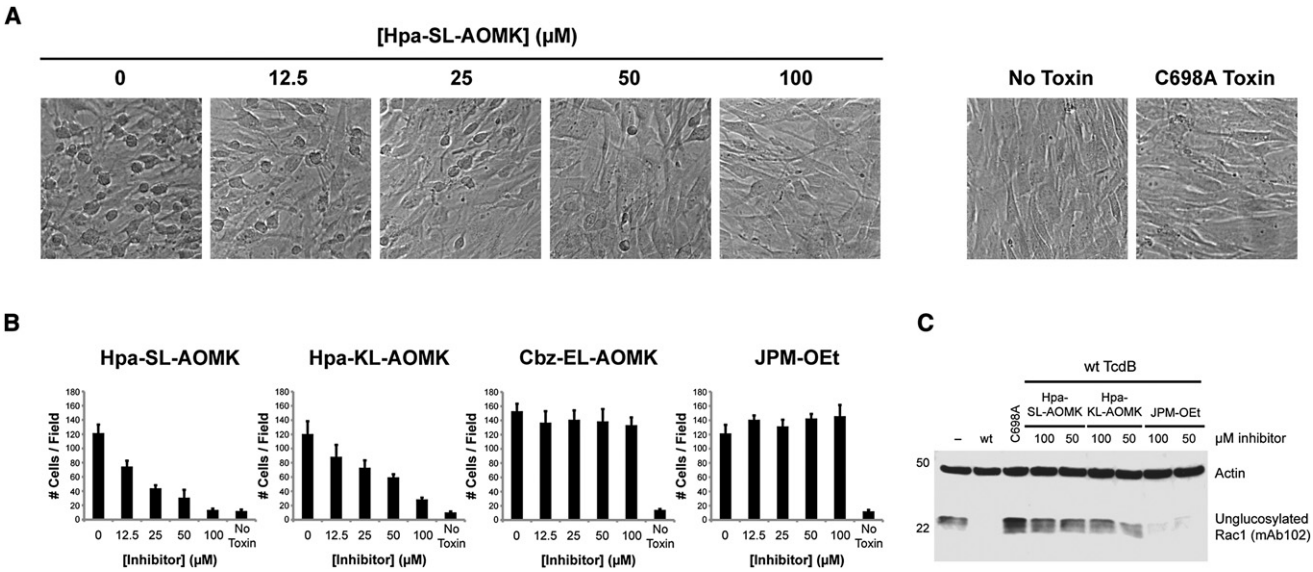


Figure 5. TcdB CPD Inhibitors Block Holotoxin Function

(A) Primary HFFs pretreated with Hpa-SL-AOMK are protected from holotoxin-mediated cell rounding. HFFs treated with the catalytically inactivated C698A holotoxin or left untreated exhibit minimal cell rounding.

(B) Quantification of inhibitor effects on holotoxin-mediated HFF cell rounding. Pretreatment with Hpa-SL-AOMK or Hpa-KL-AOMK was protective, whereas Cbz-EL-AOMK and the pan-cathepsin inhibitor JPM-OEt had no effect on toxin function. Data represent the mean of three experiments ± standard deviation.

(C) Addition of wild-type TcdB holotoxin to HFFs results in complete glucosylation of the Rho GTPase Rac1, as seen by western blot with the glucosylation-sensitive α-Rac1 monoclonal antibody mAb102. Pretreating the cells with Hpa-SL-AOMK or Hpa-KL-AOMK protected HFFs from toxin effector domain activity, whereas inhibiting cathepsin activity with JPM-OEt did not. See also Figure S3.

produced the unexpected observation that inhibitors containing basic P2 or P3 residues and bulky hydrophobic N-terminal caps could potently block TcdB autoprocessing. This SAR profile provides a starting point for the development of compounds suitable for therapeutic applications. We note that relying solely on substrate-specificity profiling for these domains may not have produced such promising results because potent substrates do not always translate into viable inhibitors (Drag et al., 2010). Furthermore, fluorogenic substrate cleavage assays lack sensitivity in detecting bacterial CPD activity because these autoprocessing enzymes exhibit poor transcleavage efficiency (Babe and Craik, 1997; Lupardus et al., 2008). Our attempts to develop optimized substrates for TcdB CPD produced substrates with poor K_m values (~1 mM; data not shown). The approach described here may also prove applicable to other protease domain-containing bacterial toxins (Lebrun et al., 2009).

