nature chemical biology

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

Aimee Shen[1], Patrick J Lupardus[2], Victoria E Albrow[1], Andrew Guzzetta[1], James C Powers[3], K Christopher Garcia[2] & Matthew Bogyo[1]

MARTX toxins modulate the virulence of a number of Gram-negative *Vibrio* species. This family of toxins is defined by the presence of a cysteine protease domain (CPD), which proteolytically activates the *Vibrio cholerae* MARTX toxin. Although recent structural studies of the CPD have uncovered a new allosteric activation mechanism, the mechanism of CPD substrate recognition or toxin processing is unknown. Here we show that interdomain cleavage of $MARTX_{Vc}$ enhances effector domain function. We also identify the first small-molecule inhibitors of this protease domain and present the 2.35-Å structure of the CPD bound to one of these inhibitors. This structure, coupled with biochemical and mutational studies of the toxin, reveals the molecular basis of CPD substrate specificity and underscores the evolutionary relationship between the CPD and the clan CD caspase proteases. These studies are likely to prove valuable for devising new antitoxin strategies for a number of bacterial pathogens.

Bacterial toxins are critical mediators of the host-pathogen interface. Recently, a new family of toxins, the multifunctional autoprocessing repeats in toxins (MARTX) toxins, was identified in the genomes of Gram-negative bacterial pathogens, including bacteria of the *Vibrio*, *Aeromonas*, *Photorhabdus* and *Yersinia* genera[1]. Although only a few MARTX family members have been characterized, MARTX toxins modulate the virulence of a number of bacterial pathogens. The MARTX toxin of the marine pathogen *Vibrio anguillarum* induces hemolysis and is essential for virulence in Atlantic salmon[2], while the MARTX toxin of the opportunistic zoonotic pathogen *Vibrio vulnificus* causes cytotoxicity and is required for full virulence in mice[3–5]. Similarly, the MARTX toxin of *Vibrio cholerae*, the etiological agent of cholera, promotes colonization of mice[6,7] and is produced by nearly all clinical and environmental isolates[8,9].

MARTX toxins are large secreted proteins that are defined by specific structural features[1]. Glycine-rich repeat regions in the N and C termini of MARTX toxins likely form a pore within host cell membranes that transfers the toxin central region into the eukaryotic cytoplasm. This central region is comprised of multiple activity domains that presumably impart distinct functionalities to a given toxin. However, only the effector domains within *V. cholerae* MARTX toxin have been characterized. Two of these domains alter host actin dynamics: the actin crosslinking domain (ACD) covalently crosslinks actin monomers[10], while the Rho-inactivating domain (RID) inhibits the function of small Rho protein GTPases[11]. A third domain, the cysteine protease domain (CPD)[12], functions as an autoprocessing cysteine protease that is required for activation of $MARTX_{Vc}$ toxin in eukaryotic cells[12]. The proteolytic function of the CPD is proposed to activate MARTX toxins by liberating $MARTX_{Vc}$ effector domains from the plasma membrane[12]. Notably, the CPD is completely conserved in all MARTX family members and is always found adjacent to the C-terminal glycine-rich repeat region[1].

The CPD is a new protease that is regulated by a unique allosteric activation mechanism[13]. Binding of the eukaryotic-specific small molecule inositol hexakisphosphate ($InsP_6$) to a basic cleft within the CPD induces a structural rearrangement that exposes the protease active site to its substrates. The responsiveness of the CPD to $InsP_6$ spatially restricts $MARTX_{Vc}$ toxin function to the eukaryotic cytosol[13]. Notably, distantly related homologs of the CPD are found in the glucosylating toxins of *Clostridium* sp. (**Supplementary Fig. 1** online)[12]. Similar to $MARTX_{Vc}$ toxin, the CPD domains of clostridial toxins are activated by $InsP_6$, and activation of the CPD is required for *Clostridium difficile* toxin B function[14–16].

