Structural Insights into the Transition of *Clostridioides difficile* Binary Toxin from Prepore to Pore

David M. Anderson[1,2,*], **Michael J. Sheedlo**[1,2,*], **Jaime L. Jensen**[1], **D. Borden Lacy**[1,2,†]

[1]Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN 37232

[2]The Veterans Affairs Tennessee Valley Healthcare System, Nashville, TN 37232

Clostridioides (formerly *Clostridium*) *difficile* is a gram-positive, spore-forming anaerobe and a leading cause of hospital-acquired infection and gastroenteritis-associated death in U.S. hospitals[1]. The disease state is most often preceded by disruption of the host microbiome in response to antibiotic treatment and is characterized by mild to severe diarrhea. *C. difficile* infection (CDI) is dependent on the secretion of one or more AB-type toxins: toxin A (TcdA), toxin B (TcdB), and the *C. difficile* transferase toxin (CDT)[2]. While TcdA and TcdB are considered the primary virulence factors, recent studies suggest that CDT increases the severity of CDI in some of the most problematic clinical strains[3]. To better understand how CDT functions, we used cryo-electron microscopy (cryo-EM) to define the structure of the cell-binding component of CDT (CDTb). We obtained structures of several oligomeric forms that highlight the conformational changes that permit conversion from a prepore to a β-barrel pore. The structural analysis also reveals a glycan binding domain and residues involved in binding the host cell-receptor, lipolysis-stimulated lipoprotein receptor (LSR). Together, these results provide a framework to understand how CDT functions at the host cell interface.

CDT belongs to the Iota family of binary toxins and consists of two proteins: an ADP-ribosyltransferase (CDTa) and a cell binding, pore-forming, and CDTa translocating protein (CDTb)[4]. CDTb engages host cells through LSR[5] and undergoes proteolytic cleavage and oligomerization to facilitate CDTa binding. The complex is internalized by endocytosis, and endosome acidification leads to CDTb pore formation and CDTa translocation. Inside the cell, CDTa catalyzes ADP-ribosylation of G-actin which leads to disruption of the cytoskeleton, formation of microtubule protrusions, and increased *C. difficile* adherence

(Extended Data Figure 1)[6–8]. Some of this mechanistic understanding has been inferred from comparisons to *Bacillus anthracis* protective antigen (PA), a tripartite binary toxin which mediates the delivery of two different enzymes, edema factor (EF) or lethal factor (LF)[9]. PA has a four-domain structure with homology to B-subunits of the Iota toxin family and forms heptameric[10,11] and octameric[12] oligomers that convert from a prepore to a β-barrel pore at low pH[13]. This leads to unfolding and translocation of EF/LF across the endosomal membrane[14,15]. It has long been hypothesized that these mechanisms would also be true for CDTb. However, differences in receptor specificity, coupled with the necessity of CDTb to translocate different cargo suggests some mechanistic differences must exist between the two.

Oligomerization of CDTb was initiated by *in vitro* trypsinization at neutral pH (20 mM HEPES, pH 8.0), and the resulting particles were imaged by cryo-EM (Supplementary Data Table 1). Initial 2D classification revealed a mixture of smaller non-oligomerized or partially-oligomerized particles and larger heptameric particles (Extended Data Figure 2a, b, c). Views of the smaller particles were restricted and led to only low-resolution reconstructions. Classification of the larger particles revealed a dimer of heptamers that segregated into populations of either 'short' particles (245 Å) or 'long' particles (270 Å) (Extended Data Figure 2c). Three-dimensional classification and refinement of these data led to reconstructions at 3.7 Å for the short form and 3.9 Å for the long form (Fig. 1a, e, Extended Data Figure 3, 4a–c and e–h). In the short form, the local resolution extended to 3.4 Å, allowing us to discern and construct a model in which we built one heptamer in the pore state (Fig. 1b) and one heptamer in the non-inserted prepore state (Fig. 1c). In the long form, we observed two nearly identical heptamers where only a portion of the β-barrel had formed (Fig. 1e–g). Within both models, we observed five domains denoted as D1-D3, D3', and D4 (Fig. 1d, h). Domains D1-D3 are arranged in a manner similar to PA with RMS deviations of 2.0 Å2 and 1.8 Å2, for the prepore and pore states, respectively (Extended Data Figure 5a, b). Phenylalanines involved in PA pore formation or pore gating are structurally conserved in CDTb, and mutation of either mitigates the toxicity of CDT in a monolayer disruption assay (Extended Data Figure 5e, f). While the overall folds for D1-D3 are similar, the electrostatic surfaces of CDTb and PA differ at the top (where EF/LF bind PA) and on the inside and outside of the β-barrel (Extended Data Figure 5c, d).

The D3' domain represents a previously unidentified discrete domain in CDTb which is composed of two β-sheets arranged in a β-roll fold (Fig. 1d, Extended Data Figure 5a, and Fig. 2a, inset). Both a sequence search as well as a DALI (distance matrix alignment) query suggested D3' might encode for a galactose binding domain. Glycan array screening using purified D3' revealed a capacity to bind L-fucose, D-fucose, L-rhamnose, chitobiose, chitotriose and lacto-N-tetraose (Fig. 2a). Though D3' is not observed in PA, it is likely present in similar Iota family toxins such as CSTb (*C. spiroforme*) and Ib (*C. perfringens*) which share 68% and 72% identity, respectively. D3' is weakly analogous (16% identity) to the 590–721 residue segment of C2 toxin from *C. botulinum* (Extended Data Figure 1, inset). While carbohydrate binding has not been demonstrated previously in the Iota toxins, the C2 toxin has been reported to engage the cell surface using carbohydrate receptors[16]. Sequence alignments in the context of this structural framework suggest that *C. botulinum* C2 toxin may engage carbohydrate receptors using a domain like the CDTb D3' domain

rather than a protein receptor binding domain (D4). In CDT, the presence of excess fucose or chitobiose did not reduce the monolayer disruption activity of the toxin (Fig. 2b). In fact, the presence of excess chitobiose led to a slight increase in activity. Further studies are needed to evaluate the mechanism and physiological relevance of this observation.

The toxicity of CDT[17], CST[18] and Ib[5] is mediated by interactions between the D4 domain and the cell surface receptor, LSR. The monolayer disruption assay was developed to monitor CDT activity and define the LSR binding site on D4 (Extended Data Figure 6a, c). In this assay, the Caco-2 monolayers remain intact when adding CDTa or CDTb alone. Additionally, the monolayers remain intact when CDTa is added to the CDTb double heptamer, suggesting that the double heptamer is non-functional and unlikely to be relevant in the context of infection. When CDTa is added with monomeric CDTb, thus allowing oligomerization to occur on the cell surface, cytopathic effects are observed[5,19,20]. This assay was used to quantify the effect of D4 surface residue mutations (Fig. 3a, Extended Data Figure 6a, b). Mutations in non-conserved residues F774, I852 or T854 resulted in decreased activity, and these mutants were chosen for further analysis by flow cytometry. In these experiments F774D, I852E and T854E exhibited a decreased binding capacity, while mutation of a proximal residue (Y812) did not impact cellular binding (Fig. 3b, Extended Data Figure 7). Efforts to test the impact of the CDTb mutations on LSR binding affinity using purified proteins were hampered by the poor expression of the LSR ectodomain and the hydrophobic nature of both LSR and the CDTb heptamer. Nevertheless, a preliminary experiment using bio-layer interferometry is consistent with the conclusion that the F774D mutant is impaired in its capacity to bind LSR (Extended Data Figure 6c). The CDTb F774 residue is also involved in the D4-D4 *trans* interaction implicated in the formation of double heptamers *in vitro* (Extended Data Figure 9c). In summary, while mutation of many conserved CDTb D4 surface residues did not alter the cytopathic effects of the toxin, mutations on the bottom face of the D4 domain decreased the ability of the toxin to disrupt the cell monolayer, bind cells, and in the case of F774D, bind LSR *in vitro*. Together, these data suggest LSR binding involves the base of the D4 domain (Fig. 3c).

In an effort to obtain a structure of LSR bound to CDTb, the ectodomain of LSR was purified and incubated with CDTb oligomer prior to concentration. Cryo-EM analysis of this sample resulted in a 4.2 Å structure (Extended Data Figures 2d, 3, 4d, i). While the LSR and D4 domains were not visible, the structure represents a 'pre-insertion state'; the D2 domains have moved ~15 Å inward, adopting a conformation similar to their position in the pore state (Fig. 4), and the pore forming loop (PFL, residues 312–383) that will form the β-barrel pore has been dislodged from its position between D3 and D3' (Extended Data Figure 8a). This 'pre-insertion' state can also be compared to the state containing the partial β-barrel that was reconstructed from the high molecular weight 'long' form (Fig. 1e–h). In this conformation, we observe residues 312–327 and 367–383 of the PFL folding into a partial β-barrel structure (Fig. 4, Extended Data Figure 8a). The isolation of such a state suggests that β-barrel formation begins at the base of the oligomer and continues down toward the cell surface through a 'zipper' like mechanism which ends by burying a hydrophobic surface into the cellular membrane (Fig. 4, Extended Data Figure 8a).