Information about the inhibitor sensitivity profile of TcdB CPD was bolstered by our crystal structure of the enzyme bound to the Ac-GSL-AOMK inhibitor (Figure 3). This structure permitted the molecular-docking studies that helped rationalize the increased potency of inhibitors with basic P2 residues and bulky hydrophobic N-terminal caps (Figure 4). Furthermore, given the overall similarity between the inhibitor structure of $InsP_6$-bound TcdB CPD presented here and $InsP_6$-bound TcdA CPD (Pruitt et al., 2009), many of our findings may be translatable to inhibiting this closely related toxin. Our most efficacious inhibitor Hpa-SL-AOMK will likely exhibit similar potency against TcdA due to its identical P1/P2 substrate sequence (Figure 2A), increasing its value as a *C. difficile* virulence-targeting agent.

This crystal structure also permits comparisons to be made with the inhibitor-bound structure of the related CPD of *Vibrio cholerae* MARTX toxin (Shen et al., 2009). Both inhibitor structures reveal that the primary substrate specificity determinant for these proteases is the P1 leucine, and residues involved in recognizing this leucine are conserved in both proteases. The inhibitor structures differ in that $MARTX_{Vc}$ CPD makes more backbone interactions with its inhibitor than TcdB CPD and in the S1′ subsite, the region that recognizes residues C terminal to the scissile bond. For $MARTX_{Vc}$ CPD this region is relatively flat and featureless, whereas in TcdB CPD, Asp656 and Glu657 directly extend into this region and may thus occlude substrate or inhibitor binding (Figure S6). This acidic extension may explain why such a large difference in potency was observed between the Cbz-EAL aza-epoxide and AOMK derivatives for TcdB CPD (Figure 1), whereas both warheads inhibit $MARTX_{Vc}$ CPD to a similar extent (Shen et al., 2009).

The CPDs of both the *C. difficile* large glucosylating and MARTX toxin families appear to "sense" the eukaryotic cell environment and activate toxin function accordingly (Egerer and Satchell, 2010; Shen, 2010). However, this unique allosteric activation process is difficult to study using traditional biochemical approaches because it is posttranslationally regulated. Our ABPs overcome this problem because they afford the direct visualization of CPD activation by $InsP_6$ in complex mixtures and possibly in vivo. Furthermore, the probes provide a robust readout for CPD activity that could be used in screening applications for competitive (Schneider and Craik, 2009) or allosteric inhibitors (Lee and Craik, 2009) of the CPD, in lieu of fluorogenic

