Structural organization of the functional domains of *Clostridium difficile* toxins A and B

Rory N. Pruitt\(^{a}\), Melissa G. Chambers\(^{b}\), Kenneth K.-S. Ng\(^{c}\), Melanie D. Ohl\(^{d,1}\), and D. Borden Lacy\(^{c,d,1}\)

Departments of \(^{a}\)Microbiology and Immunology, \(^{b}\)Cell and Developmental Biology, and \(^{c}\)Biochemistry, Vanderbilt University School of Medicine, Nashville, TN 37232; and \(^{d}\)Department of Biological Sciences and the Alberta Ingenuity Centre for Carbohydrate Science, University of Calgary, Calgary, AB, Canada T2N 1N4

Edited by Stephen C. Harrison, Harvard Medical School, Boston, MA, and approved May 18, 2010 (received for review February 24, 2010)

*Clostridium difficile* toxins A and B are members of an important class of virulence factors known as large clostridial toxins (LCTs). Toxin action involves four major steps: receptor-mediated endocytosis, translocation of a catalytic glucosyltransferase domain across the membrane, release of the enzymatic moiety by autoproteolytic processing, and a glucosyltransferase-dependent inactivation of Rho family proteins. We have imaged toxin A (TcdA) and toxin B (TcdB) holotoxins by negative stain electron microscopy to show that these molecules are similar in structure. We then determined a 3D structure for TcdA and mapped the organization of its functional domains. The molecule has a “pincher-like” head corresponding to the delivery domain and two tails, long and short, correspondingly to the receptor-binding and glucosyltransferase domains, respectively. A second structure, obtained at the acidic pH of an endosome, reveals a significant structural change in the delivery and glucosyltransferase domains, and thus provides a framework for understanding the molecular mechanism of LCT cellular intoxication.

C. difficile is the primary causative agent of hospital-acquired diarrhea and pseudomembranous colitis (1, 2). Virulence is associated with the activity of two large exotoxins, TcdA and TcdB (308 and 270 kDa, respectively) (1–3). The proteins are homologous glucosyltransferases that inactivate small GTPases of the Rho/Rac family. The resulting disruption in signaling causes a loss of cell–cell junctions, dysregulation of the actin cytoskeleton, and apoptosis (4, 5).

The action of TcdA and TcdB on mammalian target cells depends on a multistep mechanism of receptor-mediated endocytosis, membrane translocation, autoproteolytic processing, and monoglucosylation. Many of these functional activities have been ascribed to discrete regions within the primary sequence, suggesting that the toxins will adopt multimodal 3D structures (Fig. 1A). The toxins first bind the surface of the target cell through a highly repetitive C-terminal domain (6, 7) and are internalized by endocytosis (8, 9). The low pH of the endosome is proposed to induce structural changes that lead to pore formation and translocation of the N terminus across the membrane (9–11). The central regions of these toxins have been dubbed “delivery” domains on the basis of the presence of a hydrophobic sequence that could adopt a transmembrane structure during pore formation (12). The N-terminal glucosyltransferase domains are translocated and released into the host cell cytosol, where they can target small GTPases such as RhoA, Rac1, and Cdc42 (13, 14). Release is triggered by an autoproteolytic processing event in which eukaryotic inositol hexakisphosphate (InsP\(_6\)) binds the domain adjacent to the monoglucosyltransferase domain and activates an intramolecular cleavage reaction (15, 16).

TcdA and TcdB belong to a family of large clostridial toxins (LCTs) that also includes the hemorrhagic and lethal toxins of *C. sordelli* and the α-toxin from *C. novyi*. LCTs are important virulence factors, but with the exception of an analysis of TcdB by small angle x-ray scattering (SAXS) (17), the holotoxin structures of these molecules have not been characterized. We have imaged TcdA and TcdB using EM and show that they share many similar structural features. We have determined a 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 (6, 18), the TcdA autoprotease domain (19), and the TcdB glucosyltransferase domain (20) 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. Because 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.

Results

Visualization of TcdA and TcdB by Negative Stain EM. TcdA was purified from *C. difficile* culture supernatant (Fig. 1B) and shown to be active in a cell-rounding assay (Fig. S1B). Toxin was adsorbed to carbon-coated glow-discharged grids and stained with uranyl formate. Negative stain EM revealed homogeneous particles with a nonsymmetrical shape and an elongated “tail” (Fig. 1C). 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 untitled specimens were classified into 12 class averages. These class averages revealed that although 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” (Fig. S2A). 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 (Fig. S2A, marked with a “*”) and used them as references for another cycle of multireference alignment (Fig. S2B). The largest of the three resulting classes (4,956 particles; Fig. 1D and Fig. S2B, marked with a “*”) showed a TcdA particle with many clear structural features (Fig. 1D). The TcdB holotoxin was also purified (Fig. 1E) and shown to be active in a cell-rounding assay (Fig. S1C). Although negative stain EM revealed a more heterogeneous field of particles than seen with TcdA (compare Fig. 1C 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 (Fig. S2D, classes 1 and 5). To improve the alignment we selected four classes as ref-

