Supplementary Material

Supplementary Figures

**Supplementary Figure 1** TcdB autoprocessing assay. (a) Schematic of TcdB1-804 autocleavage substrate. The domains within the substrate and L543 cleavage site are noted. (b) Activation of TcdB1-804 autoprocessing by InsP₆. The autocleavage substrate was incubated with InsP₆ for 1 hr at 37°C, and reactions were resolved by SDS-PAGE and visualized by Coomassie staining. The percent cleaved was determined by comparing the relative amounts of TcdB1-543 to full-length substrate using densitometry. The activation constant (EC₅₀), the concentration of InsP₆ at which half-maximal activity occurs, was determined to be 0.18 ± 0.02 μM for TcdB1-804. This InsP₆ concentration is well below the cytosolic concentration of InsP₆ (5-100 μM¹,²), indicating that the InsP₆-induced activation of the CPD occurs at physiologically relevant concentrations.
Supplementary Figure 2  Autoprocessing of TcdB<sub>1-804</sub> during expression in *E. coli*. Cultures expressing either wildtype or C698A TcdB<sub>1-804</sub> were grown at 37°C with shaking. At time 0’, IPTG was added, and aliquots were removed at the indicated timepoints and analyzed for production of His<sub>6</sub>-tagged TcdB<sub>1-804</sub> using anti-His<sub>6</sub> Western blot analysis. After 3 hrs of expression, the cultures were harvested, and His<sub>6</sub>-tagged TcdB<sub>1-804</sub> was affinity purified. The resulting eluate was run on an SDS-PAGE gel and stained using Coomassie. Wildtype TcdB<sub>1-804</sub> is processed at the natural cleavage site sequence by the CPD upon production in *E. coli* whereas C698A is not. TcdB<sub>1-804</sub> produced by both strains is subject to degradation by proteases within *E. coli* (marked with asterisks).
Supplementary Figure 3 AWP19 labeling of *V. cholerae* MARTX (MARTX\(_{Vc}\)) CPD. (a) InsP\(_6\)-induced labeling of MARTX\(_{Vc}\) CPD by AWP19. Recombinant MARTX\(_{Vc}\) CPD (1 \(\mu\)M) was mixed with AWP19 (10 \(\mu\)M) and the indicated concentration of InsP\(_6\) and incubated for 1 hr at 37\(^\circ\)C. The labeling reactions were resolved by SDS-PAGE, and fluorescence was measured using a flatbed scanner, followed by visualization of total protein using Coomassie staining. (b) Labeling of *V. cholerae* MARTX CPD by AWP19 incubated in the presence or absence of InsP\(_6\) for the indicated time period. No increase in labeling of MARTX\(_{Vc}\) CPD is observed in the absence of InsP\(_6\).
Supplementary Figure 4 AWP19 labeling of TcdB CPD variants used in limited proteolysis assay. TcdB CPD (20 µM) was either untreated, treated with N-ethylmaleimide (NEM; 200 µM), or treated with Ac-GSL-AOMK (200 µM) for 16 hr. InsP₆-activated TcdB CPD was treated with InsP₆ for 1 hr. The TcdB CPD variants were then re-purified using the His₆-tag. Re-purified variants were either untreated or treated with InsP₆ (10 µM final) and then mixed with AWP19 (10 µM) for 1 hr at 37°C. The labeling reactions were resolved by SDS-PAGE, and fluorescence was measured using a flatbed scanner, followed by visualization of total protein using Coomassie staining.
Supplementary Figure 5 Structural comparison of InsP$_6$-bound TcdB CPD with InsP$_6$-bound TcdB CPD modified with an Ac-GSL-AOMK inhibitor. InsP$_6$-bound TcdB CPD (green) was overlaid with Ac-GSL, InsP$_6$-bound TcdB CPD (cyan, PDB ID 35A8)$^3$. Ac-GSL, InsP$_6$, the β-flap, and the N- and C-termini are labeled for orientation. Overall C-α RMSD was calculated to be 0.6 Å. Structures overlaid using Pymol.
Supplementary Fig. 6 Electron density of TcdB CPD in the InsP₆ binding site. 2Fo-Fc density contoured at 1.5σ.
Supplementary Tables