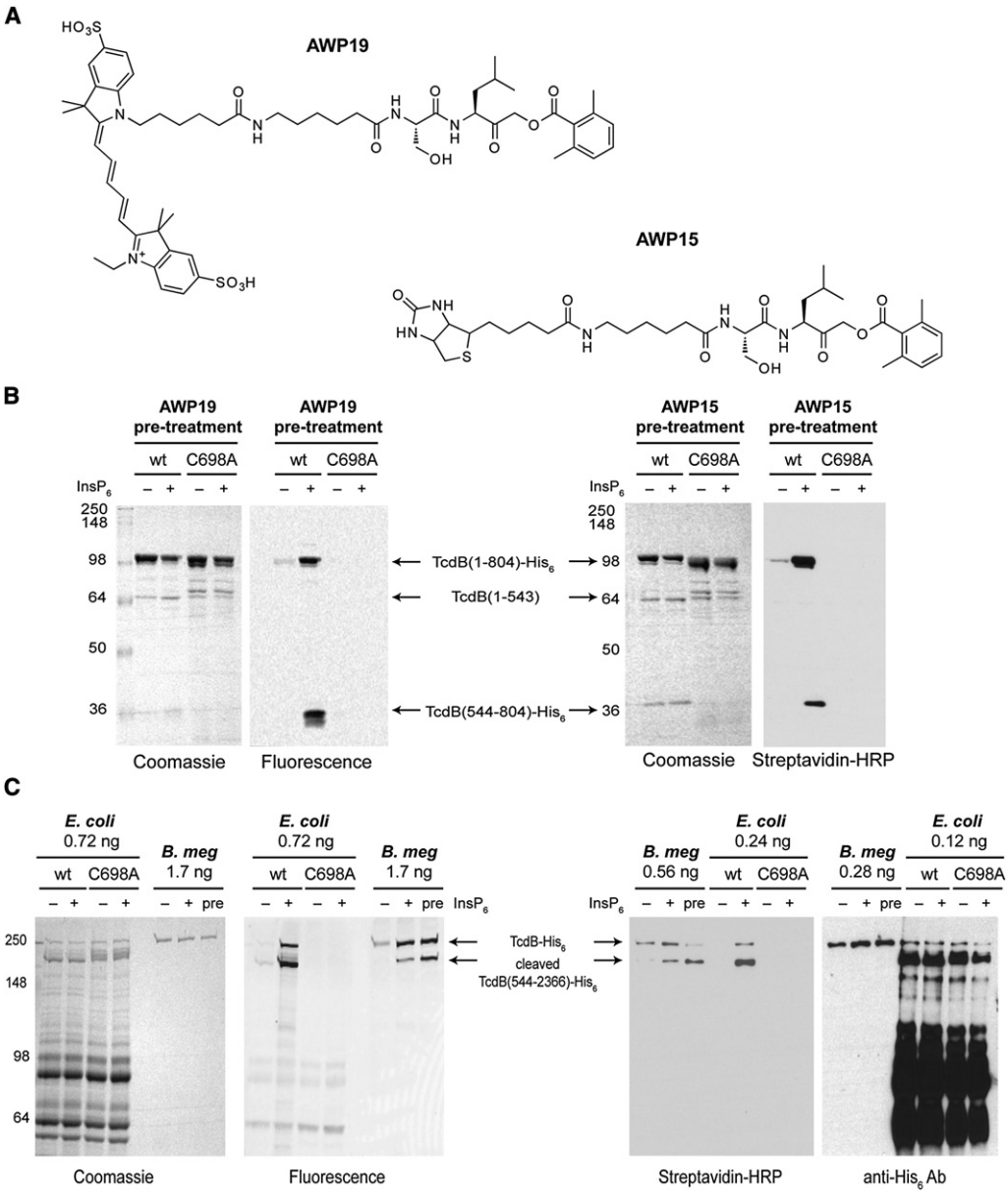


Figure 6. Probe Labeling of Recombinant TcdB

(A) Structures of Cy5-labeled AWP19 and biotin-labeled AWP15 TcdB CPD probes.

(B) Labeling of either wild-type or catalytic mutant C698A TcdB(1-804) (0.5 μM) by probes (10 μM) in the presence (25 μM, +) or absence (–) of $InsP_6$. Fluorescence scanning was used to detect AWP19 labeling, whereas streptavidin-HRP blotting was used to detect AWP15 labeling; total protein was detected by Coomassie staining. A small fraction of cleaved $TcdB_{544-804}$ is detectably labeled by both probes even though it is not detectable by Coomassie staining.

(C) Labeling of holotoxin produced in *E. coli* (wild-type and catalytic mutant C698A) and *Bacillus megaterium* by probes (10 μM) in the presence (25 μM, +) or absence (–) of $InsP_6$. The amount of protein loaded is indicated. "Pre" refers to *B. megaterium*-produced holotoxin that was pretreated with 25 μM $InsP_6$ (to induce autoprocessing) for 1 hr at 37°C prior to labeling, after which 10 μM of probe was added to the sample. Fluorescence scanning was used to detect AWP19 labeling, whereas streptavidin-HRP blotting was used to detect AWP15 labeling. Total protein was visualized using Coomassie staining, whereas His_6-tagged holotoxin was visualized using anti-His antibody western blotting. See also Figures S4 and S5.

substrates (which react poorly with the protease) or autocleavage assays (which are less sensitive).