In the prepore, partial β-barrel, and pore structures, the D4 domain forms a static ring about the center of the molecule. It is possible that the static D4 ring is stabilized by the dimer of heptamers and that it does not form under physiological conditions. On the other hand, it is notable that while the D4-D4 interfaces are nearly identical between the three structures (Extended Data Figure 9), the position of the ring relative to the rest of the molecule is not (Extended Data Figure 8b). These changes in position present the intriguing possibility that the D4 ring may play a role in regulating pore formation. First, we hypothesize that if the D4 ring is present in the prepore, it will need to disassociate to permit the conformational change in the PFL. This possibility is supported by our observation that the D4 ring cannot be detected in the pre-insertion state, a sample which contained the LSR ectodomain (Fig. 4). Second, if the D4 ring disassociates and then reforms after rearrangement of the PFL, the ring formation could contribute to insertion of the PFL into the membrane. When comparing the position of D4 relative to D2 in the partial β-barrel and pore states, we noted a rotation about the pore into an apparently domain-swapped arrangement (Extended Data Figure 8b, c). In the context of a physiological membrane, this twisting motion could promote PFL insertion into the membrane. We note that one consequence of the domain-swap is that it pulls the ring away from the membrane and closer to the main body of the toxin, leaving ~32 Å between the base of the D4 and the cell surface (as defined by the hydrophobic surface of the CDTb β-barrel, Fig. 4). If LSR remains bound after pore formation, we predict this space would be sufficient to accommodate the LSR ectodomain. Further studies in the context of cellular membranes will be needed to test these ideas. In addition, we note that all structural states observed in this study were obtained at pH 8.0. Further work will be needed to evaluate whether acidic pH is, in fact, required for pore formation in the context of cells or if receptor binding can serve as a trigger. It is possible that acidic pH is needed for unfolding and translocation of the CDTa cargo but is not essential for CDTb pore formation.

In summary, these studies provide four distinct molecular structures of the CDTb heptamer. They reveal a D3' carbohydrate binding domain not present in PA, followed by a long linker that allows the CDTb D4 domain to adopt multiple conformations. We identify residues in the LSR-binding interface, and detail a set of four structures that outline a process for the conversion of CDTb from a prepore to a fully extended β-barrel. We anticipate these structures will impart a framework for considering mechanisms of neutralization and inhibition which could inform approaches for the treatment and prevention of CDI.

Methods

Sequence analysis

To determine the identity of the Iota toxin family all toxin sequences were acquired from UniProt (CSTb - O06498_9FIRM, Ib - F7J0A3_CLOPF, PA - PAG_BACAN). All sequences were aligned with the sequence of CDTb (CDTb - Q9KH41_CLODI) using ClustalΩ[21]. The degree of identity was obtained from the protein identity calculation server Ident and Sim within the Sequence Identification Suite (SMS) [22].

Plasmid construction, generation of point mutants and protein purification.

Genes for CDTa and CDTb were amplified from the genomic DNA of *C. difficile* strain R20291 (gifted from Dr. Eric Skaar, Vanderbilt University). CDTa was cloned into plasmid pLM302, a modified pET28 vector with an N-terminal MBP tag and PreScission cleavage site (Vanderbilt University Center for Structural Biology). The resulting plasmid (pBL926) was transformed into BL21 Star, grown at 220 RPM in LB medium until OD600 = 0.5 and induced with IPTG for 16 h at 16 °C. Cells were harvested at 6800 × g for 15 min, 10 °C and stored at −80 °C. Pellets were thawed in lysis buffer (20 mM Tris-HCl pH 7.4, 200 mM NaCl, 1 mM EDTA) with DNase I and protease inhibitors, and lysed with an Avestin Emulsiflex high-pressure homogenizer. Supernatants were collected after a 20 min, 10° C, 30,000 × g centrifugation and incubated with amylose resin for 1 h at 4°C. Resin was collected by gravity flow, washed once with 10 column volumes (CVs) lysis buffer, and protein was eluted with 10 CVs of lysis buffer containing 10 mM maltose. The eluate fraction was loaded into dialysis tubing and incubated for 16 h at 4 °C with PreScission Protease in 20 mM Tris- HCl pH 7.4, 200 mM NaCl. The sample was then applied to an SP Sepharose column (GE) equilibrated in 20 mM HEPES pH 7.0, 100 mM NaCl and eluted with a 100–500 mM NaCl gradient. Peak fractions were pooled and applied to an S200 gel filtration column (GE) in 20 mM HEPES pH 8.0, 100 mM NaCl (size-exclusion chromatography (SEC) buffer). Peak fractions were collected, concentrated, sterile filtered, and flash-cooled in liquid nitrogen.

CDTb (residues 43–876) was cloned into pET28a (pBL870) and transformed into BL21 (DE3) RIL *E. coli*. All point mutations were made by the QuikChange method using primers listed in Supplementary Data Table 2 and confirmed by Sanger sequencing. Cultures were expressed and harvested similarly to CDTa. Lysates were produced in 20 mM Tris-HCl pH 8.0, 500 mM NaCl buffer (with DNase I and protease inhibitors) and purified via metal affinity chromatography. The elution fractions were concentrated via centrifugal filters, diluted 10-fold into 50 mM Bis-Tris pH 6.5, and applied to a Q Sepharose column run with a 0–400 mM NaCl gradient. Peak fractions were pooled and applied to an S200 column in SEC buffer. Proteolytic activation was accomplished by incubation with bovine trypsin for 45 min in 20 mM HEPES pH 8.0, 100 mM NaCl at 37°C with a 1:5 w/w ratio of trypsin to toxin. PMSF was added at a final concentration of 1 mM to stop the reaction, and the sample was applied to an S200 column in SEC buffer. Peak fractions were collected, concentrated, sterile filtered, and flash-cooled in liquid nitrogen. Proteins utilized for flow cytometry were labeled for 1 h at room temperature with Alexa Fluor 647 NHS ester (ThermoFisher) and buffer exchanged into SEC buffer via PD-10 columns (GE Healthcare) prior to flash-cooling in liquid nitrogen. Typical labeling ratios were 0.5–1 mol dye per toxin as assessed by absorption using the extinction coefficients of 88,190 cm-1M-1 for toxin and 270,000 cm-1M-1 for Alexa 647.

A construct of the D3' domain was cloned into pAT109 (Vanderbilt University Center for Structural Biology) using BamHI and KpnI to generate a construct containing an N-terminal GST tag followed by a PreScission Protease site. This construct (pBL915) was transformed into BL21 (DE3) RIL which were cultured at 37°C to OD600 = 0.4–0.6 before protein expression was induced with a final concentration of 250 μM IPTG. The temperature was

then reduced to 18 °C, and the cultures were incubated an additional 16–18 h before harvesting by centrifugation. The resulting cell pellet was suspended in 20 mM HEPES pH 8.0, 100 mM NaCl, 10% glycerol and loaded onto glutathione Sepharose resin. The column was washed with 5 CV resuspension buffer and eluted with the same buffer supplemented with 10 mM reduced glutathione. This sample was then concentrated, and the GST tag removed by incubation with PreScission Protease for 16 h concurrent with dialysis into resuspension buffer to remove glutathione. The sample was passed back over glutathione Sepharose resin to remove the cleaved GST before being further purified by size exclusion chromatography in SEC buffer.

A construct encoding LSR (residues 1–259, pBL839) with a N-terminal human serum albumin secretion signal and C- terminal histidine tag was synthesized into pcDNA3.4 (Life Technologies). This vector was mixed with an equal volume of ExpiFectamine (ThermoFisher) diluted in OptiPro media and applied to a suspension of ExpiCHO cells to yield a final concentration of 1 μg DNA/ml culture. The culture was then incubated overnight and ExpiCHO transfection optimizer and feed added (ThermoFisher). Protein expression was carried out over the course of 8 days at which point the cells were removed by centrifugation and the supernatant applied to chelator resin charged with nickel. The resin was then washed with 20 mM HEPES pH 8.0, 100 mM NaCl, 2 mM imidazole to remove impurities. LSR was eluted with 150 mM imidazole and concentrated to a final volume of 1 ml. PNGase F was used to deglycosylate the LSR and the resulting sample purified by size exclusion chromatography.

To form the pre-insertion complex, CDTb was trypsinized to form oligomers as described above. The sample was then purified by size exclusion chromatography, and the sample containing only the pre-insertion state was immediately mixed with LSR at a 1:1.5 molar ratio (CDTb:LSR) and the sample was concentrated and screened by negative stain EM.

Cryo-EM sample preparation and data collection

Trypsin-activated CDTb was applied to glow discharged Quantifoil R2/1 grids using a Vitrobot Mark IV at a final concentration of 177 nM. The grids were plunged into liquid ethane and stored under liquid nitrogen. Data were collected over the course of two sessions using a FEI Titan Krios operating at 300 keV and outfitted with a Gatan K2 direct electron detector and Bioquantum energy filter set to a slit width of 20 eV (Washington University Center for Cellular Imaging; see Supplementary Data Table 1). Data were collected with a pixel size of 1.45 Å/pixel and a total dose of either 111.5 or 109.4 e-/Å2. Movies were collected in counted mode with a frame rate of 300 ms/frame for a total of 32 frames.

The CDTb-LSR complex sample was applied to glow discharged Quantifoil R1.2/1.3 grids using a Vitrobot Mark IV at a final concentration of 800 nM and plunged into liquid ethane cooled by liquid nitrogen. Data were collected on a Titan Krios as described above with a pixel size of 1.13 Å/pixel and a total dose of 73.4 e-/Å2. Movies were collected with a frame rate of 200 ms/frame for a total of 32 frames.

Cryo-EM data processing

All CDTb micrograph movies were aligned using MotionCor2 with dose weighting applied [23] and then CTF corrected with CTFFind4 [24,25]. The monomer/dimer classes were generated by first picking a small subset of particles by hand that included both forms. These particles were then aligned and classified in Relion 3.0 [26]. The particles which represented either the monomer or the dimer were separately selected from this initial alignment and further classified. Alignment in 3D was attempted, but due to problems with preferred orientation, a high-resolution map could not be reconstructed. A small subset of the oligomeric particles were picked by hand from the same dataset and aligned and classified in Relion 2.1 [27]. The resulting classes were used as templates to pick particles from the remaining micrographs from this dataset.

These particles were then classified in 2D, revealing stacked heptamers in two distinct conformations. Though differences could be discerned in some classes, it was impossible to differentiate between the two forms in all 2D classes obtained.