Author contributions: R.N.P. and D.B.L. designed research; R.N.P., M.G.C., and M.D.O. performed research; K.K.-S.N. contributed new reagents/analytic tools; R.N.P., M.G.C., and D.B.L. analyzed data; and R.N.P., M.D.O., and D.B.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence may be addressed. E-mail: melanie.ohl@vanderbilt.edu or borden.lacy@vanderbilt.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1002199107/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1002199107
TcdA (Fig. 1D). 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 (21) and is presented in Fig. 2 and Movie S1. The Fourier shell correlation (FSC) curve calculated from our final density map suggests a resolution of 42 Å on the basis of the FSC = 0.5 criterion (22) (Fig. S2C). The face view of the 3D density map shows structural features very similar to those seen in the projection average (compare upper left panel in Fig. 2 with Fig. 1D), suggesting that the 3D reconstruction was successful. The 3D density map of TcdA adopts an elongated, asymmetric structure that is ≈ 310 Å × ≈ 150 Å × 90 Å in dimension. The structure contains three prominent features: a “head” domain, a long kinked “tail,” and a short inner “tail”. The head, ≈ 90 Å × 90 Å × 60 Å in size, seems 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.

**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 (18). 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 (6, 18) (Fig. 3A). Several features of this predicted model are seen in the long tail observed by EM (Fig. 3A and B). To confirm that the long tail is indeed the binding domain, we expressed the TcdA binding domain (amino acids 1832–2710) in *Escherichia coli*, purified it, and subjected it to negative stain EM (Fig. 3C and 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 (18) (Fig. 3A–C). Thus, the long kinked tail domain found in our structure represents the C-terminal binding domain of TcdA.

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 (Fig. 4A), but classification of 1,523 particles into five class averages resulted in two classes of 139 and 314 particles (Fig. 4B and C) that recapitulate the size, shape, and pincher-like features of the holotoxin head domain (Fig. 4D). An alternative approach, wherein we labeled the toxin

![Fig. 1.](https://www.pnas.org/ cgi/doi/10.1073/pnas.1002199107)  
**Fig. 1.** Purification and characterization of TcdA and TcdB. (A) The proposed domain organization for TcdA and TcdB. The numbers are for amino acids that mark domain boundaries. (B) SDS/PAGE of purified TcdA, visualized by Coomassie staining. (C) Typical electron micrograph showing TcdA particles in negative stain. A few of the particles are circled in black. (Scale bar, 500 Å.) (D) Representative class average of TcdA particles (4,956) selected from images of un tilted 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.

![Fig. 2.](https://www.pnas.org/ cgi/doi/10.1073/pnas.1002199107)  
**Fig. 2.** 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, 5 nm.)
with a monoclonal antibody specific for the delivery domain, is presented as part of the supplemental data for this article (Fig. S3).

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 (Fig. 4E). The most populated class average from this analysis (405 particles out of 1,002 segregated into the image shown in Fig. 4F) 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 (Fig. S4). The glucosyltransferase domain was removed by gel-filtration chromatography, and the larger fragment was analyzed by negative stain EM (Fig. 5A).

Upon examining the images of cleaved TcdA (Fig. 5A and B), it is clear that whereas 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.

A model of the binding domain of TcdA was placed into the 3D EM map (Fig. 6). 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 with 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 (Fig. 6). The domain is 90 Å long, consistent with the ~100 Å length of the short tail, and is ~25 Å wide at the extreme N terminus. The C-terminal side of the domain is considerably larger with a diameter of ~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.

**Fig. 3.** Characterization and localization of the TcdA C-terminal binding domain. (A) Predicted model of the TcdA binding domain (18). (B) Average of 4,956 TcdA particles (from Fig. 1D). (C) Image of a single negative stained particle of the TcdA binding domain. (D) Typical electron micrograph area of the recombinitely expressed C-terminal domain of TcdA (1832–2710). A few of the particles are circled in black. (Scale bar, 500 Å.)

**Fig. 4.** Visualization of the TcdA delivery domain by negative stain EM. (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 and 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 after exposure to a low-pH buffer. Side length of panel is 27.2 nm.

**Fig. 5.** Localization of the TcdA N-terminal glucosyltransferase domain by autoproteolysis. Cleavage of TcdA was induced by the addition of InsP6 and DTT. The large fragment containing residues S43–2710 was isolated, applied to a grid, stained with uranyl formate, and visualized by EM. (A) Representative image of cleaved TcdA particles in negative stain. (Scale bar, 500 Å.) (B) 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.