Because of the covalent nature of these probes, they provide a direct readout of when the toxin has encountered $InsP_6$. This is valuable within the context of studies directed at dissecting toxin trafficking and the molecular mechanisms underlying the allosteric regulation of toxin function by $InsP_6$ (Giesemann et al., 2008; Jank and Aktories, 2008). For example it is notable that the cleaved form of TcdB CPD is more effectively labeled by the ABP than the full-length protein (Figure 6). This may imply that the CPD is held in an inhibitory conformation within the native holotoxin and that autoproteolytic cleavage relieves this

inhibition. Alternatively, the active site may be more accessible to the ABP following cleavage, which would suggest that the conformation of full-length TcdB in vitro occludes substrate binding. *C. difficile*-glucosylating toxins undergo significant conformational rearrangements during the pH-dependent toxin translocation process (Pruitt et al., 2010; Qa'Dan et al., 2000). Based on our observation that cleaved TcdB toxin (aa 544-2366) was also more readily labeled by the probe in the presence of $InsP_6$ (Figure 6C), it is tempting to speculate that the CPD is subject to additional regulation at the level of toxin conformation. Further studies using these promising tools will provide a more detailed understanding of the regulation of the CPD by $InsP_6$ and in the context of the full-length toxin.

SIGNIFICANCE

The large glucosylating toxins TcdA and TcdB are the primary virulence factors of the antibiotic-resistant bacterium *Clostridium difficile*. These toxins are autoproteolytically activated by an internal cysteine protease domain (CPD) in a step that is critical for toxin function. We synthesized a focused library of substrate-based compounds in order to determine a structure-activity relationship for TcdB CPD inhibitors and then gained further insight by co-crystallizing the domain with one of these inhibitors. This rational approach yielded compounds potent enough to inhibit toxin function in cell culture, validating the clostridial-glucosylating toxins as druggable targets. Our results provide a promising starting point for the development of therapeutics that minimize the selective pressure on *C. difficile* to develop resistance. We also used the inhibitor data to develop covalent activity-based probes (ABPs) that can directly measure the allosteric activation of the protease by the small molecule inositol hexakisphosphate. Because these ABPs monitor the posttranslational activation of the toxin, they will be useful in studies directed at understanding this unique regulatory mechanism in both biochemical and cell-based assays.

EXPERIMENTAL PROCEDURES

Compound Synthesis

The aza-leu epoxide inhibitors were synthesized in solution using standard chemistry as previously described (Asgian et al., 2002). The AOMK inhibitors were synthesized using solid-phase synthesis as previously described (Kato et al., 2005). The ABP AWP19 was synthesized by combining H_2N-aminohexanoic-SL-AOMK (1 equivalent) with Cy5-NHS (0.9 equivalents) and DIEA (5 equivalents) in DMSO for 1 hr and then purifying directly by HPLC. The identity and purity of all compounds were characterized using HR-MS and LCMS.

Protein Expression and Purification

An overnight culture of pET28a-$TcdB_{1-804}$ was diluted 1:500 into 4 L 2YT media and grown shaking at 37°C. When an OD_{600} of 0.6–0.9 was reached, IPTG was added to 250 μM, and cultures were grown for 3 hr at 225 rpm at 30°C. Cultures were pelleted, resuspended in 60 ml lysis buffer (500 mM NaCl, 50 mM Tris-HCl [pH 7.5], 15 mM imidazole, and 10% glycerol) and flash frozen in liquid nitrogen. Lysates were thawed, then lysed by sonication and cleared by centrifugation at 15,000 × g for 30 min. C-terminally His_6-tagged TcdB(1-804) was affinity purified by incubating the lysates in batch with 2.0 ml Ni-NTA Agarose beads (QIAGEN) with shaking for 3 hr at 4°C. The binding reaction was pelleted at 1500 × g, and the pelleted Ni-NTA agarose beads were washed three times with lysis buffer. His_6-tagged CPD was eluted from the beads by the addition of 400 μl high imidazole elution buffer (500 mM NaCl, 50 mM Tris-HCl [pH 7.5], 175 mM imidazole, and 10% glycerol). The elution was repeated three times; the eluate was pooled, buffer exchanged in gel-filtration buffer (200 mM NaCl, 10 mM Tris [pH 7.5], and 5% glycerol), and concentrated to 750 μl. The concentrated prep was pelleted at 13,000 × g for 10 min at 4°C prior to loading on a HiPrep S200 16/60 Sephacryl column (GE Healthcare). Purified His_6-tagged CPD was concentrated and stored at −20°C in gel-filtration buffer.