Therefore, all particles were classified in 3D, and the highest resolution classes were selected. Particles from a second dataset were then added after an initial round of 2D classification. The resulting particles were purified over two rounds of 3D classification and refined with no symmetry applied to yield a final global resolution of 4.5 Å for the short form particles. These particles were further refined with C7 symmetry applied to a final resolution of 3.7 Å. The long form particles were further classified in 3D and then refined with C7 symmetry to a resolution of 6.3 Å. Masks were created in Relion 3.0 pertaining to the 'body' of the toxin on either side of the particle as well as the D4 rings. These masks were applied in Relion MultiBody refinement and refined to a final resolution of 3.9 Å. Portions of the resulting maps were combined in Phenix[28] to obtain a complete map of the long form of CDTb which was used to refine the entire structure. Processing of raw movies of the preinsertion-LSR complex was carried out following a virtually identical protocol as described above. Particles were again picked by hand in Relion 3.0 and an initial 2D classification carried out to generate templates. The data were autopicked in Relion 3.0 and the dataset cleaned with 2D classification. The resulting particles were classified in 3D and refined with applied C7 symmetry. Particles from the highest resolution class were selected and underwent one more round of 3D classification and then 3D refinement with C7 symmetry applied. All cryo-EM data processing and reconstruction software were accessed as a part of the SBGrid consortium[29].

Model building and refinement

Models of single protomers of the CDTb oligomeric state and pore were constructed in COOT[30] using the short oligomer map[30]. These models underwent iterative rounds of refinement in COOT as well as Phenix real space refine (see Supplementary Data Table 1) with secondary structure and Ramachandran restraints applied over the course of refinement[28]. The heptameric structure was generated using ApplyNCS and refined in Phenix to generate models of both the oligomeric and pore states of CDTb. To build the partial β-barrel states, the CDTb pore was first truncated and docked into the map in UCSF Chimera[31]. The model was then refined in COOT and Phenix as described above. Similarly,

the pre-insertion state-LSR complex model was constructed by first docking the CDTb oligomer into the map followed by refinement in COOT and Phenix as described above. All software used in model building and refinement were accessed as a part of the SBGrid consortium[29].

Flow cytometry binding assay

Caco-2 cells were obtained from ATCC (HTB-37) and were not further authenticated or tested for mycoplasma contamination. They were grown in DMEM with 10% FBS. Twenty-four h prior to the assay, cells were split into 10 cm dishes. Cells were washed once in PBS before incubation for 1 h at room temperature with CellStripper dissociation solution (Corning). Cells were then scraped off each plate, pooled, harvested by centrifugation (1000 × g for 3 min), suspended in room temperature RPMI media without FBS and phenol red, and passed through a 70 μM filter. Cells were then enumerated via a digital cell counter (Corning) and adjusted to 5×10^5 cells/ml. Aliquots of cells were removed for compensation controls of no stain and Alexa-647-labeled toxin-only. Viability dye (7AAD, ThermoFisher) was then added to the remaining pool of cells at 0.25 μg/million cells and an aliquot was removed for the 7AAD-only control. Cells were allocated to tubes containing various amounts of toxin to incubate at room temperature for 15–30 minutes prior to data collection on a 3 laser BD LSR II cytometer. All data analysis was performed using Flowing Software (Perttu Terho, Cell Imaging Core, Turku Centre for Biotechnology), and gating for display images utilized FlowJo_V10 (FlowJo, LLC).

Monolayer disruption assay

Caco-2 cells were seeded into 48 well plates at 40,000 cells per well and allowed to grow for 48 h. CDTa and CDTb were diluted into fresh media to a concentration of 10 nM each and applied to the cells for imaging on a Cytation5 plate reader, with image acquisition at 20X magnification every 30 min for 14 h. The glycan competition assay was conducted under identical conditions in the presence of either 10 mM L-fucose (Sigma) and 10 mM D-fucose (Sigma) combined together, or in the presence of 10 mM chitobiose (Carbosynth). All images were assembled into a stack using ImageJ software, cropped and background corrected using the built-in plugin BaSiC [32]. Individual stacks were saved into a separate folder corresponding to each well. The last image from each stack was compiled into a separate stack for training of a pixel classifier using the built-in Weka segmentation plugin [33]. The free-hand drawing tool was used to select areas corresponding to cells or background to train the classifier before it was used to batch process stacks corresponding to each well in the assay using Fiji macros (Supplemental Software File, Script #1). Each stack was classified into a stack of binary images to quantify pixel coverage of cells over time (Supplemental Software File, Script #2). Linear regression analysis and one-way ANOVA followed by Dunnett's multiple comparisons test were performed using GraphPad Prism 6.0.

Glycan binding array

Purified D3' domain was labelled with NHS-Alexa 647 in 20 mM HEPES pH 8.0, 100 mM NaCl with a 4:1 molar excess of dye. Unreacted dye was removed by passing the sample over a PD-10 column, and the resulting protein was concentrated. The sample was analyzed on a glycan 100 binding array provided by Zbiotech (Aurora, CO) at 50 μg/ml.

Mean fluorescence signal from each of the glycan array spot replicates was plotted and compared by one-way ANOVA with Dunnett's multiple comparisons test.

Bio-layer interferometry assay

Recombinantly purified LSR was biotinylated with sulfo-BMCC biotin (ThermoFisher) at a 1:2 protein:biotin ratio for 3 h at room temperature. Streptavidin biosensors were loaded with biotinylated LSR (120 nM) for 300 sec, followed by a 30 sec wash step. The sensors were submersed in a solution of trypsin-activated wild-type or F774D CDTb for 600 seconds, followed by a dissociation phase in fresh buffer for 600 seconds. The assay was conducted with an Octet RED96 (ForteBio) in 20 mM HEPES pH 8.0, 100 mM NaCl at 30°C.

Data Availability

The data that support the findings of this study are available from the corresponding author upon request. Protein Data Bank accession codes for all structures shown here are available as follows: 6O2O, 6O2N, 6OKR, 6O2M, 6OKT, 6OKS and 6OKU. Maps are available at the EM databank with accession codes 0608, 0609, 0610, 20102, 20103, 20104, and 20105.

Code Availability

The ImageJ macro scripts used in this work are available as Supplementary Software Files.
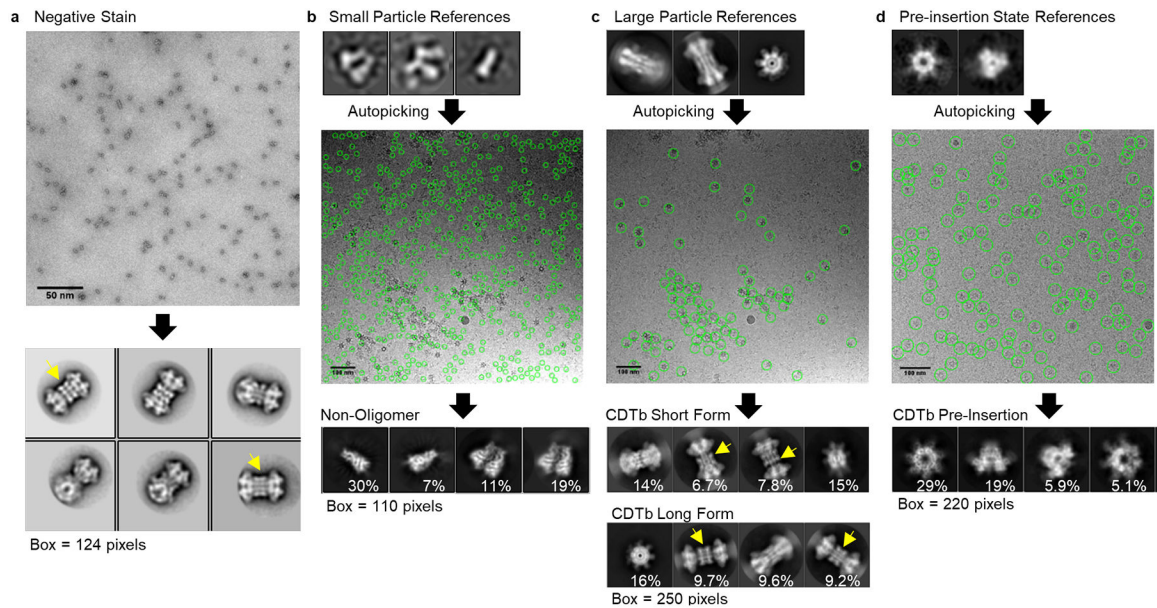
Extended Data
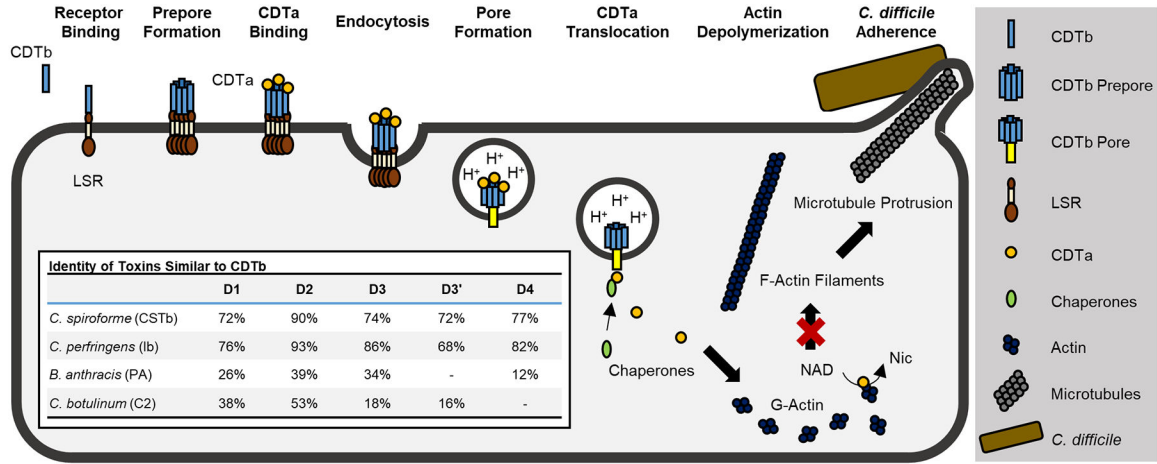
**Extended Data Fig. 1. A schematic representation of CDT intoxication.**

Once CDTb is secreted from the bacterium it targets the cell surface by binding the lipolysis stimulated lipoprotein receptor (LSR). CDTb then undergoes proteolytic cleavage allowing for oligomerization. It is thought that CDTa then binds the prepore, and this entire complex enters the cell through endocytosis. Upon acidification in the endosome, CDTb transitions from the prepore to the lipid-inserted pore state.

