Supplementary Information

Mechanistic and structural insights into the proteolytic activation of *Vibrio cholerae* MARTX toxin

Aimee Shen, Patrick J. Lupardus, Victoria E. Albrow, Andrew Guzzetta, James C. Powers, K. Christopher Garcia & Matthew Bogyo
Supplementary Figures

Supplementary Fig. 1

Supplementary Figure 1 Phylogenetic tree of selected putative CPD domains and clan CD proteases. Proteins above the red dashed line include putative CPDs with homology to MARTX$\text{V}_{\text{c}}$ CPD. FhaL proteins denoted in blue are adhesins containing domains homologous to MARTX CPDs that are encoded in the genomes of *Bordetella* sp. pathogens. CPDs found in MARTX toxins are denoted in green, while CPDs found in the large glucosylating toxins of *Clostridium* sp. are denoted in purple. MARTX$\text{V}_{\text{c}}$ CPD and *Clostridium difficile* CPDs are classified as clan CD proteases in Family C80.
Representative peptidases from each of the major clan CD protease families are found below the red dashed line\textsuperscript{11}. 
Supplementary Fig. 2

AC\textsubscript{en} (10 \, \mu M JCP485) = 1.1 \pm 0.1 \, nM

AC\textsubscript{en} (10 \, \mu M JCP479) = 21 \pm 3 \, nM

AC\textsubscript{en} (10 \, \mu M JCP650) = 11 \pm 2 \, nM

AC\textsubscript{en} (10 \, \mu M JCP654) = 158 \pm 19 \, nM

AC\textsubscript{en} (10 \, \mu M JCP657) = 5.2 \pm 0.3 \, nM
Supplementary Fig. 2 continued
Supplementary Figure 2  AC$_{50}(I)$ measurements for InsP$_6$-activated CPD autocleavage in the presence of 10 µM inhibitor mutants. Recombinant MARTX$_{Vc}$ CPD (amino acids 3391-3650) was pre-treated with 10 µM inhibitor then incubated with increasing concentrations of InsP$_6$. Autocleavage (performed in triplicate) was assessed by SDS-PAGE and Coomassie staining (inset). The amount of autocleaved protein relative to the total protein amount was analyzed by densitometry and plotted versus concentration of InsP$_6$. The AC$_{50}(I)$ was determined as the concentration of InsP$_6$ that produced half-maximal autocleavage in the presence of 10 µM inhibitor.
Supplementary Fig. 3

a

b
Supplementary Figure 3 (a) Stereo view of electron density (contoured at 1.5 $\sigma$) near aza-Leu epoxide inhibitor covalently bound to Cys3581 of the MARTX$_{Vc}$ CPD. (b) Superposition of unbound activated MARTX$_{Vc}$ CPD (red)$^{10}$ and inhibitor-bound activated MARTX$_{Vc}$ CPD (blue) (root-mean-square-deviation of 0.5 Å). Electron density around the Leu epoxide inhibitor is contoured at 1.5 $\sigma$. 
Supplementary Figure 4 Activation of *P. luminescens* MARTX CPD autocleavage by InsP₆.

Wildtype or catalytic-dead (CS) recombinant Plu3324 and Plu1341 CPDs from *P. luminescens* TTO1 (aa 2369-2636 and aa 2524-2775, respectively), and wildtype Plu1344 CPD (aa 2902-3173) MARTX toxins were incubated in the presence or absence of 10 μM InsP₆, and autocleavage was assessed by SDS-PAGE and Coomassie staining.
Supplementary Fig. 5

**AC$_{60}$ CPD = 0.85 ± 0.08 nM**

![Graph showing % cleaved wt vs. Concentration InsP$_6$ (nM) for AC$_{60}$ CPD]

**AC$_{60}$ αβ-CPD = 0.76 ± 0.27 nM**

![Graph showing % cleaved αβ-CPD vs. Concentration InsP$_6$ (nM) for AC$_{60}$ αβ-CPD]

**AC$_{60}$ ACD-RID = 88 ± 13 nM**

![Graph showing % cleaved ACD-RID vs. Concentration InsP$_6$ (nM) for AC$_{60}$ ACD-RID]

**AC$_{60}$ ACD-RID (LHAL/S) = 213 ± 13 nM**

![Graph showing % cleaved ACD-RID (LHAL/S) vs. Concentration InsP$_6$ (nM) for AC$_{60}$ ACD-RID (LHAL/S)]