TcdB Autocleavage Assay

Recombinant TcdB(1-804) was diluted to a final concentration of 0.5 μM in assay buffer (60 mM NaCl, 20 mM Tris [pH 7.5], and 250 mM sucrose) in a 96 well plate. A total of 0.5 μl of a 100× inhibitor stock in DMSO was then added to each well in triplicate, and the samples were incubated at 37°C for 30 min. $InsP_6$ (0.5 μl, Calbiochem) was then added to a final concentration of 25 μM, and the reaction was incubated at 37°C for 1 hr. Samples were then diluted in 4× SDS-PAGE loading buffer and resolved by SDS-PAGE on 12% gels. Cleavage reactions were visualized by Coomassie staining and quantified using the program ImageJ (http://rsb.info.nih.gov/ij/). For each sample the amount of autocleaved protein relative to the total protein amount was plotted versus the concentration of inhibitor and globally fit using the sigmoidal function in KaleidaGraph.

Protein Purification, Crystallization, and Data Collection

An overnight culture of pET22b-$TcdB_{544-797}$ was diluted 1:500 into 3 liters of 2YT media and grown shaking at 37°C. When an OD_{600} of 0.6–0.9 was reached, IPTG was added to 250 μM, and cultures were grown for 12–16 hr (225 rpm) at 18°C–20°C. Cultures were pelleted, resuspended in 50 ml lysis buffer (500 mM NaCl, 50 mM Tris-HCl [pH 7.5], 15 mM imidazole, and 10% glycerol), and flash frozen in liquid nitrogen. Lysates were thawed, then lysed by sonication and cleared by centrifugation at 15,000 × g for 30 min. $TcdB_{544-797}$ was purified as described above except that it was concentrated to 1 mM, and the gel-filtration buffer was 150 mM NaCl, 10 mM Tris (pH 7.5). Gel-filtration purified TcdB(544-797) was treated with 2 mM $InsP_6$ and 2 mM Ac-GSL-AOMK (inhibitor stock was at 200 mM in DMSO). The inhibitor was added slowly due to poor solubility; reaction with the protease improved inhibitor solubility. The reaction was allowed to proceed for 1 hr at room temperature after which excess inhibitor was pelleted by centrifuging the reaction at 13,000 × g for 10 min at 4°C. Crystallization screening was carried out using the sitting drop vapor-diffusion method, and initial hits were observed in 0.1 M Tris HCl (pH 8.0) and 30% PEG2000 MME as the precipitant. Crystals used for data collection were grown in 1 μl drops by mixing equal volumes of protein with mother liquor and appeared only after ~45–60 days. Crystals were cryoprotected in 45% PEG2000 MME, flash frozen in liquid nitrogen, and data collected under cryo-cooled conditions at 100 K at beamline 8.2.1 at the Advanced Light Source (University of California-Berkeley). Diffraction data were processed using MOSFLM (Leslie, 1991) and SCALA (Potterton et al., 2003), and processing statistics are listed in Table 1.

Structure Determination and Refinement

Initial phases were obtained by molecular replacement with PHASER (McCoy et al., 2007) using the structure of the TcdA CPD (PDB ID 3HO6) as a search model (Pruitt et al., 2009). Using the molecular replacement phases, the TcdB CPD was built by ARP/wARP (Perrakis et al., 1999) to approximately 85% completeness, followed by rounds of model building and adjustment with COOT (Emsley and Cowtan, 2004) and refinement with PHENIX (Adams et al., 2002). Restraints for the Ac-GSL molecule were obtained from the PRODRG server (Schuttelkopf and van Aalten, 2004). The final model underwent restrained and translation/libration/screw refinement in REFMAC5 (Murshudov et al., 1997), resulting in final R/Rfree values of 18.8% and 23.3%. Ramachandran analysis of model geometry by MolProbity (http://molprobity.biochem.duke.edu/) indicates that 99.0% of residues reside in the most favorable regions, with none in the disallowed regions. Refinement statistics can be found in Table 1. Structural figures were prepared with PyMOL (DeLano, 2002). The final model contains two TcdB CPD molecules in the asymmetric unit (ASU), with each bound to one $InsP_6$ molecule, one Ac-GSL inhibitor molecule, two sodium ions, and a calcium ion bridging the crystal contact between the

molecules in the ASU. Chain A of the Tcd CPD is used for all figures in the paper.