CDTa unfolds, is translocated into the cytosol, and then refolds with the help of cytosolic chaperones such as HSP90 and cyclophilin A. Upon refolding, CDTa ADP-ribosylates actin, leading to actin depolymerization and the formation of microtubule protrusions. These protrusions promote increased adherence of the bacterium. Similar Iota toxin family members such as CSTb (*C. spiroforme*) and Ib (*C. perfringens*) are thought to utilize a similar process based on a high degree of sequence identity (identity shown in insert). PA (*B. anthracis*) and C2 (*C. botulinum*) possess a much lower level of identity and are not considered part of the Iota toxin family.

The following table appears within the figure:

Identity of Toxins Similar to CDTb

	D1	D2	D3	D3'	D4
C. spiroforme (CSTb)	72%	90%	74%	72%	77%
C. perfringens (Ib)	76%	93%	86%	68%	82%
B. anthracis (PA)	26%	39%	34%	-	12%
C. botulinum (C2)	38%	53%	18%	16%	-

Extended Data Fig. 2. Autopicking and 2D classification of CDTb particles.
CDTb oligomers were first screened using negative stain EM as shown in panel (**a**). Representative 2D classes from this analysis revealed double heptamer particles in both the long and short form as denoted by the placement of the D4 ring (yellow arrows). CDTb particles were first picked manually to generate templates (shown at the top) to be used to autopick each dataset. The images corresponding to the larger molecular weight oligomers were analyzed twice, once for the smaller particles (**b**) and once for the larger particles (**c**). Micrographs collected on the smaller molecular weight oligomer were similarly picked (**d**). All data were classified and aligned in 2D as an initial means of purification (bottom). Analysis of the larger particles revealed two separate forms referred to as the short and long forms (**c**, middle) which differ in the location of a central ring as indicated by the yellow arrow. Though some views allowed for easy differentiation between the two forms, these features could not be observed in all classes. Classes representing different views observed during 2D classification are shown with percentages reflecting the number of particles per class divided by the total number of particles in the analysis. (**d**) Classification of the smaller molecular weight oligomeric sample of CDTb in complex with the cell surface receptor LSR. The classification revealed particles in only the pre-insertion state with the LSR, D3' and D4 domains not visualized, likely due to flexibility.
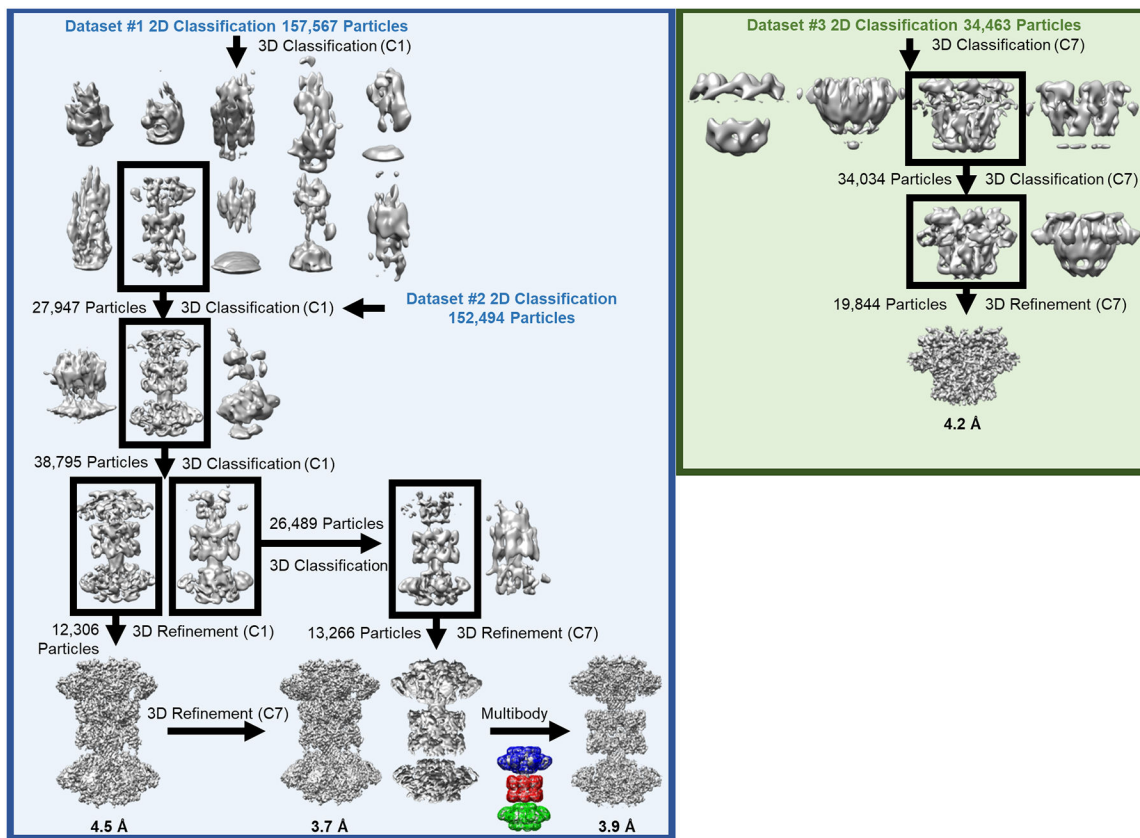
Extended Data Fig. 3. Cryo-EM reconstruction of CDTb density maps.
Cryo-EM maps were reconstructed in Relion from three distinct datasets. Datasets 1 and 2 were collected on the larger molecular weight oligomeric sample (shown in blue) and dataset 3 on the pre-insertion state (shown in green). Particles were autopicked from all three datasets in Relion and the purified particles from 2D classification used to generate initial 3D models. The first dataset was subjected to one round of 3D classification without symmetry. Particles were then selected from the best 3D class and joined with particles from the second dataset for further processing. These particles were then subjected to two more rounds of 3D classification without applying symmetry. Two forms of particles were obtained corresponding to the short form and long form. The short form particles were first refined using C1 symmetry to a resolution of 4.5 Å. After applying C7 symmetry, the resolution of the reconstruction reached 3.7 Å. An initial map of the long form was obtained at 6.3 Å. These data were further refined using multi-body refinement to yield a reconstruction at resolution of 3.9 Å. The pre-insertion state was obtained from particles picked from dataset 3. These particles were classified over two rounds of 3D classification with C7 symmetry applied. The set of particles was refined to a resolution of 4.2 Å with applied C7 symmetry.
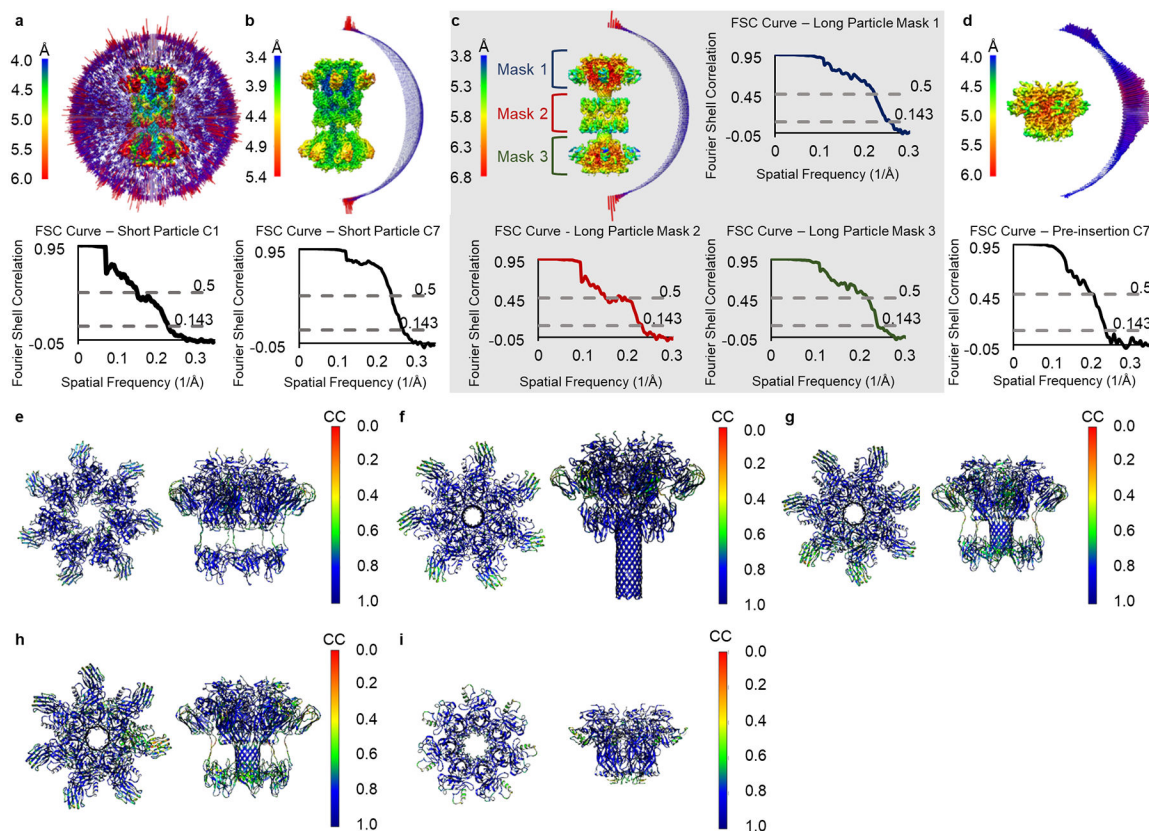
Extended Data Fig. 4. Cryo-EM resolution and model validation.
The local resolutions of the short form refined with C1 symmetry (**a**), the short form refined with C7 symmetry (**b**), the long form of CDTb refined with C7 (**c**) and the pre- insertion state (**d**), as determined by Relion LocalRes and displayed in UCSF Chimera. The angular distribution of particles used in the reconstruction of these maps is also shown. The corresponding Pearson's correlation coefficient (CC) for each residue is plotted on each model described in this manuscript: the CDTb oligomer (**e**), the pore (**f**), the partial β- barrel states (**g,h**), and the pre-insertion state (**i**).
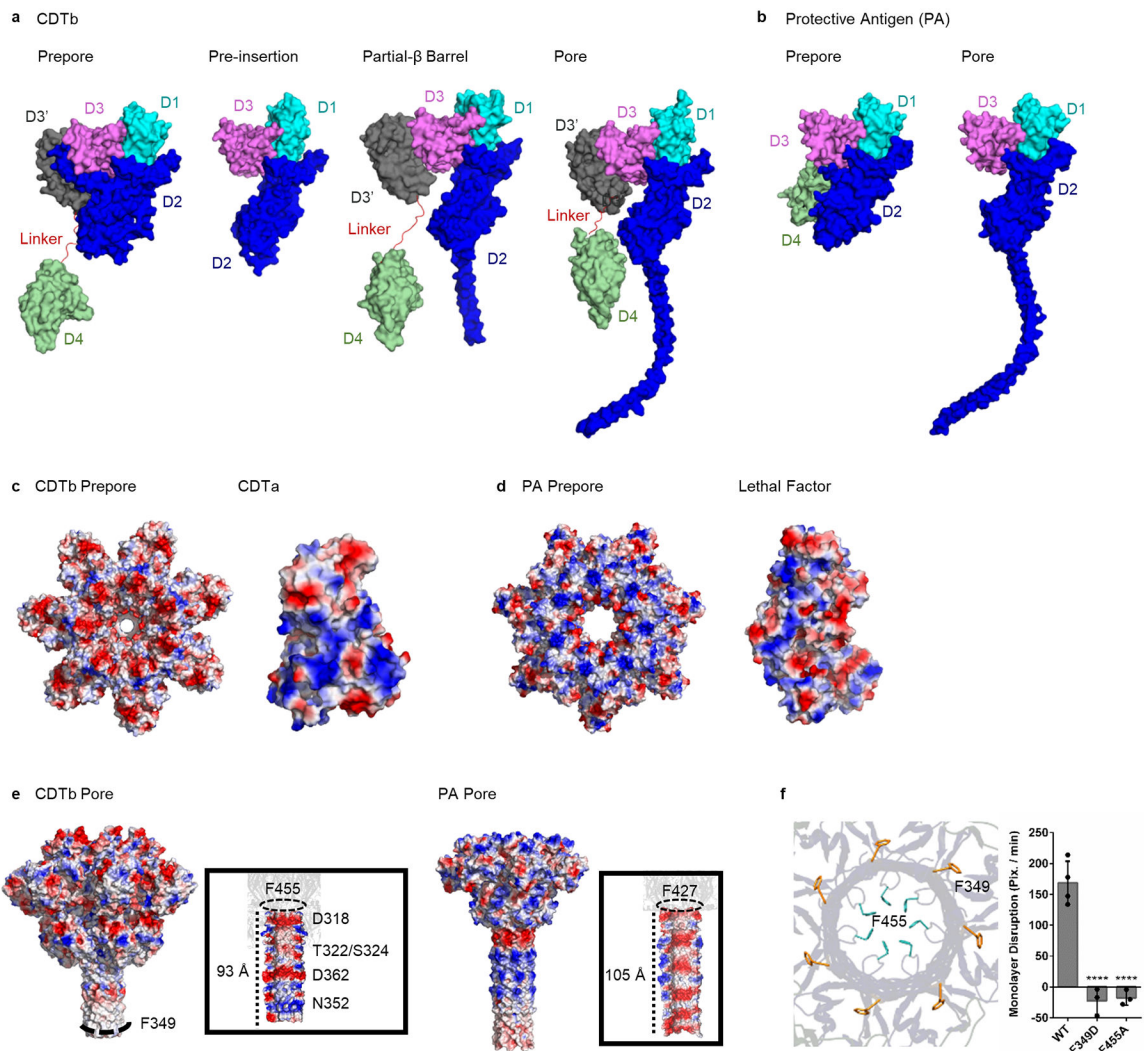
Extended Data Fig. 5. Comparison of CDTb to PA.