### Supplementary Table 1 Primers used in this study.

| #  | Name                          | Sequence      | RE  |
|----|-------------------------------|---------------|-----|
| 01 | 5' NdeI TcdB 544 aa           | ATACATAATGGGTGAAAGATGATAATCTTT | NdeI |
| 02 | 3' XhoI TcdB 797 aa           | GCAGCTCTGATAGCTCGTAAATTTTCTTAA | XhoI |
| 03 | 5' R575Q SOE                  | TTAGCACAGAAGTTTCAGAGCAAAGGATATATACTATATTTGTT | |
| 04 | 3' R575Q rev OES              | AACAATATAATGTAATATCTTTCTTGGCTCAAA | |
| 05 | 5' K600N SOE                  | GCATGTTACTTTATTTGCAAATACCTCTTATTGGATGATCTTGG | |
| 06 | 3' K600N rev OES              | CATGATTAACATATTTTCAATTTAATCTATATTTTCTTATATTTG | |
| 07 | 5' K645N SOE                  | AGTTAGCAATTAATATTTTCAATTACATATTTTCTTATATTTG | |
| 08 | 3' K645N rev OES              | ATGGCGAATTTTTTTCAATTACATATTTTCTTATATTTG | |
| 09 | 5' K647N SOE                  | ATGGCGAATTTTTTTCAATTACATATTTTCTTATATTTG | |
| 10 | 3' K647N rev OES              | ATGGCGAATTTTTTTCAATTACATATTTTCTTATATTTG | |
| 11 | 5' R751Q SOE                  | GTTAAATATGGAAGGACAAAGAAGATATTTGATGTAATTTTCTTATATTTG | |
| 12 | 3' R751Q rev OES              | ATGGCGAATTTTTTTCAATTACATATTTTCTTATATTTG | |
| 13 | 5' R751C SOE                  | GTTAAATATGGAAGGACAAAGAAGATATTTGATGTAATTTTCTTATATTTG | |
| 14 | 3' R751C rev OES              | GTTAAATATGGAAGGACAAAGAAGATATTTGATGTAATTTTCTTATATTTG | |
| 15 | 5' R752Q SOE                  | GTTAAATATGGAAGGACAAAGAAGATATTTGATGTAATTTTCTTATATTTG | |
| 16 | 3' R752Q rev OES              | ATGGCGAATTTTTTTCAATTACATATTTTCTTATATTTG | |
| 17 | 5' K764N SOE                  | GTTAAATATGGAAGGACAAAGAAGATATTTGATGTAATTTTCTTATATTTG | |
| 18 | 3' K764N rev OES              | ATGGCGAATTTTTTTCAATTACATATTTTCTTATATTTG | |
| 19 | 5' K775N SOE                  | ATGGCGAATTTTTTTCAATTACATATTTTCTTATATTTG | |
| 20 | 3' K775N rev OES              | ATGGCGAATTTTTTTCAATTACATATTTTCTTATATTTG | |

*Abbreviations: RE - Restriction Enzyme*
Restriction enzyme sequences are underlined, and point mutations are in bold. Restriction site
### Supplementary Table 2: E. coli strains used in study.