While the general details of CPD activation have been established, the mechanisms underlying CPD-mediated $MARTX_{Vc}$ toxin activation, substrate recognition and catalysis remain unknown. $MARTX_{Vc}$ CPD exhibits weak structural similarity to clan CD proteases, including caspases and gingipain-R. This observation suggests that, in spite of their disparate mechanisms of activation, these proteases may share similar catalytic mechanisms[13]. However, $MARTX_{Vc}$ CPD has proven resistant to all known inhibitors of these clan CD proteases[12]. In this study, we identified a series of new inhibitors of CPD activity by screening a highly focused library of small-molecule protease inhibitors. Using a combination of chemical, structural and mutational approaches, we defined the substrate specificity of $MARTX_{Vc}$ CPD and mapped multiple CPD cleavage sites within $MARTX_{Vc}$. These data demonstrate that the CPD cleaves exclusively after a P1 leucine within interdomain regions, an event that is required for optimal

[1]Department of Pathology and [2]Department of Molecular and Cellular Physiology and Howard Hughes Medical Institute, Stanford School of Medicine, Stanford, California, USA. [3]Department of Chemistry, Georgia Institute of Technology, Atlanta, Georgia, USA. Correspondence should be addressed to M.B. (mbogyo@stanford.edu) or A.S. (ashen2@stanford.edu).

Received 4 November 2008; accepted 25 March 2009; published online 24 May 2009; doi:10.1038/nchembio.178

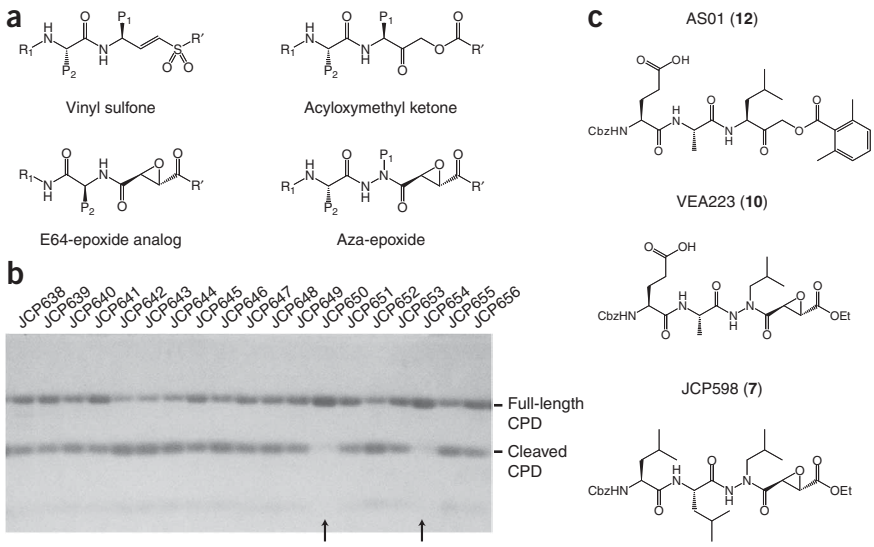

Figure 1 Identification of $MARTX_{Vc}$ CPD autoprocessing inhibitors. (**a**) General structures of the main classes of covalent cysteine protease inhibitors in the library used for screening. (**b**) Sample gel from CPD autoprocessing inhibitor screen. Recombinant $MARTX_{Vc}$ CPD (amino acids 3391–3650) was pretreated with 100 μM inhibitor for 15 min, after which GTPγS was added at 200 μM to activate autoprocessing. Cleavage reactions were resolved by SDS-PAGE and visualized by Coomassie stain. Compound numbers in the database are shown; arrowheads indicate hits in the screen. (**c**) Structures of the most potent inhibitors of CPD-mediated autoprocessing.

activity of a given domain. Our analyses also indicate that chemically inhibiting CPD function prevents $MARTX_{Vc}$ toxin activation; chemical inhibition of the CPD likely occurs through a mechanism similar to that of caspases. This study furthers our understanding of protease-mediated activation of bacterial toxins, validates the CPD domain as a target for developing antitoxin therapies and provides a structural basis for developing improved inhibitors of this and other related virulence factors.