Surface rendering of the four CDTb cryo-EM structures colored by the five CDTb domains (**a**). Analogous domains for the prepore (PDB: 1TZO) and pore state (PDB: 3J9C) of PA (**b**). Vacuum electrostatic calculations of the surface of the CDTb (**c, left**) and PA (**d, left**) prepore show a distinct difference in electrostatic character between the two toxins. This surface may allow for the recognition of different cargo as suggested by the change in electrostatics observed in CDTa (**c, right**) when compared to LF (**d, right**). A band of hydrophobicity is observed at the bottom of the pore form of both CDTb (**e, left**) and PA (**e, right**). This is presumed to be the surface inserted into the lipid bilayer. Of note, the interior of the prepore is not hydrophobic and there are minimal contacts with the portion of the CDTb pore that spans the membrane, suggesting that the *in vitro* formation of double heptamers is driven by the *trans* interactions in the D4 ring. The pores of CDTb and PA differ by ~12 Å in length as shown in the inset. The difference could be important for accommodating the size of different receptors or it could be important for the translocation of different lengths of cargo. Two acidic bands line the interior of the β-barrel pore of CDTb (**left, inset**) as opposed to the five observed in PA (**right, inset**). Proteins with mutations at

residues corresponding to the phenylalanine gate (F455) located at the top of the β-barrel (top panel) and a residue at the bottom of the β-barrel (F349) hypothesized to be inserted into the lipid bilayer were generated (**e, insets**). The ability of these mutants to disrupt monolayers (right panel) was significantly lower than that observed with wild type toxin (one-way ANOVA ****=P 0.0001) (**f**). Each data point represents the average of four technical replicates. The bar height represents the average of the 3 or 4 independent assays, and the error bars indicate the standard deviation
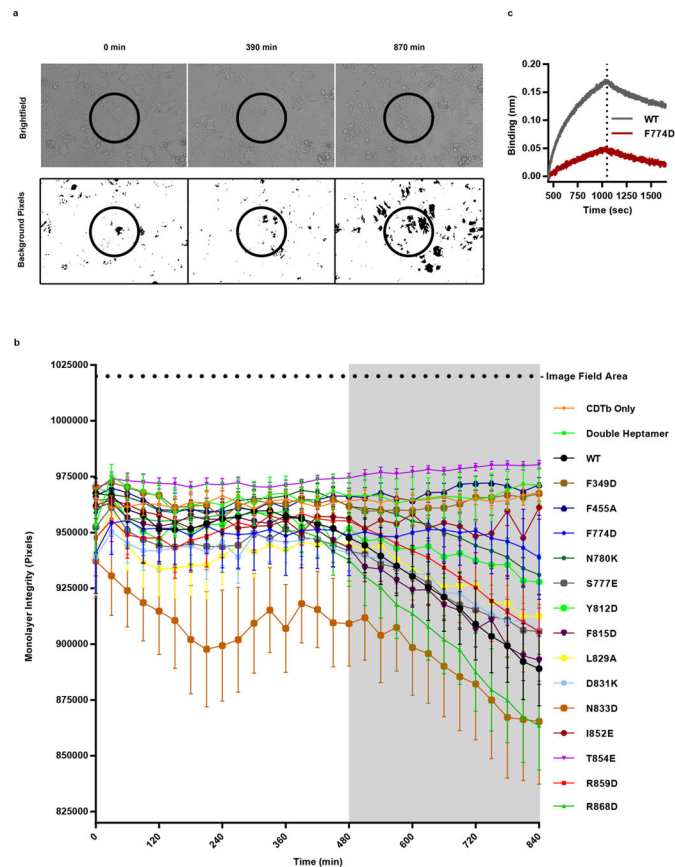
Extended Data Fig. 6. Cellular binding and intoxication assays.
(**a**) Representative bright-field images captured at the time-point indicated on a Cytation5 (BioTek) imaging plate reader (top) with the corresponding classified image showing background pixels in black (bottom). Data shown depicts cell rounding in the presence of 10 nM CDTa and 10 nM CDTb as highlighted by the area circled. (**b**) Classified images were quantified to produce a readout of monolayer integrity over the time course of the assay. The points represent the average of 3 or 4 independent biological replicates, and the error bars reflect the standard error of the mean. The region highlighted in gray was used for linear fitting to quantify the slope of pixel loss (cell coverage) over time. (**c**) Bio-layer interferometry (BLI) sensorgram comparing wild-type and F774D CDTb binding to NHS-biotinylated LSR immobilized on the sensor tip.