| Strain | Genotype and relevant features | Reference |
|--------|--------------------------------|-----------|
| BL21(DE3) |                             | Novagen   |
| 041 | DH5α                          | D.E. Cameron |
| 007 | pET22b in DH5α                | D.E. Higgins |
| 269 | pET28a in DH5α                | E. Ponder  |
| 300 | pET28a-TcdB<sub>1-804</sub> (wt) |           |
| 303 | pET28a-TcdB<sub>1-804</sub> (C698A) | 3         |
| 313 | pET22b-TcdB<sub>544-797</sub> (wt) | This study |
| 320 | pET22b-TcdB<sub>544-797</sub> (C698A) | This study |
| B317 | pET22b-TcdB<sub>544-797</sub> (R575Q) | This study |
| 336 | pET22b-TcdB<sub>544-797</sub> (K600N) | This study |
| 337 | pET22b-TcdB<sub>544-797</sub> (K645N) | This study |
| 318 | pET22b-TcdB<sub>544-797</sub> (K647N) | This study |
| 349 | pET22b-TcdB<sub>544-797</sub> (R751Q) | This study |
| 350 | pET22b-TcdB<sub>544-797</sub> (R752Q) | This study |
| 342 | pET22b-TcdB<sub>544-797</sub> (K764N) | This study |
| 338 | pET22b-TcdB<sub>544-797</sub> (K775N) | This study |
| 341 | pET22b-TcdB<sub>544-797</sub> (E749A) | This study |
| 348 | pET22b-TcdB<sub>544-797</sub> (E753Q) | This study |
| 367 | pET22b-TcdB<sub>544-797</sub> (E753N) | This study |
| 421 | pET22b-TcdB<sub>544-797</sub> (E753R-R753E) | This study |
| 368 | pET22b-TcdB<sub>544-797</sub> (N747S) | This study |
| 390 | pET22b-TcdB<sub>544-797</sub> (N747A) | This study |
| 355 | pET22b-TcdB<sub>544-797</sub> (R745K) | This study |
| 356 | pET22b-TcdB<sub>544-797</sub> (R745Q) | This study |
| 369 | pET22b-TcdB<sub>544-797</sub> (R745N) | This study |
| B319 | pET22b-TcdB<sub>544-797</sub> (W761F) | This study |
| B320 | pET22b-TcdB<sub>544-797</sub> (W761A) | This study |
| 430 | pET22b-TcdB<sub>544-797</sub> (L755A) | This study |
| 431 | pET22b-TcdB<sub>544-797</sub> (L755S) | This study |
| 400 | pET22b-TcdB<sub>544-797</sub> (L755C) | This study |
| 399 | pET22b-TcdB<sub>544-797</sub> (L755C-W761C) | This study |
| 407 | pET22b-TcdB<sub>544-797</sub> (W761C) | This study |
| 420 | pET22b-TcdB<sub>544-797</sub> (R745C-W761C) | This study |
| 397 | pET22b-TcdB<sub>544-797</sub> (R745C-E753C) | This study |
| 398 | pET22b-TcdB<sub>544-797</sub> (R751C-E765C) | This study |
| 396 | pET22b-TcdB<sub>544-797</sub> (I762C-E766C) | This study |
Supplementary Methods

Strain construction. Primers used are listed in Supplementary Table 1; strains constructed are listed in Supplementary Table 2. To construct the strain expressing His6-tagged TcdB544-797, primers #01 and #02 were used to amplify the sequence encoding amino acids 544-797 of TcdB using pET28a-TcdB1-804 as a template. The amplified DNA was digested with NdeI and XhoI, ligated to pET22b digested with the same enzymes, and transformed into DH5α. The resulting pET22b-TcdB544-797 plasmid was transformed into BL21(DE3). Point mutations within TcdB CPD544-797 were constructed using PCR SOE3,4.

Analysis of TcdB1-804 autocleavage during production in E. coli. Overnight cultures of pET28a-TcdB1-804 (wildtype or C698A mutant)3 were diluted 1:500 in 1 L 2YT media and grown shaking at 37°C. When an OD600 of 0.6 was reached, a 1 mL sample was pelleted and resuspended in 60 µL 1X FSB (– sample). IPTG was added to the culture at 250 µM, the culture was transferred to 30°C, and culture aliquots were removed every 30 min. The samples were re-suspended in 1X FSB to normalize for cell density, boiled at 95°C for 5 min prior to loading on a 10% SDS-PAGE gel. Following electrophoresis, the samples were transferred to nitrocellulose membrane and analyzed by Western blotting using an anti-His antibody.

Purification of TcdB1-804 autocleavage substrate. Cultures were grown as described above; cells were lysed after 3 hr growth at 30°C following IPTG induction. His6-tagged TcdB1-804 variants were purified as described in the Methods section.
**In vitro TcdB autoprocessing assay.** Recombinant autocleavage substrate was mixed with the indicated concentrations of InsP₆ and incubated at 37°C for 1 hr. Autocleavage was assessed by SDS-PAGE followed by Coomassie staining. Reactions were performed in triplicate, and the amount of cleaved substrate relative to full-length substrate was quantified using densitometry, averaged, and plotted against concentration of InsP₆.