RESULTS

Chemical inhibitors of $MARTX_{Vc}$ CPD

Many bacterial toxins undergo proteolytic activation upon encountering a eukaryotic cell[17]. Whereas most toxins are activated by host proteases, the MARTX toxin family is auto-activated by an internal cysteine protease domain[12]. Because genetic inactivation of the catalytic cysteine of the CPD prevents *V. cholerae* MARTX function, we sought to chemically inhibit the protease activity of $MARTX_{Vc}$. To identify inhibitors of CPD function, we screened a unique library of 498 cysteine protease inhibitors[18] for the ability to block recombinant CPD autoprocessing *in vitro*. This library is composed of cysteine protease–specific peptide vinyl sulfones, acyloxymethyl ketones (AOMKs), aza-peptide epoxides and epoxysuccinates (**Fig. 1a**)[19]. Compounds were screened by pretreating recombinant pro-enzymes (containing 50 residues upstream of the CPD cleavage site) with each compound and then adding guanosine 5′-[γ-thio]triphosphate (GTPγS, 1) to induce activation of the protease activity. Although GTPγS is a much less potent activator of the CPD than inositol hexakisphosphate[13,20] ($InsP_6$, 2), GTPγS was used as the activating compound because at the time this was the only known activator of CPD protease activity. Compounds that blocked autoprocessing were identified by SDS-PAGE analysis (**Fig. 1b**). This screen identified eight aza-peptide epoxides that exhibited reproducible, dose-dependent inhibitory activity in our assay (**Table 1**). Notably, all eight compounds contained leucine in the P1 position, which suggests a high degree of selectivity of this protease for the residue directly adjacent to the scissile amide bond.

To compare the potencies of each inhibitor, we measured the concentration of $InsP_6$ required to activate half-maximal cleavage of the CPD in the presence of 10 μM inhibitor (AC_{50}(I), **Table 1** and **Supplementary Fig. 2** online). A large AC_{50}(I) is indicative of a better CPD inhibitor, since more $InsP_6$ is required to activate cleavage in the presence of a fixed amount of inhibitor. It should be noted that this assay only measures *cis* autocleavage events, as autocleavage of recombinant $MARTX_{Vc}$ CPD in *trans* is strongly disfavored due to steric hindrance[13]. Based on these measurements, we generated a small structure-activity relationship series using the eight inhibitors identified in our screen (**Table 1**). Most notably, inhibitor potency correlated with peptide length: addition of a P3 leucine increased inhibitor potency by ~40-fold (11 ± 2 nM versus 457 ± 80 nM; JCP650 versus JCP598). Inhibitor potency was also dependent on the regio- and stereochemistry at the epoxide moiety, with the order of inhibition being *S,S* > trans ≫ *R,R* (**Table 1**). Notably, this same preference for the *trans S,S* aza-peptide epoxide has been observed for the caspases[21], implying that the CPD and caspases share similar mechanisms of substrate recognition.

Based on this observation, we hypothesized that functional groups previously used as caspase inhibitors might also inhibit CPD protease

Table 1 AC_{50}(I) values for CPD inhibitors identified in screen and for inhibitors designed to block CPD autoprocessing activity