**AC$_{60}$ RID-αβ = 511 ± 76 nM**

![Graph showing % cleaved RID-αβ vs. Concentration InsP$_6$ (nM) for AC$_{60}$ RID-αβ]
Supplementary Figure 5 AC₅₀ measurements for CPD-mediated cleavage of MARTXᵥc-derived polypeptides. Recombinant MARTXᵥc CPD (aa 3391-3650) and recombinant fragments of full-length MARTXᵥc (wildtype and mutant proteins) were incubated with increasing concentrations of InsP₆. Autocleavage was assessed by SDS-PAGE and Coomassie staining (inset). The amount of autocleaved protein relative to the total protein amount was analyzed by densitometry and plotted versus concentration of InsP₆. The AC₅₀ was determined as the concentration of InsP₆ that produced half-maximal autocleavage. * indicates a secondary cleavage at Leu3435 as determined by FT-MS.
Supplementary Fig. 6

Supplementary Figure 6 Model for CPD-mediated proteolytic activation of MARTX<sub>vc</sub> actin crosslinking activity. The inactive pro-toxin is secreted intact into the extracellular environment. Upon encountering a eukaryotic cell, the N- and C-terminal MARTX<sub>vc</sub> conserved regions insert into the eukaryotic plasma membrane and form a pore that permits translocation of the toxin central region across the cell membrane. Following entry into the eukaryotic cytoplasm, the CPD binds inositol hexakisphosphate (InsP<sub>6</sub>), resulting in activation of its protease activity. The CPD first cleaves at Leu3441, the preferred autoprocessing site, and then transcleaves MARTX<sub>vc</sub> at Leu2447, followed by Leu3099 (the lowest affinity cleavage site). Processing of MARTX<sub>vc</sub> by the CPD releases the Rho-inactivating domain (RID) and α/β hydrolase-like domain (α/β) domains into the cytosol, while the actin crosslinking domain (ACD) and CPD remain tethered to the membrane. Cleavage of MARTX<sub>vc</sub>, particularly at Leu2447, activates the actin crosslinking activity of the ACD.
Supplementary Figure 7 Comparison of the Cys-His catalytic dyad and active site topology of MARTX<sub>Vc</sub> CPD (aa 3391-3650) to the active sites of caspase-3 (PDB ID 2C1E)<sup>12</sup> and <i>P. gingivalis</i> gingipain R (PDB ID 1CVR)<sup>13</sup>.
**Supplementary Table 1** Data collection and refinement statistics (molecular replacement)

| Data collection | Native        |
|-----------------|---------------|
| Space group     | P2₁2₁2₁       |
| Cell dimensions |               |
| a, b, c (Å)     | 48.6, 65.8, 254.9 |
| α, β, γ (°)     | 90, 90, 90    |
| Resolution (Å)  | 30-2.35 (2.43-2.35)* |
| R_sym or R_merge| 0.099 (0.454) |
| R_meas          | 0.112 (0.511) |
| I / σI          | 13.6 (2.3)    |
| Completeness (%)| 89.0 (61.4)   |
| Redundancy      | 4.4 (3.6)     |

| Refinement      |               |
|-----------------|---------------|
| Resolution (Å)  | 30-2.35       |
| No. reflections | 29632         |
| R_work / R_free | 22.1/26.5     |
| No. atoms       |               |
| Protein         | 6220          |
| Ligand/ion      | 316           |
| Water           | 145           |
| B-factors       |               |
| Protein         | 49.4          |
| Ligand/ion      | 52.2          |
| Water           | 37.3          |
| R.m.s. deviations|             |
| Bond lengths (Å)| 0.010         |
| Bond angles (°) | 1.30          |

*Highest-resolution shell is shown in parentheses.*
Bacterial and eukaryotic cell growth conditions Bacterial strains were grown at 37°C in Luria-Bertani (LB) broth. Antibiotics were used at 100 μg/mL carbenicillin for pET22b and pGP704-Sac28 vectors in *E. coli* and *V. cholerae* and 30 μg/mL kanamycin for pET28a vectors in *E. coli*. HFF cells were grown in DMEM supplemented with glutamine and 10% fetal bovine serum (FBS) and maintained at 37°C in a 5% CO₂-air atmosphere.