**Purification of TcdB(544-797).** For purification of TcdB(544-797) variants, overnight cultures of the appropriate BL21(DE3) strain was diluted 1:500 in 2L 2YT media and grown shaking (225 rpm) at 37°C. When an OD₆₀₀ of 0.6-0.9 was reached, IPTG was added to 250 µM, and cultures were grown for 12-16 hr at 16°C. Cultures were pelleted, resuspended in 25 mL lysis buffer [500 mM NaCl, 50 mM Tris-HCl, pH 7.5, 15 mM imidazole, 10% (v/v) glycerol] and flash frozen in liquid nitrogen. Lysates were thawed, then lysed by sonication and cleared by centrifugation at 15,000 x g for 30 minutes. His₆-tagged CPD was affinity purified by incubating the lysates in batch with 1.0 mL Ni-NTA Agarose beads (Qiagen) with shaking for 3 hr at 4°C. The binding reaction was pelleted at 1,500 x g, the supernatant was set aside, and the pelleted Ni-NTA agarose beads were washed 3 x with lysis buffer. His₆-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, 10% (v/v) glycerol]. The elution was repeated four times; the eluate was pooled, buffer exchanged in gel filtration buffer [200 mM NaCl, 10 mM Tris pH 7.5, 5% (v/v) glycerol], and concentrated to 750 µL. The concentrated prep was pelleted at 13,000 x g for 10 min at 4°C prior to loading on a Superdex 200 10/30 column (GE Healthcare).
Purified His$_6$-tagged CPD was concentrated (between 200 and 600 µM) and stored at -20°C in gel filtration buffer.

For crystallization studies, C-terminally His$_6$-tagged TcdB CPD (544-797 aa) was affinity purified as described earlier. His$_6$-tagged TcdB CPD was gel purified using a HiPrep S200 16/60 Sephacryl column (GE Healthcare); the gel filtration buffer was 150 mM NaCl, 10 mM Tris pH 7.5. The purified protein was concentrated to 1 mM, and InsP$_6$ was added to a final concentration of 2 mM for crystallization screening.

**Apo-CPD labeling with AWP19.** A mastermix of 100 µL CPD (1 µM) was incubated with AWP19 (10 µM) in the absence (apo) or presence of 10 µM InsP$_6$. Thirty µL aliquots were removed at the indicated timepoints and added to 10 µL 4X FSB. The samples were loaded in triplicate and resolved by SDS-PAGE as described above. The labeling reactions were quantitated using ImageJ, and the percentage labeling was determined by comparing the signal of apo-CPD to InsP$_6$-bound signal at the same timepoint. To determine the residual activity of TcdB CPD after prolonged incubation with Ac-GSL-AOMK inhibitor, the experiment was performed as described above except that Ac-GSL-AOMK was used instead of AWP19. At the indicated timepoint, 10 µM of AWP19 and 10 µM InsP$_6$ were added to the reactions; the labeling reaction was allowed to proceed for 30 min at room temperature, after which aliquots were removed into sample buffer. The amount of residual activity was determined by comparing the amount of AWP19 labeling in the presence of Ac-GSL-AOMK treatment relative to the untreated sample at the same timepoint.
**Linear rates of AWP19 labeling.** A mastermix of 200 µL CPD (1 µM) and AWP19 (10 µM) was made and then InsP$_6$ was added to the indicated concentration. A 20 µL aliquot was immediately removed into 7.5 µL 4X FSB for the 0’ timepoint and then 20 µL aliquots were removed into 7.5 µL 4X FSB at the indicated timepoints. Ten µL of sample was resolved by SDS-PAGE as described above. Labeling reactions were quantitated using the ImageJ program and plotted using Kaleidagraph. The linear rates of labeling were determined using the Linear function on Kaleidagraph. These rates were then plotted against InsP$_6$ concentration, and the $K_D$ was determined using the Michaelis-Menten function on Kaleidagraph. Results are representative of three independent experiments.