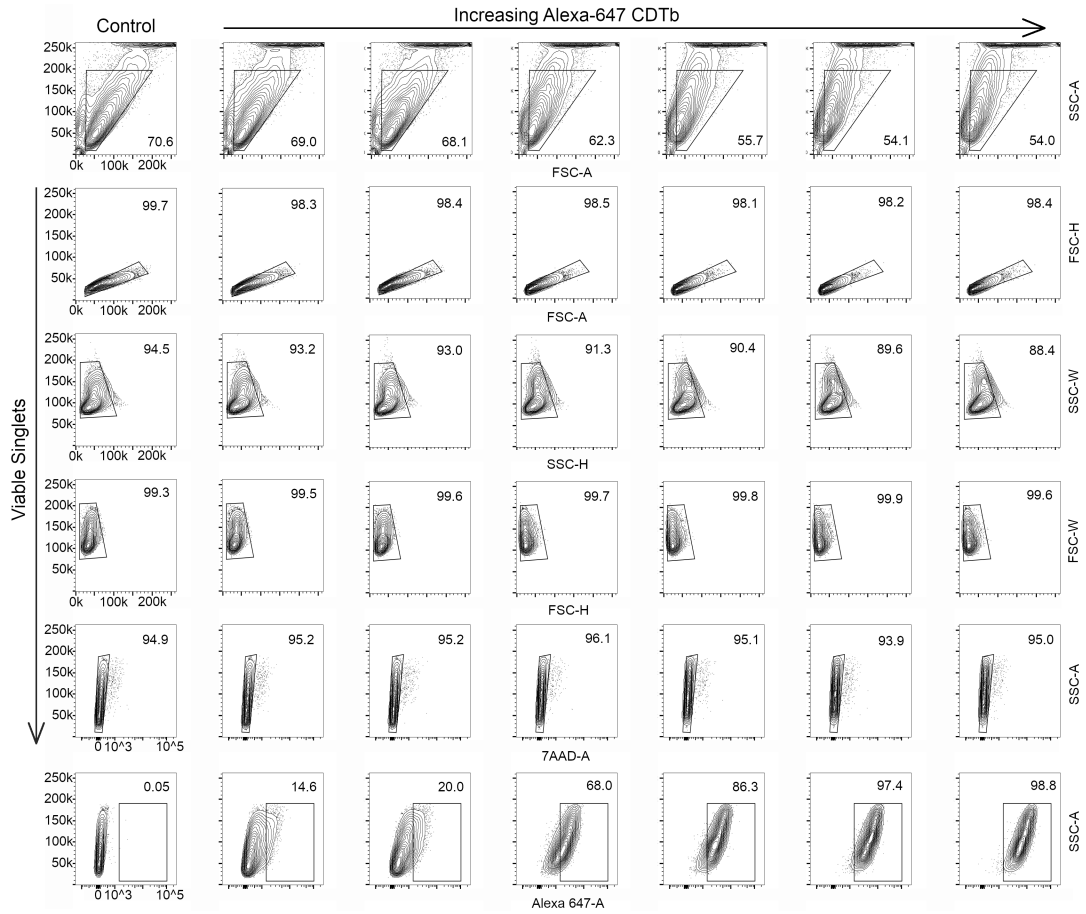
Extended Data Fig. 7. Representative flow cytometry scatter plots for the wild-type CDTb-Alexa-647 titration.
The far-left column represents the toxin-free control, while columns ascending to the right are plots of a two-fold increasing amount of toxin starting at 1.5 nM and ending at 50 nM toxin. Candidate cells were initially gated via side scatter area versus forward scatter area, followed by multiplet elimination via forward scatter height versus forward scatter area and side and forward scatter width versus height gating. Further intact, live cells were isolated next by a side scatter area versus 7AAD area gate. Lastly, the fraction of bound cells was calculated by the number of cells present in a side scatter area by Alexa-647 area gate. While this figure represents one experiment, the assay was independently performed twice for each protein and the data points are shown in Fig. 3b.
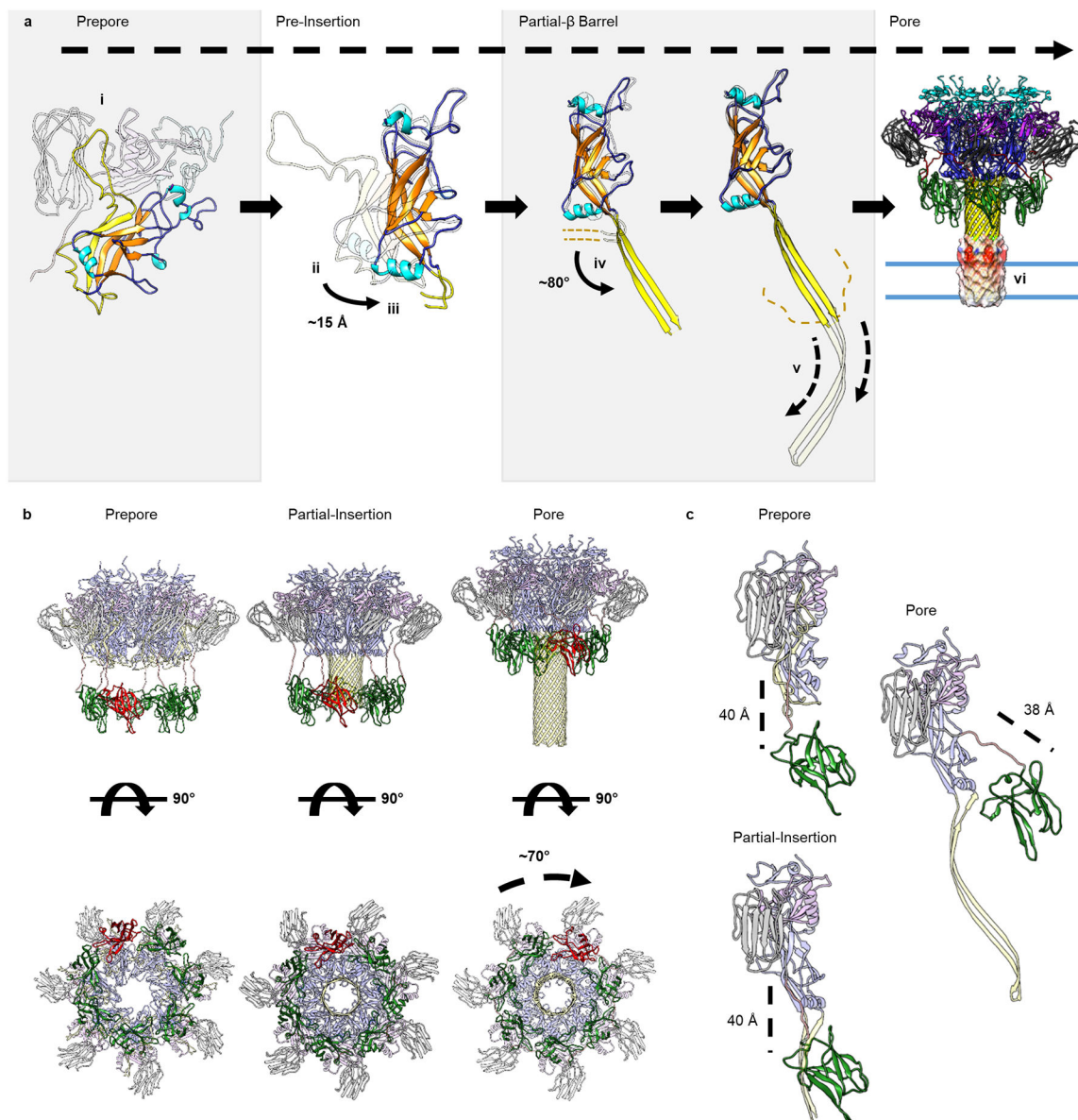
Extended Data Fig. 8. Transitions and flexibility of the D4 domain.

Conversion from the prepore state to the β-barrel pore state is shown in panel (**a**). The transition begins with the prepore form of CDTb in which the pore forming loop is docked into a cleft between the D3 and D3' domains (i). Next, a ~15 Å shift of the D2 domain (ii) to a position similar to that observed in the pore configuration. This results in displacement of the pore forming loop to a flexible conformation (iii). Upon formation of the β- barrel pore, the pore forming loop rotates ~80° relative to the eventual barrel (iv). We predict the remainder of the barrel then forms through a "zipper" like mechanism (v) which buries the hydrophobic tip of the barrel in the membrane (vi). We also observe conformational changes in the D4 domains as illustrated in panel (**b**). Though we observe a static arrangement of the D4 ring in the prepore and pre-insertion states, the flexibility of this ring is evident as it was rotated ~70° about the center of the ring, in the pore state. As a result, the extended linker
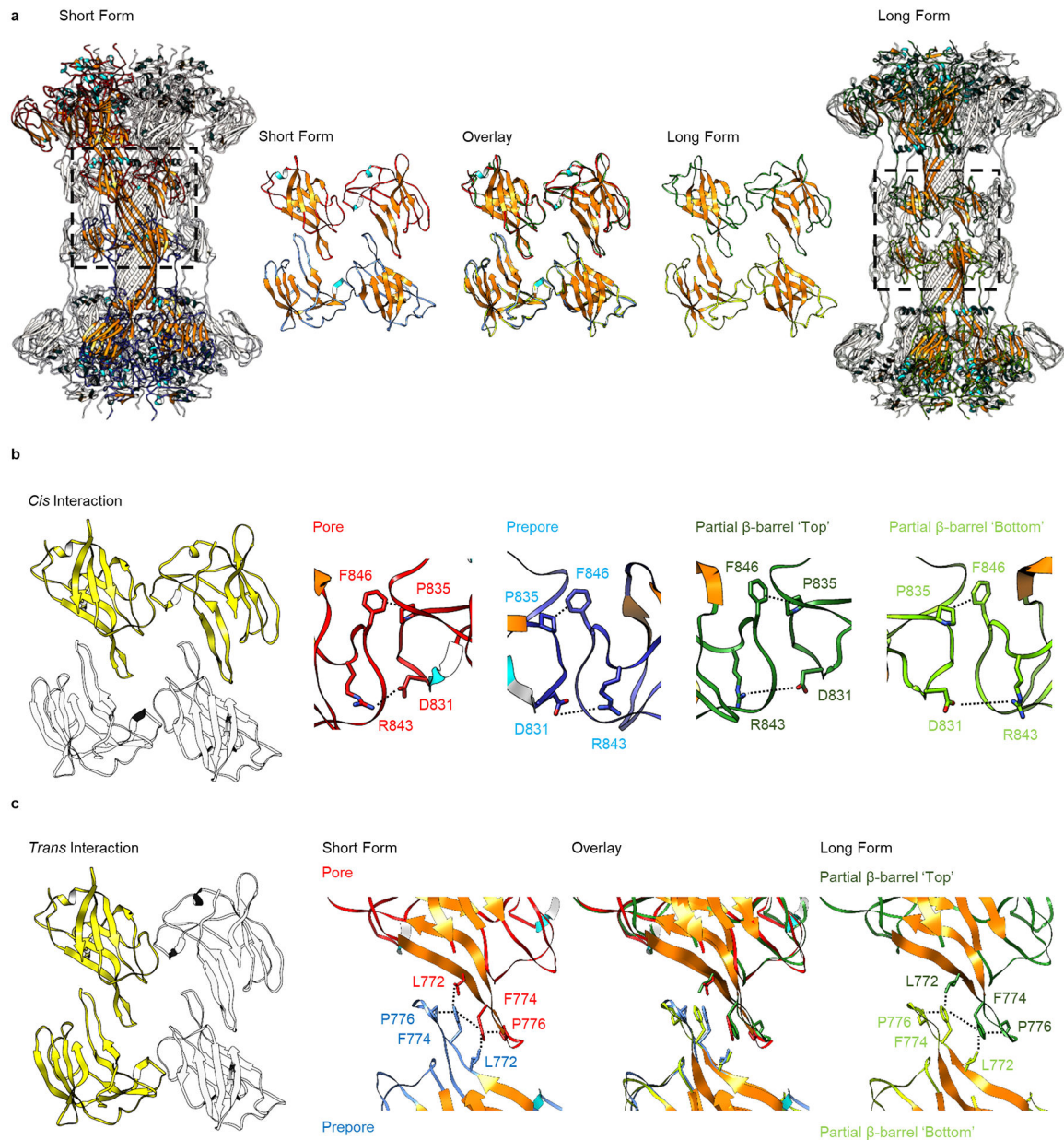
adopts a horizontal conformation as shown in panel (**c**). Though the orientations differ, the length of the linker remains relatively constant between all forms described here.