Number	Designation	Inhibitor[a]	EP[b]	AC_{50} (nM) at 10 μM inhibitor[c]
–	–	No inhibitor	n/a	0.9 ± 0.1
3	JCP485	Z-ALeu-EP-COO-Et	*trans*	1.1 ± 0.1
4	JCP479	Ac-Leu-ALeu-EP-COO-Et	*S,S*	21 ± 3
5	JCP650	Z-Leu-ALeu-EP-COO-Et	*S,S*	11 ± 2
6	JCP654	Z-Leu-Leu-ALeu-EP-COO-H	*S,S*	158 ± 19
7	JCP598	Z-Leu-Leu-ALeu-EP-COO-Et	*S,S*	457 ± 80
8	JCP657		*R,R*	5.2 ± 0.3
9	JCP599		*trans*	74 ± 8
10	VEA223	Z-Glu-Ala-ALeu-EP-COO-Et	*S,S*	429 ± 83
11	AS04	Z-Ala-Leu-AOMK	n/a	187 ± 30
12	AS01	Z-Glu-Ala-Leu-AOMK	n/a	529 ± 108
13	AS02	Z-Lys-Glu-Ala-Leu-AOMK	n/a	290 ± 52

[a]Z, Ph-CH_2-O-C(O)CO–; Aleu, aza-Leu; EP, epoxide; Et, ethyl. [b]The *trans* epoxide is a mixture of *S,S* and *R,R*, while the *cis* epoxide is a mixture of *R,S* and *S,R*. n/a, not applicable. [c]AC_{50}(I) represents the concentration of $InsP_6$ required to activate half-maximal cleavage of the CPD in the presence of 10 μM of inhibitor. Since the extent of inhibition depends on the concentration of $InsP_6$ in the assay, the higher the AC_{50}(I) value, the better the inhibitor. AC_{50}(I) values were determined from triplicate experiments (± s.d.).

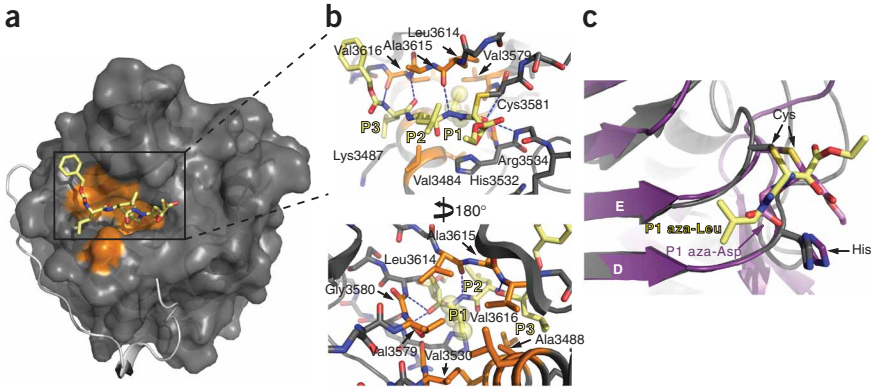

Figure 2 Structure of activated $MARTX_{Vc}$ CPD bound to an aza-peptide epoxide inhibitor. (**a**) Surface topology of the CPD active site. Hydrophobic residues in the substrate binding cleft are highlighted in orange. The aza-peptide epoxide inhibitor (JCP598) is shown as a stick model bound in the substrate binding pocket. The N terminus is shown as a gray ribbon, terminating at Ile5 and highlighting the threading of this region along the surface of the core domain. (**b**) Close-up 'top' and 'bottom' views of the S1 pocket. Hydrophobic residues in the S1 pocket are shown as orange sticks, and the side chain atoms of the P1 aza-leucine residue are shown as transparent spheres. Hydrogen bonds between the inhibitor backbone and the protein are shown as dashed lines. (**c**) Superposition of the D and E β-strands of caspase-3–aza-Asp epoxide (PDB ID 2C1E) and CPD–aza-leucine epoxide inhibitor structures shown as a cut-away view of the thioether inhibitor adduct bound in the S1 pocket. Caspase-3 is colored purple, and the aza-Asp inhibitor is colored pink. The $MARTX_{Vc}$ CPD is colored gray, and the aza-leucine inhibitor is colored yellow.

activity. Thus, we synthesized AOMK inhibitors[19] carrying the P4-P1 (KEAL) residues of the Leu3441 cleavage site and evaluated their efficacy in the CPD autocleavage assay. We also synthesized an aza-peptide epoxide containing the P3-P1 positions of the Leu3441 cleavage site (VEA223) to directly compare the contribution of the functional group to inhibitor strength (**Fig. 1c**). As with the aza-leucine epoxide inhibitors, the presence of the P3 residue increased inhibitor potency (529 ± 108 nM versus 187 ± 30 nM; AS01 versus AS04). Addition of the P4 residue, however, did not improve inhibitor potency, perhaps because the hydrophobic Cbz (Ph-CH_2-O-C(O)) group of AS01 was replaced with a basic lysine residue in AS04 (290 ± 52 nM versus 529 ± 108 nM; AS02 versus AS01).