**Complete Trypsin Proteolysis.** TcdB CPD (1 mM) was either untreated, treated with 2 mM InsP$_6$ (diluted from a 25 mM stock of InsP$_6$ in water), a two-fold excess of InsP$_6$ and Ac-GSL-AOMK (diluted from a 200 mM stock in DMSO), or a two-fold excess of InsP$_6$ and NEM (diluted from 50 mM stock in DMSO). The proteins were diluted and buffer exchanged into gel filtration buffer then purified by gel filtration as described above. The proteins were concentrated to between 300 and 500 µM. The TcdB CPD variants were diluted to 80 µM in limited proteolysis buffer (10 mM Tris pH 7.5, 150 mM NaCl) and split into two aliquots. To one tube, 10 µL of 1 mg mL$^{-1}$ trypsin was added, and 15 µL aliquots were removed at the indicated timepoints after incubation at room temperature. Ten microlitres of sample was resolved on a 14% SDS-PAGE gel and visualized by Coomassie staining. Ten microlitres of 1 mg mL$^{-1}$ trypsin was added to the second tube and incubated at room temperature for 90 minutes. The sample was then boiled at 95°C
for 10 min, and precipitated protein was pelleted by centrifuging the boiled mixture for 5 min at 13,000 rpm. The supernatant was removed and used in the InsP₆ transfer assay.

**InsP₆ transfer assay.** MARTXᵥc CPD activity was measured using the fluorogenic cleavage assay described in ⁵. Specifically, MARTXᵥc CPD was diluted into 7.2 mL CPD buffer to a final concentration of 2.35 μM; Z-GGL-AMC substrate was diluted to a final concentration of 100 μM (BioMol from a 50 mM DMSO stock), and 85 μL of the mastermix was aliquoted into a 96 well plate. Five microlitres of the TcdB CPD variants used in the complete trypsin digestion assay (either the supernatant from the trypsinized, heat-inactivated TcdB CPD or the untreated form), was added to MARTXᵥc CPD in triplicate. The samples were mixed and then either 10 μL of CPD buffer or 10 μL of InsP₆ (250 μM stock) was added. The plate was mixed again then hydrolysis of Z-GGL-AMC by recombinant CPD was followed using a Molecular Devices fmax plate reader at 37°C. Cleavage of the substrate as a function of time was followed by monitoring the emission at 460 nm following excitation at 380 nm. The initial velocity was determined from the linear portion of progress curve. The linear rates of Z-GGL-AMC cleavage were measured in triplicate and averaged for each TcdB CPD variant ± trypsin digestion; the concentration of InsP₆ transferred from TcdB CPD to MARTXᵥc CPD was determined by using the Michaelis-Menten equation derived from titrating InsP₆ against MARTXᵥc CPD in the fluorogenic cleavage assay. The InsP₆ titration curve was measured in triplicate. Varying amounts of InsP₆ were added to MARTXᵥc CPD (not exposed to TcdB CPD), and Z-GGL-AMC cleavage was measured as described. The Michaelis-Menten function on KaleidaGraph was used to derive the equation used for determining the concentration of InsP₆ transferred from TcdB CPD variants.
Supplementary References

1. Irvine, R.F. & Schell, M.J. Back in the water: the return of the inositol phosphates. *Nat Rev Mol Cell Biol* 2, 327-38 (2001).

2. Michell, R.H. Inositol derivatives: evolution and functions. *Nat Rev Mol Cell Biol* 9, 151-61 (2008).

3. Puri, A.W. et al. Rational Design of Inhibitors and Activity-Based Probes Targeting Clostridium difficile Virulence Factor TcdB. *Chem Biol accepted* (2010).

4. Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K. & Pease, L.R. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77, 61-68 (1989).

5. Lupardus, P.J., Shen, A., Bogyo, M. & Garcia, K.C. Small Molecule-Induced Allosteric Activation of the Vibrio cholerae RTX Cysteine Protease Domain. *Science* 322, 265-8 (2008).