Docking Simulations

The different homology models of TcdB CPD were built from the crystal structure of the protease bound of Ac-GSL-AOMK using the default parameters of the MOE software. In each model the covalent linkage to Cys698 and the position of leucine in the P1 pocket were initially fixed as those of the crystal structure. The N-terminal cap and P2 side chain were manually built into the active site of the protease. Energy minimization was performed first on the modified region of the inhibitor and all side chains within 4.5 Å, and second on the entire inhibitor and all side chains within 4.5 Å.

Cell Rounding Assay

Primary HFFs were seeded in 96 well plates at a density of 1–2 × 10^4 cells/well in DMEM supplemented with 10% FBS. Prior to the assay the cells were washed three times with 100 μl DMEM alone. One microliter of a 100× inhibitor stock in DMSO was then added to each well in triplicate, and the cells were incubated at 37°C for 30 min. Recombinant TcdB holotoxin expressed in *E. coli* was then added to the cells at a final concentration of 0.3 pM. The samples were incubated at 37°C for 2 hr and then imaged using a 20× objective on an inverted microscope. Four fields were imaged per well, and the average number of rounded cells per field was calculated.

Rac1 Glucosylation Assay

HFF cells were seeded into 24 well treated plates (7.5 × 10^5) and grown to 100% confluency overnight, washed once with pre-warmed DMEM, and left in 0.25 ml DMEM per well. The indicated concentration of inhibitor was added as a 1:100 dilution from a DMSO stock and incubated for 30 min. Recombinant TcdB holotoxin purified from *E. coli* was added to the cells (0.3 pM) and incubated for 90 min at 37°C. The media were removed and then the cells were lysed in 25 μl 1× FSB by scraping the cells in concentric circles. Cell lysates were boiled for 5 min at 95°C, and 15 μl of lysate was resolved on a 14% SDS-PAGE gel and transferred to nitrocellulose. Unglucosylated Rac1 was detected using a 1:1000 dilution of mAb102 (Millipore) and a 1:5000 dilution of anti-mouse IgG HRP (BioRad). Actin was simultaneously visualized using a polyclonal anti-actin antibody at 1:2000 dilution (Sigma), and a 1:5000 dilution of anti-rabbit IgG HRP (BioRad).

Purification of TcdB Holotoxin from *E. coli*

Overnight cultures of pET28a-TcdB wild-type or C698A holotoxin were diluted 1:500 into 3 liters of 2YT media and grown shaking at 37°C. C-terminally His_6-tagged holotoxin was purified as described for His_6-tagged TcdB(1-804) with the exception that β-mercaptoethanol was added to the lysis buffer at 2 mM.

AWP19 and AWP15 Labeling of Recombinant TcdB(1-804)

Wild-type and C698A TcdB 1-804 (0.5 μM) were incubated with 10 μM of the indicated probe. $InsP_6$ was added to a final concentration of 25 μM (1:100 dilution) in a total volume of 50 μl where indicated, and probe labeling proceeded for 1 hr at 37°C. Fifteen microliters of 4× final sample buffer were added and then the sample was boiled for 3 min at 95°C. The samples were resolved on a 10% gel by SDS-PAGE. For AWP19 labeling, fluorescent labeling was visualized using a fluorescent scanner followed by Coomassie staining. For AWP15 staining the sample was loaded in duplicate, and one set was visualized by Coomassie staining (5 μl sample loaded), whereas the other set was transferred to nitrocellulose (2.5 μl sample loaded) and blotted using streptavidin-HRP (Sigma) at 1:3000.