**Fig. 6.** 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 Fig. 7A (compare with Fig. 1C). To examine the structural homogeneity of TcdA in a low-pH state, ~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 (Fig. S5A). 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 Fig. 7 with Fig. 1D). To more carefully address the structural changes in TcdA structure at endosomal pH, additional images of TcdA in a low-pH state were collected at 60° and 0°. A total of 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
low pH. Side length of panel is 57.3 nm. A carbon-coated glow-discharged grid, washed with 50 mM sodium citrate at pH 4.5. A few of the particles are circled in black. (Scale bar, 500 Å.)

**Fig. 6.** Placement of the functional domains of TcdA within EM map. (A and B) 3D reconstruction of TcdA filtered to 25 Å, 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 docked into the density. B shows two views of the toxin enlarged to show the placement of the glucosyltransferase and protease domains.

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 three to five short repeats and one long repeat) as compared with seven in the TcdA sequence.
TcdA (18). Truncating the TcdA structural model after four 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 (Fig. S2 D and E). 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 (17) because ab initio envelope calculations from scattering data can be problematic in flexible systems in which domain orientations differ between conformers (25, 26).

TcdA and TcdB have been proposed to undergo pH-dependent conformational changes to form a pore through which the glucosyltransferase domain is translocated (10). 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 (Fig. 4 E and F and Fig. 8). This might be accomplished through a decoupling of the two lobes of the head, effectively opening the pincher.

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 reveals that these are two regions with distinct homology profiles: residues 801–1831 (D1) and 1401–1831 (D2) (Fig. S7) (27). 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 (8) and contains an Asp-Ser-Gly motif, which may be involved in toxin delivery into the cytosol (16). Further study is needed to dissect the respective roles of these delivery domains in the context of the membrane.

Release of the glucosyltransferase domain into the cell is mediated by the adjacent cysteine protease domain (15). 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 (19). The TcdA autoprotease domain is known to undergo significant rearrangement upon exposure to InsP6 (19). Analysis of cleaved TcdA in negative stain reveals that the particles are much more heterogeneous 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, because 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 nonpore forming state, a model that is further supported by our data showing that 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 (28–30). 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. Despite this mechanistic similarity, anthrax toxin, botulinum neurotoxin, and diphtheria toxin have been noted for their diversity in structure (31).

The unique structural features observed in this study, namely a 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.

**Methods**

Cloning. DNA corresponding to the TcdA receptor-binding domain (amino acids 1832–2710), TcdA delivery domain (amino acids 799–1859), and TcdB holotoxin was amplified and used as template to prepare in vitro transcription-translation reactions.

**Protein Expression and Purification.** Native TcdA was obtained from the supernatant of *C. difficile* strain 10643 grown in dialysis sac culture, as described previously (1). TcdA was purified by multiple rounds of anion-exchange chromatography, followed by gel-filtration chromatography in 150 mM NaCl and 20 mM Tris, pH 8.0. The identity of TcdA was verified by mass spectrometry. TcdB was expressed in *Bacillus megaterium* as described previously (32) except cells were harvested at 4 h after induction. All recombinant single domains were expressed from *E. coli* as reported previously for the autoprotease domain (19), except BL21(DE3)-CodonPlus cells (Strategene) 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 and 50 mM potassium phosphate, pH 7.0, by gel filtration chromatography.

**Autoproteolysis of TcdA.** Purified TcdA (100 μg) in 150 mM NaCl and 20 mM Tris, pH 8.0, was mixed with 100 mM InsP6 and 100 mM DTT and incubated at 37 °C for 3 h. Some of the protein precipitated during this process. Protein that remained soluble was run on a 2000 g gel-filtration column in 150 mM NaCl and 20 mM Tris, pH 8.0.

**Specimen Preparation and EM.** Uranyl formate (0.7% wt/vol) was used for conventional negative staining as previously described (33). 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 LaB6 filament and operated at an acceleration voltage of 120 kV. Images were taken under low-dose conditions at a magnification of ×67,000 using a defocus value of −1.5 μm. Images were recorded on DITABIS digital imaging plates. The plates were scanned on a DITABIS micrometer scanner, 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 2,048 × 2,048-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 ×100,000 and a defocus value of −1.5 μm. Images were collected using a Gatan 4K × 4K 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 (34) and windowed with 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 (35).