Strain construction Primers used are listed in Supplementary Table 2; strains constructed are listed in Supplementary Tables 3 and 4. To construct the strain expressing His₆-tagged ACD domain that was used to generate anti-ACD antibodies, primers #01 and #02 were used to amplify the sequence encoding amino acids 1964-2375. The amplified DNA was digested with NdeI and XhoI, ligated to pET28a, and transformed into DH5α. The resulting pET28a-ACD plasmid was transformed into BL21(DE3). To construct the strain expressing the His₆-tagged RID domain used to generate anti-RID antibodies, primers #03 and #04 were used to amplify the sequence encoding amino acids 2552-3099. To construct the strain expressing His₆-tagged (i) ACD-α/β, primers #01 and #06 were used to amplify the sequence encoding amino acids 1964-3403, (ii) RID-α/β, primers #03 and #06 were used to amplify the sequence encoding amino acids 2459 to 3403, (iii) ACD-RID, primers #01 and #04 were used to amplify the sequence encoding amino acids 1964 to 3403. Strains expressing mutant His₆-tagged MARTX polypeptides were constructed by using splicing by overlap extension (PCR SOE)¹ to introduce point mutations into the His₆-tagged constructs. For example, to construct the pET22b-ACD-RID (L2447Δ) mutant, primers #01 and #09
were used to amplify a 5’ fragment comprising the ACD domain, and primers #08 and #04 were used to amplify a 3’ fragment comprising the RID domain. The resulting fragments were used as the templates for a second PCR reaction using the flanking primers, #01 and #04, to amplify an ACD-RID fragment carrying the L2447A mutation.

To construct strains expressing His$_6$-tagged *V. vulnificus* MARTX CPD, primers #20 and #21 were used to amplify DNA encoding residues 4044-4299 of VV20479 from *V. vulnificus* CMCP6 genomic DNA (P. Gulig). The catalytic Cys mutation was introduced by PCR SOE using internal SOE primers #22 and #23. To construct strains expressing His$_6$-tagged *P. luminescens* MARTX CPDs, primer pairs #24 and #25, #30 and #31, #34 and #35, and #36 and #37, were used to amplify DNA encoding residues 2390-2630 of Plu3217, residues 2369-2636 of Plu3324, residues 2524-2775 of Plu1341, and residues 2902-3173 of Plu1344, respectively. The catalytic Cys and L2408 cleavage site mutations for Plu3217 were introduced using PCR SOE primer pairs #26 and #27; and #28 and #29, respectively. The catalytic Cys mutation for Plu3324 was introduced by PCR SOE using primer pair #32 and #33, and the catalytic Cys mutation for Plu1341 was introduced by PCR SOE using primer pair #36 and #37.

For all His$_6$-tagged constructs, the procedures described for pET28a-ACD were followed, with the exception that some sequences were cloned into the pET22b expression vector.

To construct point mutations in *V. cholerae*, double homologous recombination using the sacB-lethality counterselection method was used$^2$. In general, ~0.5-0.6 kB of sequence flanking the point mutation or region to be deleted was amplified and then fused together in a second reaction using PCR SOE. For the L2447A mutant, the flanking primers were
#40 and #41, while the internal SOE primers were #08 and #09. For the L3099A mutant, the flanking primers were #42 and #43, and the internal SOE primers were #16 and #17. For the L3441A mutant, the flanking primers were #44 and #45, and the internal SOE primers were #18 and #19. SOE PCR fragments were ligated into the sucrose-based counter selectable plasmid pGP704-Sac28 using the Ncol/XbaI sites for the L3099A and L3441A vectors and Ncol/AgeI for the L2447A vector. Allelic exchange plasmids were transformed into *E. coli* SM10λpir and mated into the recipient *V. cholerae* strain.

Consistent with previous studies of *V. cholerae* MARTX function, the *V. cholerae* N19691 strain referred to as wildtype contains an in-frame deletion in *hapA* to reduce heterologous proteolysis of MARTXVc in culture supernatants. Sucrose-based counter selection was done essentially as described, and mutations were confirmed by PCR or genomic DNA sequencing. To construct the *V. cholerae* L2447A/L3099A strain, *E. coli* strain AS192 was conjugated to *V. cholerae* strain AS179 (L2447A). To construct the double mutants L2447A/L3441A and L3099A/L3441A, *E. coli* strain AS208 was conjugated to *V. cholerae* strains AS179 (L2447A) and AS187 (L3099A), respectively. The triple mutant (3X) was constructed by conjugating *E. coli* strain AS208 to *V. cholerae* strain AS200 (L2447A/L3099A).