Extended Data Fig. 9. The 'cis' and 'trans' dimerization interfaces of the D4 domain.
The D4 receptor binding domain interacts with other D4 domains within the same oligomer (referred to as the 'cis' interaction) as well as the adjacent oligomer (referred to as the 'trans' interaction) in both the short and long forms (**a**). These interactions appear to be relatively weak, though, nearly identical in both forms. The cis interactions are comprised of electrostatic interactions between D831 and R843 as well as Van der Waals interactions between P835 and F846 in all observed orientations (**b**). Similarly, all trans interactions appear to exploit similar interactions consisting of Van der Waals interactions occurring between L772, F774, and P776.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References and Notes:

1. Lessa FC, Winston LG & McDonald LC Burden of Clostridium difficile infection in the United States. N. Engl. J. Med 372, 825–834 (2015). [PubMed: 25714160]

2. Chandrasekaran R & Lacy DB The role of toxins in clostridium difficile infection. FEMS Microbiol. Rev 41, 723–750 (2017). [PubMed: 29048477]

3. Gerding DN, Johnson S, Rupnik M & Aktories K Clostridium difficile binary toxin CDT: Mechanism, epidemiology, and potential clinical importance. Gut Microbes 5, 15–27 (2013). [PubMed: 24253566]

4. Aktories K, Papatheodorou P & Schwan C Binary Clostridium difficile toxin (CDT) - A virulence factor disturbing the cytoskeleton. Anaerobe 53, 21–29 (2018). [PubMed: 29524654]

5. Papatheodorou P et al. Lipolysis-stimulated lipoprotein receptor (LSR) is the host receptor for the binary toxin Clostridium difficile transferase (CDT). Proc. Natl. Acad. Sci 108, 16422–16427 (2011). [PubMed: 21930894]

6. Perelle S, Gibert M, Bourlioux P, Corthier G & Popoff MR Production of a complete binary toxin (actin- specific ADP- ribosyltransferase) by Clostridium difficile CD196. Infect. Immun 65, 1402–1407 (1997). [PubMed: 9119480]

7. Schwan C et al. Clostridium difficile toxin CDT induces formation of microtubule-based protrusions and increases adherence of bacteria. PLoS Pathog. 5, (2009).

8. Schwan C et al. *Clostridium difficile* toxin CDT hijacks microtubule organization and reroutes vesicle traffic to increase pathogen adherence. Proc. Natl. Acad. Sci 111, 2313–2318 (2014). [PubMed: 24469807]

9. Friebe S, van der Goot FG & Bürgi J The ins and outs of anthrax toxin. Toxins (Basel). 8, (2016).

10. Petosa C, Collier RJ, Klimpel KR, Leppla SH & Liddington RC Crystal structure of the anthrax toxin protective antigen. Nature 385, 833–838 (1997). [PubMed: 9039918]

11. Lacy DB, Wigelsworth DJ, Melnyk RA, Harrison SC & Collier RJ Structure of heptameric protective antigen bound to an anthrax toxin receptor: A role for receptor in pH-dependent pore formation. Proc. Natl. Acad. Sci 101, 13147–13151 (2004). [PubMed: 15326297]

12. Kintzer AF et al. The Protective Antigen Component of Anthrax Toxin Forms Functional Octameric Complexes. J. Mol. Biol 392, 614–629 (2009). [PubMed: 19627991]

13. Jiang J, Pentelute BL, Collier RJ & Hong Zhou Z Atomic structure of anthrax protective antigen pore elucidates toxin translocation. Nature 521, 545–549 (2015). [PubMed: 25778700]

14. Krantz BA et al. Microbiology: A phenylalanine clamp catalyzes protein translocation through the anthrax toxin pore. Science (80-.). 309, 777–781 (2005).

15. Wynia-Smith SL, Brown MJ, Chirichella G, Kemalyan G & Krantz BA Electrostatic ratchet in the protective antigen channel promotes anthrax toxin translocation. J. Biol. Chem 287, 43753–43764 (2012). [PubMed: 23115233]

16. Eckhardt M, Barth H, Blöcker D & Aktories K Binding of Clostridium botulinum C2 toxin to asparagine-linked complex and hybrid carbohydrates. J. Biol. Chem 275, 2328–2334 (2000). [PubMed: 10644682]

17. Hemmasi S et al. Interaction of the Clostridium difficile binary toxin CDT and its host cell receptor, lipolysis-stimulated lipoprotein receptor (LSR). J. Biol. Chem 290, 14031–14044 (2015). [PubMed: 25882847]

18. Papatheodorou P et al. Identification of the cellular receptor of Clostridium spiroforme toxin. Infect. Immun 80, 1418–1423 (2012). [PubMed: 22252869]

19. Pfeifer G et al. 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 (2003). [PubMed: 12941936]

20. Kaiser E et al. Membrane translocation of binary actin-ADP-ribosylating toxins from Clostridium difficile and Clostridium perfringens Is facilitated by Cyclophilin A and Hsp90. Infect. Immun 79, 3913–3921 (2011). [PubMed: 21768281]

21. Madeira F et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res (2019). doi:10.1093/nar/gkz268

22. Stothard P The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28, 1102, 1104 (2000). [PubMed: 10868275]

23. Zheng SQ et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332 (2017). [PubMed: 28250466]

24. Mindell JA & Grigorieff N Accurate determination of local defocus and specimen tilt in electron microscopy. J. Struct. Biol 142, 334–347 (2003). [PubMed: 12781660]

25. Rohou A & Grigorieff N CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol 192, 216–221 (2015). [PubMed: 26278980]

26. Zivanov J et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife 7, (2018).

27. Kimanius D, Forsberg BO, Scheres SHW & Lindahl E Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife 5, (2016).

28. Afonine PV et al. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. Sect. D Struct. Biol 74, 531–544 (2018). [PubMed: 29872004]

29. Morin A et al. Collaboration gets the most out of software. Elife 2013, (2013).

30. Emsley P, Lohkamp B, Scott WG & Cowtan K Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr 66, 486–501 (2010). [PubMed: 20383002]

31. Pettersen EF et al. UCSF Chimera - A visualization system for exploratory research and analysis. J. Comput. Chem 25, 1605–1612 (2004). [PubMed: 15264254]

32. Peng T et al. A BaSiC tool for background and shading correction of optical microscopy images. Nat. Commun 8, 14836 (2017). [PubMed: 28594001]

33. Arganda-Carreras I et al. Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification. Bioinformatics 33, 2424–2426 (2017). [PubMed: 28369169]
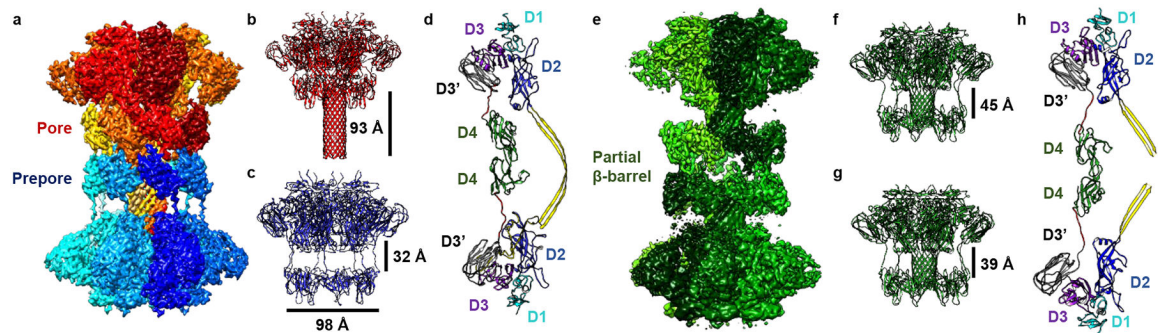
Fig. 1. Reconstruction and modelling of CDTb oligomers.
a, Side view of the cryo-EM density map obtained from the short particles described in this study, with the protomers forming the pore colored in red and those representing CDTb prepore colored in blue. **b-c**, Models of the pore (**b**) and CDTb prepore (**c**) built from the density map shown in panel **a**. **d**, Models that have been constructed detailing the protomers of the two forms described here with discrete domains labelled. **e**, A side view of the cryo-EM density of the partial β-barrel state. **f-g**, Nearly identical models of the partial β-barrel were constructed from both the 'top' (**f**) and 'bottom' (**g**) of the particle in which a 45 Å pore was observed. The D4 domains remain in an extended conformation 39 Å from the rest of the toxin. **h**, The dimerization interface in the partial β-barrel form remains largely similar to that observed in the pore-prepore complex (**d**).
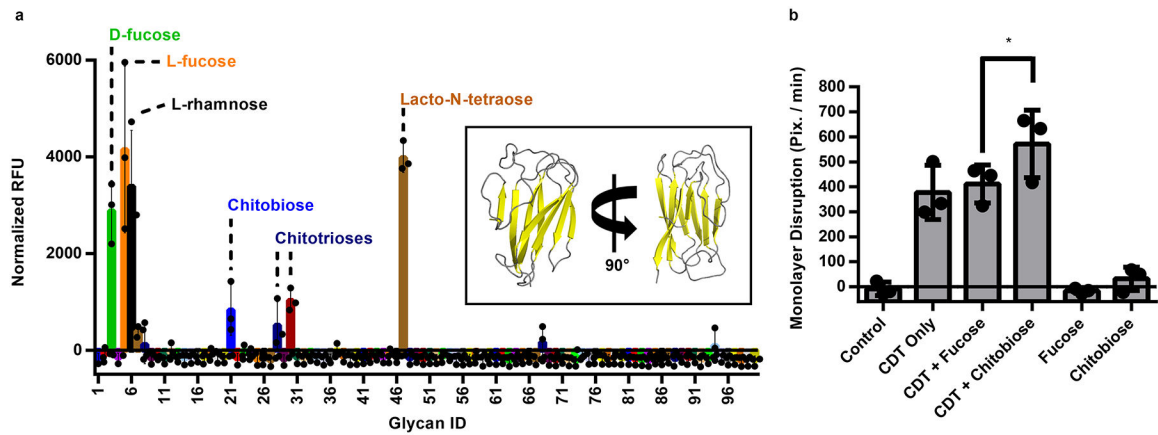
Fig. 2. The D3' domain engages glycans.