While the presence of P2 and P3 residues enhanced inhibitor potency, the protease exhibited a somewhat broad selectivity in these positions, since the EAaL (VEA223) and LLaL (JCP598) epoxides had similar AC_{50}(I) values (**Table 1**). The clan CD–specific AOMK and aza-peptide epoxide functional groups were also equally effective at inhibiting CPD function (JCP598 versus VEA223, **Table 1**). Inhibition of CPD activity was specific to these functional groups, since the proteasome inhibitors MG132 (Cbz-LLL-aldehyde, 14) and Z-L_3VS (Cbz-LLL-vinyl sulfone, 15) failed to inhibit CPD function (data not shown). Taken together, our results strongly imply that optimal inhibition of CPD activity requires compounds with a P1 leucine linked to either the AOMK or aza-epoxide functional groups.

Crystal structure of inhibitor-bound, activated $MARTX_{Vc}$ CPD

To gain insight into the mechanism of chemical inhibition of the CPD, we cocrystallized and solved the structure of activated, $InsP_6$-bound CPD in complex with the aza-leucine epoxide inhibitor JCP598 (**Fig. 2a**). The overall structure of inhibitor-bound, activated CPD is nearly identical to our previous unbound structure of activated CPD (r.m.s. deviation of 0.5 Å) (**Supplementary Fig. 3** online)[13]. This superposition indicates that the inhibitor essentially docks into an active site cleft created upon binding of $InsP_6$ to the CPD; no significant changes in active site topology are induced upon inhibitor binding.

As with most proteases, the substrate-binding cleft can be subdivided into multiple subsites, each consisting of residues involved in recognition of the substrate. The catalytic residues are positioned between the S_1 and S_1' subsites, with subsite numbering mirroring the numbering of the corresponding substrate residues. The S1 subsite consists of a deep hydrophobic pocket that buries the side chain of the P1 leucine residue. Of the 12 residues that form this hydrophobic cleft, 7 are within Van der Waals bond distance (4.4 Å) of the P1 leucine: Leu3614 and Ala3615 are contributed by the β-flap (the structural region that mediates $InsP_6$ activation)[13]; Val3579 and Gly3580 are contributed by strand E; Val3530 and Gly3531 are contributed by strand D; and Ala3488 is contributed by helix 1 (**Figs. 2** and **3**). Due to the covalent bond between Cys3581 and the aza-epoxide, the inhibitor is pulled in slightly toward the catalytic cysteine; in a native conformation, the P1 leucine likely makes more diffuse contacts with both sides of the pocket.

C-terminal to the S1 pocket, the surface topology of the CPD is relatively flat and featureless, being composed primarily of peptide backbone atoms as well as the alkyl chain of Arg3534. On the N-terminal side, a groove is formed by helix 1 and the G1 strand of the β-flap, where the backbone atoms of the P2 and P3 inhibitor residues engage in hydrogen bond interactions with the backbone of the G1 strand (**Fig. 2b**). This analysis supports our prior observation that addition of P2 and P3 residues to the inhibitor scaffold increases potency irrespective of peptide sequence (**Table 1**). The P2 leucine residue of the inhibitor points away from the protease, tangentially interacting with Trp3631 and Glu3613, while the P3 leucine interacts with Val3484, the alkyl chain of Lys3487, Ala3488, Asn3491 and Val3616 in the S3 subsite (**Fig. 2b**). Based on our structure-activity analyses of the LLaL (JCP598) versus EAaL (VEA223) inhibitors (**Table 1**), which have similar AC_{50}(I) values, it seems likely that the surface chemistry of the S3 subsite can accommodate varied side chains at the P3 position. Taken together, the inhibitor structure provides mechanistic insight into substrate recognition and reveals how the CPD specifically recognizes a leucine in the P1 position.