**Crystallization and data collection** The CPD core domain (residues 3391-3650) was concentrated to 1 mM and mixed with a 2-fold excess (2 mM) of InsP₆ and JCP598 for crystallization. Diffraction quality crystals were grown in 0.5 μL sitting drops in an equal volume of 100 mM Tris-HCl pH 8.5, 30% PEG 3350 and 15% MPD at 22°C. Crystals were cryoprotected in the mother liquor and data was collected at 100K under
cryo-cooled conditions at a wavelength of 1.0 Å at the Stanford Synchrotron Radiation Laboratory (beamline 11-1). Diffraction data was processed using HKL2000\textsuperscript{8}, and $R_{\text{meas}}$ was calculated using SCALA\textsuperscript{9}. Data processing statistics are listed in Supplementary Table 1.

**A note on inhibitor analyses** Although the CPD inhibitors blocked MARTX\textsubscript{Vc} toxin activation, they failed to irreversibly inhibit CPD protease activity in our biochemical analyses. In the inhibitor structure, however, JCP598 was irreversibly bound to the catalytic Cys of activated CPD (Fig. 2). The unique activation mechanism of the CPD likely explains these seemingly incongruous results. Due to its allosteric activation mechanism, the CPD active site is only accessible to general alkylating agents and substrates when the protease is bound to InsP\textsubscript{6}; furthermore, InsP\textsubscript{6} exhibits rapid on-off-binding kinetics for the CPD\textsuperscript{10}. Thus, under normal assay conditions, our CPD inhibitors only transiently occupy the protease active site; as a result, the CPD catalytic Cys does not have sufficient time to react with the reactive electrophile of the inhibitor. In contrast, in the protein crystal, the protease is locked in an activated conformation such that the catalytic Cys is constitutively available to alkylate the inhibitor. These observations suggest that, rather than targeting the active site, a more effective class of CPD inhibitors might target the InsP\textsubscript{6} binding site or interfere with the structural transition induced by InsP\textsubscript{6} binding.
### Supplementary Tables. Primers used in this study.

| # | Name | Sequence<sup>a</sup> | RE<sup>b</sup> |
|---|------|----------------------|-------------|
| 01 | 5’ NdeI ACD | TTCCATATGTCAGGGTGAAACCTGCAAGT | NdeI |
| 02 | 3’ XhoI ACD | CCGCTCGAGCTCATACCGTAAACAAAGC | XhoI |
| 03 | 5’ NdeI RID | TTTTCAATGTCAGGGTGAAACTGCAAGT | NdeI |
| 04 | 3’ XhoI RID | CCGCTCGAGCTCATACCGTAAACAAAGC | XhoI |
| 05 | 5’ NdeI RID up | TTTTCAATGTCAGGGTGAAACTGCAAGT | NdeI |
| 06 | 3’ XhoI α/β | CCGCTCGAGCTGACTGACAAACAGC | XhoI |
| 07 | 3’ XhoI | CCGCTCGAGCTGACTGACAAACAGC | XhoI |

<sup>a</sup> Restriction enzyme sequences are underlined, and point mutations are in bold italics.  
<sup>b</sup> Restriction site
### Supplementary Tables continued: *V. cholerae* strains used in study

| Strain  | Genotype and relevant features                  | Strain Designation | Reference        |
|---------|------------------------------------------------|--------------------|------------------|
| AS028   | $\Delta hapA$ in *V. cholerae* N19691wild type |                    | G. Schoolnik     |
| AS063   | $\Delta rtxA$ in $\Delta hapA$                 | $\Delta rtxA$      | $^{10}$          |
| AS066   | C3581A in $\Delta hapA$                        | C3581A             |                  |
| AS179   | L2447A in $\Delta hapA$                        | L2447A             | This study       |
| AS187   | L3099A in $\Delta hapA$                        | L3099A             | This study       |
| AS200   | L2447A in L3099A                               | L2447A/L3099A      | This study       |
| AS202   | L3441A in $\Delta hapA$                        | L3441A             | This study       |
| AS210   | L2447A in L3441A                               | L2447A/L3441A      | This study       |
| AS212   | L3099A in L3441A                               | L3099A/L3441A      | This study       |
| AS204   | L3441A in L2447A/L3099A                        | 3X                 | This study       |