AWP19 and AWP15 Labeling of Recombinant TcdB

For labeling of TcdB holotoxin purified from *E. coli*, wild-type and C698A TcdB was diluted to 0.3 μM in assay buffer and then the indicated probe was added to 10 μM (1:100 dilution from DMSO stock) in a total volume of 15 μl. $InsP_6$ was added at a final concentration of 25 μM where indicated. For labeling of TcdB holotoxin purified from *B. megaterium* (a gift from R. Pruitt and D.B. Lacy), 0.5 μM of toxin was diluted in CPD buffer and then 10 μM of the indicated probe was added (1:100 dilution from DMSO stock). $InsP_6$ was then added at a final concentration of 25 μM. Alternatively, $InsP_6$ (25 μM) was added to the *B. megaterium*-produced toxin (0.5 μM) and incubated for 1 hr at 37°C and then the indicated probe was added at 10 μM final concentration. Following probe addition, the labeling reactions were allowed to proceed for 1 hr at 37°C, after which 5 μl of 4× FSB was added, and the samples were boiled for 3 min at 95°C. For AWP19 labeled samples, 15 μl was resolved on a 10% gel and then visualized by fluorescence scanning followed by Coomassie staining. For AWP15-labeled samples the samples (either 5 or 2.5 μl) were loaded in duplicate and resolved on a 10% gel, then transferred to nitrocellulose. The membranes were probed with Streptavidin-HRP (Sigma) at 1:3,000 (5 μl sample loaded) or with an anti-His antibody (Pierce) and anti-rabbit IgG HRP (BioRad) at 1:10,000 (2.5 μl sample loaded).

ACCESSION NUMBERS

Coordinates and structure factors have been deposited in the Protein Data Bank (www.rcsb.org) under accession number 3PA8.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at doi:10.1016/j.chembiol.2010.09.011.

ACKNOWLEDGMENTS

We thank D. Borden Lacy and Rory Pruitt (Vanderbilt University) for the generous gift of TcdB toxin produced in *B. megaterium* and Elaine Hamm and Jimmy Ballard (Oklahoma Health Sciences University) for *Clostridium difficile* VPI 10463 genomic DNA. A.W.P. is supported by an NSF Graduate Research Fellowship. P.J.L. is a Damon Runyon Fellow, supported by the Damon Runyon Cancer Research Foundation. K.C.G. is supported by the Keck Foundation and the Howard Hughes Medical Institute. M.B. is supported by the Burroughs Wellcome Foundation and NIH grants R01EB005011 and R01AI078947. A.S. is supported by an NIH National Institutes of General Medical Sciences 1-K99-GM092934-01.

Received: August 31, 2010
Revised: September 30, 2010
Accepted: September 30, 2010
Published: November 23, 2010

REFERENCES

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.

Asgian, J.L., James, K.E., Li, Z.Z., Carter, W., Barrett, A.J., Mikolajczyk, J., Salvesen, G.S., and Powers, J.C. (2002). Aza-peptide epoxides: a new class of inhibitors selective for clan CD cysteine proteases. J. Med. Chem. *45*, 4958–4960.

Babe, L.M., and Craik, C.S. (1997). Viral proteases: evolution of diverse structural motifs to optimize function. Cell *91*, 427–430.

Berger, A.B., Sexton, K.B., and Bogyo, M. (2006). Commonly used caspase inhibitors designed based on substrate specificity profiles lack selectivity. Cell Res. *16*, 961–963.

Bogyo, M., Verhelst, S., Bellingard-Dubouchaud, V., Toba, S., and Greenbaum, D. (2000). Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs. Chem. Biol. *7*, 27–38.

Clatworthy, A.E., Pierson, E., and Hung, D.T. (2007). Targeting virulence: a new paradigm for antimicrobial therapy. Nat. Chem. Biol. *3*, 541–548.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System (San Carlos, CA: DeLano Scientific).

Drag, M., Bogyo, M., Ellman, J.A., and Salvesen, G.S. (2010). Aminopeptidase fingerprints, an integrated approach for identification of good substrates and optimal inhibitors. J. Biol. Chem. *285*, 3310–3318.

Egerer, M., and Satchell, K.J. (2010). Inositol hexakisphosphate-induced autoprocessing of large bacterial protein toxins. PLoS Pathog. *6*, e1000942.

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.

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

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.

Genth, H., Huelsenbeck, J., Hartmann, B., Hofmann, F., Just, I., and Gerhard, R. (2006). Cellular stability of Rho-GTPases glucosylated by *Clostridium difficile* toxin B. FEBS Lett. *580*, 3565–3569.