a, Solid-phase glycan binding results showing the relative levels of D3' interactions with several glycans. The data represent the average of signal from three replicate spots on a single glycan array 100 chip and are representative for what was observed on three independent glycan microarray chips. The D3' domain is shown in the inset. **b**, Caco-2 monolayers were treated with 10 nM CDTb and CDTa (indicated by CDT) in the absence (CDT only) or presence of 10 mM of select glycans identified as positive hits by the glycan binding array analysis. Conditions labeled with fucose contained 10 mM D-fucose and 10 mM L-fucose. Chitobiose was present at 10 mM in the conditions indicated. Each data point represents the average of three wells per condition, the bar height represents the average of the independent biological replicates (n = 3) and error is indicated as standard deviation. Statistical significance was determined by a one-way ANOVA, with * = P 0.05.
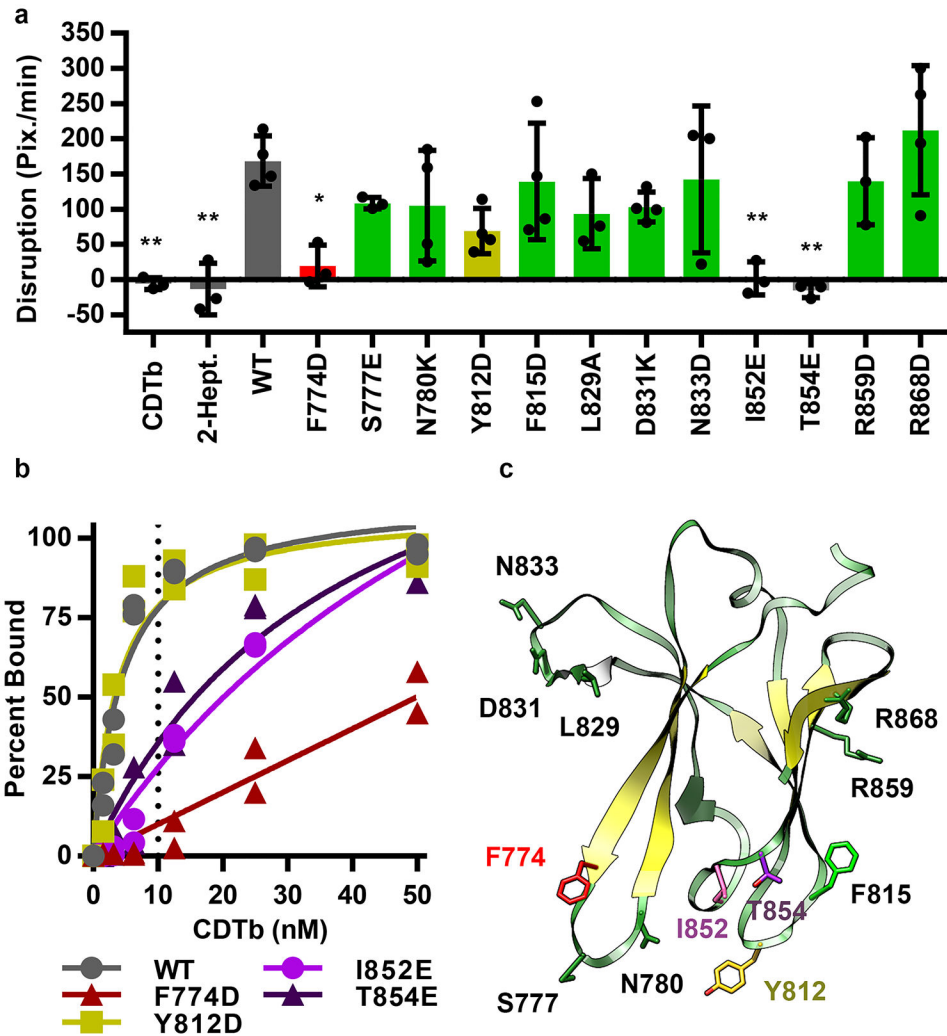
Fig. 3. The D4 domain binds LSR along the 'bottom' face.
a, Quantification of the monolayer disruption assay conducted with 10 nM CDTa and 10 nM CDTb. Bars are plotted as the average of three or four independent biological replicates (n = 3 or 4) with error bars indicating standard deviation. Each data point represents the average of technical replicates performed in quadruplicate. Statistical significance was evaluated with the Brown-Forsythe test (F=1.1, P=0.38, one-way ANOVA with *=P 0.05, **=P 0.01). '2-hept' denotes assays conducted with preformed double heptamer sample instead of monomeric CDTb. **b**, Graph of the percent of Caco-2 cells bound as a function of concentration of Alexa-647-labeled CDTb using flow cytometry. Data from two independent experiments for each toxin species were fit to a one-site specific binding model in GraphPad Prism 6.0. The derived apparent KDs were as follows: WT, 4.5 ± 0.9 nM; F774D, > 50 nM; Y812D, 3.9 ± 1.1 nM; I852E, 75 ± 34 nM; T854E, 38 ± 16 nM). **c**, Residues that were mutated in these assays are depicted on the D4 domain. Mutation of residues colored in purple or red resulted in proteins with drastically reduced cytopathic activity.
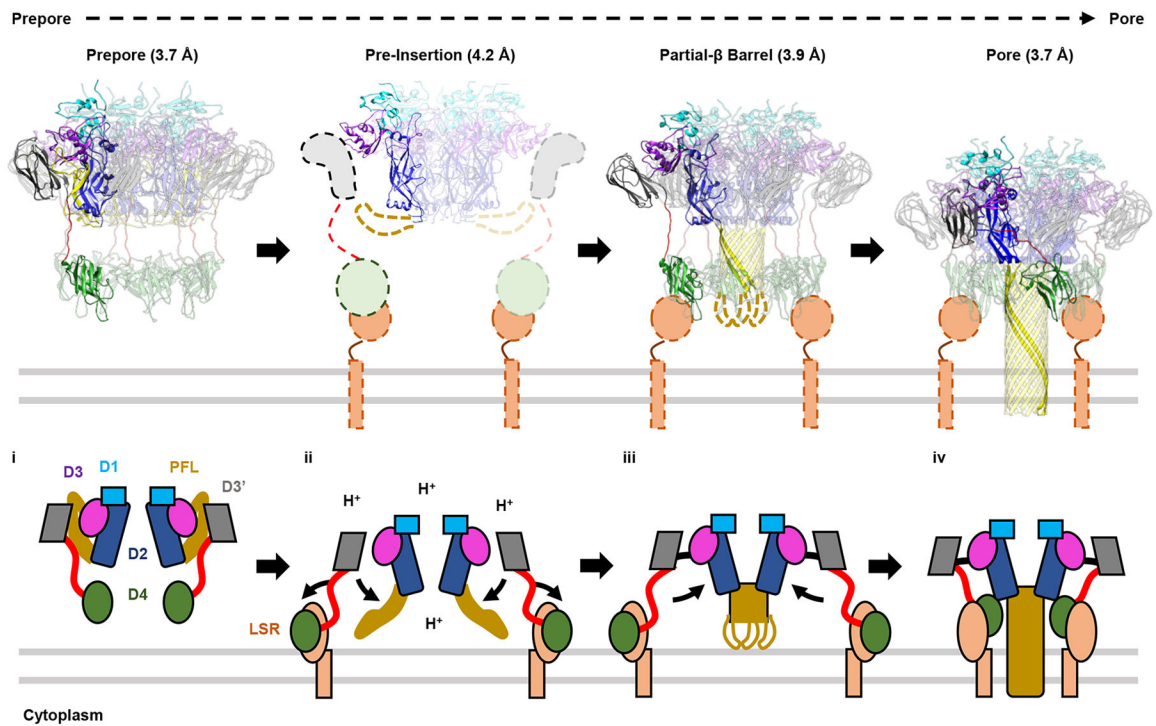
Fig. 4. Transition from the prepore to the pore state.

The structures presented in this paper (**top**) can be used to illustrate the process of CDTb pore formation (**bottom**). In the proposed mechanism, the prepore engages the receptor, LSR (**i**) which leads to a rearrangement of the D2 domain and release of the PFL (shown in yellow, **ii**). At this stage, the D4 ring is not assembled, resulting in a lack of density in the map and rearrangement of the PFL to initiate β-barrel formation. The β-barrel starts at the top and extends toward the membrane with a "zipper" like mechanism (**iii**). Upon pore formation the hydrophobic end of the β-barrel is buried within the membrane (**iv**). If receptor remains bound, the D4 ring can be accommodated at the same height from the membrane by incorporating the horizontal rotation about the y-axis.