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 × 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 (Fig. S2A). From the class averages, three representative projections were chosen and used as references for another cycle of multireference alignment (Fig. S2A, 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 (Fig. S5A, 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) (Figs. S2A and S6A) 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 (Fig. S2A). From the class averages, three representative projections were chosen and used as references for another cycle of multireference alignment (Fig. S2A, 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 (Fig. S5A, 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) (Figs. S2A and S6A) 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 (Fig. S2A). From the class averages, three representative projections were chosen and used as references for another cycle of multireference alignment (Fig. S2A, 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 (Fig. S5A, 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) (Figs. S2A and S6A) 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 (Fig. S2A). From the class averages, three representative projections were chosen and used as references for another cycle of multireference alignment (Fig. S2A, 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 (Fig. S5A, 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) (Figs. S2A and S6A) 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 (Fig. S2A). From the class averages, three representative projections were chosen and used as references for another cycle of multireference alignment (Fig. S2A, 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 (Fig. S5A, marked with a “*”).
Expression vectors for the production of the TcdA protease domain (amino acids 543–795) and delivery subdomains (amino acids 799–1460 and 771–1460) were constructed to generate recombinant proteins from *Escherichia coli*. Briefly, all clones were derived by amplification of the appropriate nucleotide sequence from *Clostridium difficile* strain 10643 genomic DNA. The DNA was cloned into a modified pET27 vector such that the resulting protein contains an N-terminal His$_{10}$ tag separated from the protein by a 3C protease cleavage site. The proteins were purified following the same protocol as for the other recombinantly expressed proteins. The nucleotide sequence encoding TcdA amino acids 1–542 was cloned into the *Bacillus megaterium* expression vector pC-His1622 (MoBiTec, BMEG20) using the restriction sites BsrGI and SphI. The plasmid was transformed into *B. megaterium*, and the protein was expressed following the manufacturer’s instructions. After French Press lysis, the lysate was centrifuged to pellet debris. Protein was purified from the supernatant by Ni-NTA affinity chromatography and gel filtration chromatography.

**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 TcdA, or 50 μg native holotoxin at 5 μg/mL were coated onto 96-well ELISA plates at 4 °C 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 Laboratories) was added to the plates in PBST and incubated 1 h. The ELISAs were developed by the addition of ABTS substrate solution [1 mM ABTS (2,2’ azino-di-(3-ethylbenzthiozoline sulfonic acid)) (Sigma) in 70 mM citrate-phosphate buffer, pH 4.2, with 0.03% H$_2$O$_2$] for 30 min, and absorbance was measured at 410 nm.

For antibody labeling, the proteins were mixed in 1:2 mass ratios and incubated at room temperature for 3 h before grid preparation. Grids were prepared as described in the main text. Particle images of TcdA bound to 15A4 were selected with Boxer (1) and windowed with SPIDER and the associated display program WEB (2).

**TcdA Glucosyltransferase Domain Solubility Assay.** Purified glucosyltransferase domain was exchanged into 150 mM NaCl and 10 mM Tris, pH 8, by gel filtration chromatography and concentrated to ≈0.38 mg/mL. Protein (18 μL) was mixed with 2 μL 1M sodium citrate at the indicated pH. After a 30-min incubation at room temperature, the samples were centrifuged at 14,000 × g. The absorbance of the supernatant was measured at 280 nM.


---

**Fig. S1.** Confirmation of TcdA and TcdB cytopathic activity. CHO cells were plated in a 96-well cell culture dish at 1,000 cells per well in DMEM with 10% FCS. Wells were treated with (A) buffer, (B) 0.2 μg TcdA, or (C) 0.2 μg TcdB for 4 h and visualized by light microscopy.
Fig. S2. EM analysis of TcdA and TcdB. (A) Average of 7,396 TcdA particles classified into 12 class averages. Boxes marked with "*" indicate averages used as references for B. (B) Multireference alignment of 7,396 particles into three 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) Fourier shell correlation (FSC) curve for neutral pH structure of TcdA. (D) Averages of 2,133 TcdB particles classified into 10 class averages. Boxes marked with "*" indicate averages used as references for E. (E) Multireference alignment of 2,133 particles into four classes. Each box has a side length of 51.1 nm.
Fig. S3. 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 (amino acids 1–542), autoprotease (543–795), delivery (799–1460), and binding (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) 2D view of 15A4 antibody position (marked with * *) on the 3D model of TcdA. (Scale bar, 5 nm.)

Fig. S4. Cleaved TcdA. Coomassie stained SDS/PAGE of native TcdA and the larger fragment of TcdA after autoproteolytic cleavage and purification by size exclusion chromatography.
**Fig. S5.** 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 4,014 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 four 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.

**Fig. S6.** pH dependence on the solubility of TcdA glucosyltransferase domain. (A) TcdA was incubated at either pH 7 or pH 4.5 for 5 min. 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.
Fig. S7. 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.

**Movie S1.** 3D random conical tilt reconstruction of TcdA in negative stain.

**Movie S2.** 3D random conical tilt reconstruction of TcdA at low pH in negative stain.