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.

Gerhard, R., Nottrott, S., Schoentaube, J., Tatge, H., Olling, A., and Just, I. (2008). Glucosylation of Rho GTPases by *Clostridium difficile* toxin A triggers apoptosis in intestinal epithelial cells. J. Med. Microbiol. *57*, 765–770.

Giesemann, T., Egerer, M., Jank, T., and Aktories, K. (2008). Processing of *Clostridium difficile* toxins. J. Med. Microbiol. *57*, 690–696.

Greenbaum, D., Medzihradszky, K.F., Burlingame, A., and Bogyo, M. (2000). Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. Chem. Biol. *7*, 569–581.

Halsey, J. (2008). Current and future treatment modalities for *Clostridium difficile*-associated disease. Am. J. Health Syst. Pharm. *65*, 705–715.

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

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.

Kato, D., Boatright, K.M., Berger, A.B., Nazif, T., Blum, G., Ryan, C., Chehade, K.A., Salvesen, G.S., and Bogyo, M. (2005). Activity-based probes that target diverse cysteine protease families. Nat. Chem. Biol. *1*, 33–38.

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

Lebrun, I., Marques-Porto, R., Pereira, A.S., Pereira, A., and Perpetuo, E.A. (2009). Bacterial toxins: an overview on bacterial proteases and their action as virulence factors. Mini Rev. Med. Chem. *9*, 820–828.

Lee, G.M., and Craik, C.S. (2009). Trapping moving targets with small molecules. Science *324*, 213–215.

Leslie, A.G. (1991). Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No. 26.

Lipinski, C.A. (2000). Drug-like properties and the causes of poor solubility and poor permeability. J. Pharmacol. Toxicol. Methods *44*, 235–249.

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.

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., et al. (2009). Toxin B is essential for virulence of *Clostridium difficile*. Nature *458*, 1176–1179.

McCoy, A.J., Grosse-Kunztleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. *6*, 458–463.

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.

Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003). A graphical user interface to the CCP4 program suite. Acta Crystallogr. D Biol. Crystallogr. *59*, 1131–1137.

Powers, J.C., Asgian, J.L., Ekici, O.D., and James, K.E. (2002). Irreversible inhibitors of serine, cysteine, and threonine proteases. Chem. Rev. *102*, 4639–4750.

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.

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. Proc. Natl. Acad. Sci. USA *107*, 13467–13472.

Puri, A.W., and Bogyo, M. (2009). Using small molecules to dissect mechanisms of microbial pathogenesis. ACS Chem. Biol. *4*, 603–616.

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.

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.

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.

Rolfe, R.D., and Song, W. (1993). Purification of a functional receptor for *Clostridium difficile* toxin A from intestinal brush border membranes of infant hamsters. Clin. Infect. Dis. *16* (*Suppl 4*), S219–S227.

Rupnik, M., Wilcox, M.H., and Gerding, D.N. (2009). *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat. Rev. Microbiol. *7*, 526–536.

Satchell, K.J. (2007). MARTX, multifunctional autoprocessing repeats-in-toxin toxins. Infect. Immun. *75*, 5079–5084.

Schneider, E.L., and Craik, C.S. (2009). Positional scanning synthetic combinatorial libraries for substrate profiling. Methods Mol. Biol. *539*, 59–78.

Schuttelkopf, A.W., and van Aalten, D.M. (2004). PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallogr. D Biol. Crystallogr. *60*, 1355–1363.

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.

Shen, A. (2010). Allosteric regulation of protease activity by small molecules. Mol. Biosyst. *6*, 1431–1443.

Shen, A., Lupardus, P.J., Albrow, V.E., Guzzetta, A., Powers, J.C., Garcia, K.C., and Bogyo, M. (2009). Mechanistic and structural insights into the proteolytic activation of Vibrio cholerae MARTX toxin. Nat. Chem. Biol. *5*, 469–478.

Thornberry, N.A., Peterson, E.P., Zhao, J.J., Howard, A.D., Griffin, P.R., and Chapman, K.T. (1994). Inactivation of interleukin-1 beta converting enzyme by peptide (acyloxy)methyl ketones. Biochemistry *33*, 3934–3940.

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.