Importantly, the aza-leucine epoxide inhibitor is found covalently bound to the catalytic cysteine through a thioether bond in the crystal structure. Nucleophlic attack by the catalytic cysteine occurs at the C3, rather than the C2, position of the epoxide; this same mechanism of catalysis is observed in the aza-peptide epoxide–bound structure of caspase-3 (ref. 22). Indeed, the catalytic cysteine and histidine residues of the CPD and caspase-3 are similarly positioned to attack the P1-P1′ peptide bond (**Fig. 2c**). Furthermore, both proteases have an optimal S1 pocket for recognition of the side chain of their respective P1 substrate residues. The CPD S1 pocket, however, is considerably deeper than the caspase-3 S1 pocket, causing the JCP598 inhibitor to be buried more deeply in the CPD structure than the aza-aspartate epoxide caspase inhibitor (**Fig. 2c**). These analyses reveal that the CPD

and caspases share similar mechanisms of catalysis and substrate binding, despite differing significantly in the size and surface properties of their S1 subsites.

The substrate specificity of MARTX CPDs is conserved

Clan CD proteases are distinguished by their strict preference for specific amino acid side chains in the S1 subsite[23]. Thus, by inference, one would predict that all MARTX CPDs are selective for leucine in the P1 position. Indeed, multiple sequence alignment of related MARTX CPDs indicates that substrate binding pocket residues are well conserved (**Fig. 3a**). To directly examine the substrate specificity of related CPDs, we expressed and purified MARTX CPDs from *Vibrio vulnificus* and *Photorhabdus luminescens*, which encodes four distinct MARTX CPDs, and analyzed their autoprocessing activity in the presence of $InsP_6$. All five MARTX CPDs tested underwent $InsP_6$-dependent autoprocessing (**Fig. 3b** and **Supplementary Fig. 4** online),