### Supplementary Tables continued: *E. coli* strains used in study

| Strain  | Genotype and relevant features                  | Reference        |
|---------|------------------------------------------------|------------------|
| AS039   | pSac28a in DH5$\alpha$                          | G. Schoolnik     |
| AS070   | pET8a-RID in BL21(DE3)                          | This study       |
| AS072   | pET28a-ACD in BL21(DE3)                         | This study       |
| AS125   | pET28a-ACD-$\alpha/\beta$ wt in BL21(DE3)       | This study       |
| AS106   | pET22b-ACD-RID$wt$ in BL21(DE3)                 | This study       |
| AS127   | pET28a-RID-$\alpha/\beta$ wt in BL21(DE3)       | This study       |
| AS196   | pET22b-ACD-RID (L2447A) in BL21(DE3)             | This study       |
| AS189   | pET22b-ACD-RID (L2447I) in BL21(DE3)             | This study       |
| AS223   | pET22b-ACD-RID (G2448S) in BL21(DE3)             | This study       |
| AS224   | pET22b-ACD-RID (G2448L) in BL21(DE3)             | This study       |
| AS214   | pET28a-RID-$\alpha/\beta$ (L3099A) in BL21(DE3) | This study       |
| AS090   | pET28a-CPD$wt$ (*V. vulnificus*) in BL21(DE3)   | This study       |
| AS092   | pET28a-CPD$CS$ (*V. vulnificus*) in BL21(DE3)   | This study       |
| AS045   | pET28a-CPD$wt$ (Plu3217) in BL21(DE3)            | This study       |
| AS051   | pET28a-CPD$CA$ (Plu3217) in BL21(DE3)            | This study       |
| AS049   | pET28a-CPD$L2408A$ (Plu3217) in BL21(DE3)        | This study       |
| AS094   | pET28a-CPD$wt$ (Plu3324) in BL21(DE3)            | This study       |
| AS096   | pET28a-CPD$CS$ (Plu3324) in BL21(DE3)            | This study       |
| AS077   | pET28a-CPD$wt$ (Plu1341) in BL21(DE3)            | This study       |
| AS078   | pET28a-CPD$CS$ (Plu1341) in BL21(DE3)            | This study       |
| AS076 | pET28a-CPD<sub>wt</sub> (Plu1344) in BL21(DE3) | This study |
|-------|------------------------------------------|------------|
| AS132 | pSac28a-L2447<sub>A</sub> in Sm10       | This study |
| AS192 | pSac28a-L3099<sub>A</sub> in Sm10       | This study |
| AS208 | pSac28a-L3441<sub>A</sub> in Sm10       | This study |

Point mutations in MARTX<sub>Ve</sub> are underlined in bold.
3 JCP485 MW = 364.4
4 JCP479, MW = 385.5
5 JCP650 MW = 477.6
6 JCP654  MW = 562.7
7 JCP598  MW = 590.7
8 JCP557 MW = 590.7
9 JCP599 MW = 590.7
10 VEA223 MW = 564.58
11 AS04 MW = 482.6
12 AS01 MW = 611.7
13 AS02 MW = 739.9
Supplementary References

1. 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).

2. Fullner, K.J. & Mekalanos, J.J. Genetic characterization of a new type IV-A pilus gene cluster found in both classical and El Tor biotypes of *Vibrio cholerae*. *Infect. Immun.* **67**, 1393-1404 (1999).

3. Nielsen, A.T. et al. RpoS controls the *Vibrio cholerae* mucosal escape response. *PLoS Pathog.* **2**, e109 (2006).

4. Boardman, B.K. & Satchell, K.J. *Vibrio cholerae* strains with mutations in an atypical type I secretion system accumulate RTX toxin intracellularly. *J. Bacteriol.* **186**, 8137-8143 (2004).

5. Fullner, K.J. & Mekalanos, J.J. In vivo covalent cross-linking of cellular actin by the *Vibrio cholerae* RTX toxin. *EMBO J.* **19**, 5315-5323 (2000).

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

7. Boardman, B.K., Meehan, B.M. & Fullner Satchell, K.J. Growth phase regulation of *Vibrio cholerae* RTX toxin export. *J. Bacteriol.* **189**, 1827-1835 (2007).

8. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326 (1997).

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

10. 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**, 265-268 (2008).

11. Rawlings, N.D., Morton, F.R., Kok, C.Y., Kong, J. & Barrett, A.J. MEROPS: the peptidase database. *Nucleic Acids Res.* **36**, D320-325 (2008).

12. Ganesan, R. et al. Exploring the S4 and S1 prime subsite specificities in caspase-3 with aza-peptide epoxide inhibitors. *Biochemistry* **45**, 9059-9067 (2006).

13. Eichinger, A. et al. Crystal structure of gingipain R: an Arg-specific bacterial cysteine proteinase with a caspase-like fold. *EMBO J.* **18**, 5453-5462 (1999).