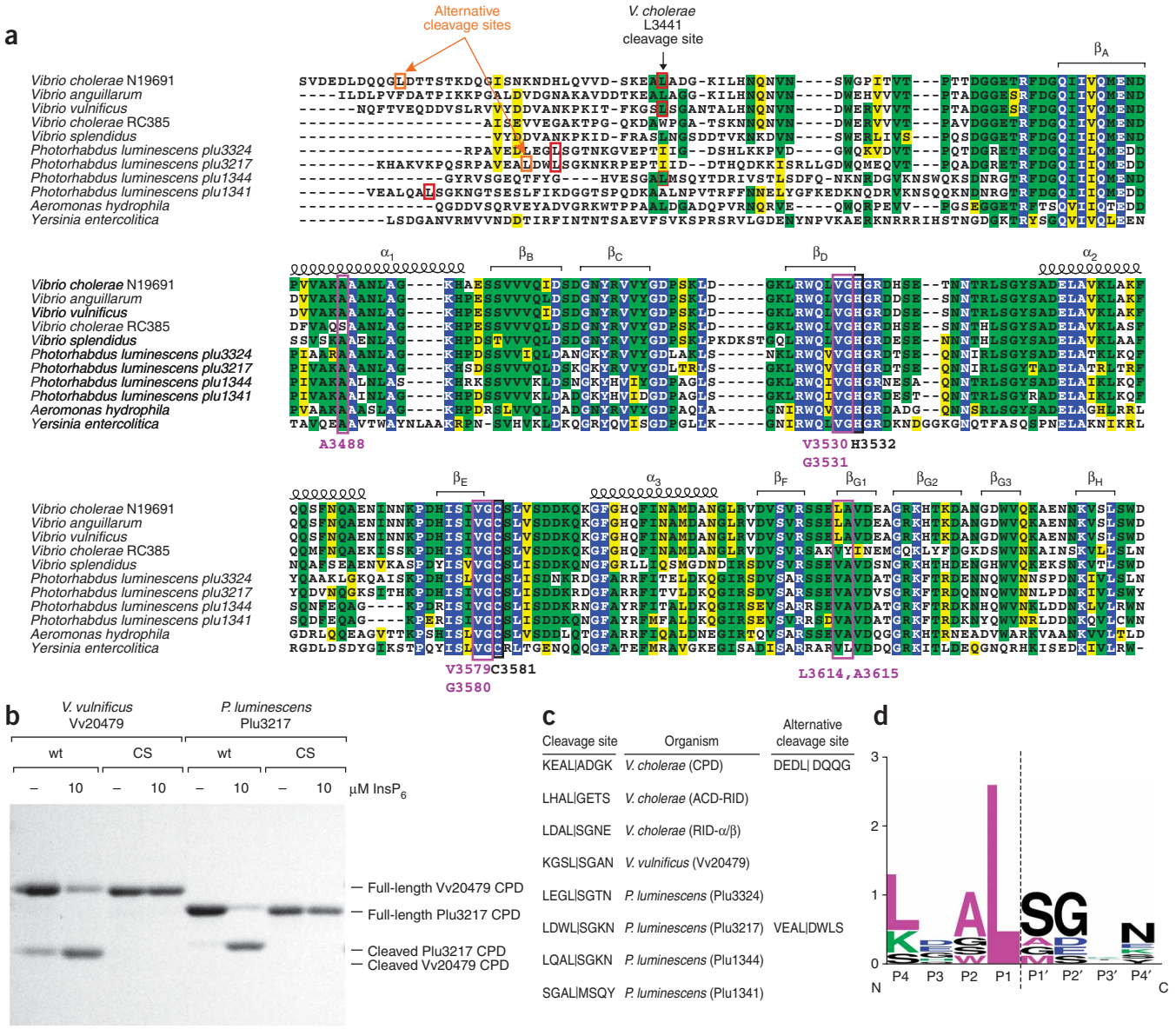

Figure 3 MARTX CPDs cleave after a P1 leucine. (**a**) Multiple sequence alignment of $MARTX_{Vc}$ CPD with homologs from related MARTX toxins. Completely conserved identical residues are blocked in blue, conserved identical residues in green and conserved similar residues in yellow. Secondary structure elements of the *V. cholerae* CPD (amino acids 3405–3647) are shown above the amino acid sequences. The catalytic cysteine and histidine residues are boxed and labeled in black; P1 leucine-interacting residues are boxed and labeled in purple. Cleavage site residues of MARTX CPDs from *V. cholerae* N16961, *V. vulnificus* and *P. luminescens* were identified by FT-MS analysis (red boxes). Alternative cleavage site residues that occur upon mutation of the P1 leucine to alanine were mapped by FT-MS (orange boxes). (**b**) Activation of MARTX CPD autocleavage by $InsP_6$. Wild-type or catalytic-dead (CS) recombinant CPDs from *V. vulnificus* CMCP6 VV20479 (amino acids 4044–4299) and *P. luminescens* TTO1 Plu3217 (amino acids 2390–2630) MARTX toxins were incubated in the presence or absence of 10 μM $InsP_6$, and autocleavage was assessed by SDS-PAGE and Coomassie staining. (**c**) Comparison of P4-P4′ residues of known MARTX CPD cleavage sites. Alternative cleavage sites that result from mutation of the primary autoprocessing site are shown for *V. cholerae* (Leu3441) and *P. luminescens* Plu3217 (Leu2408). (**d**) Sequence logo representation of MARTX CPD consensus cleavage site based on sites shown in **c**. The sequence logo was created using WebLogo. The dashed line indicates the scissile bond.

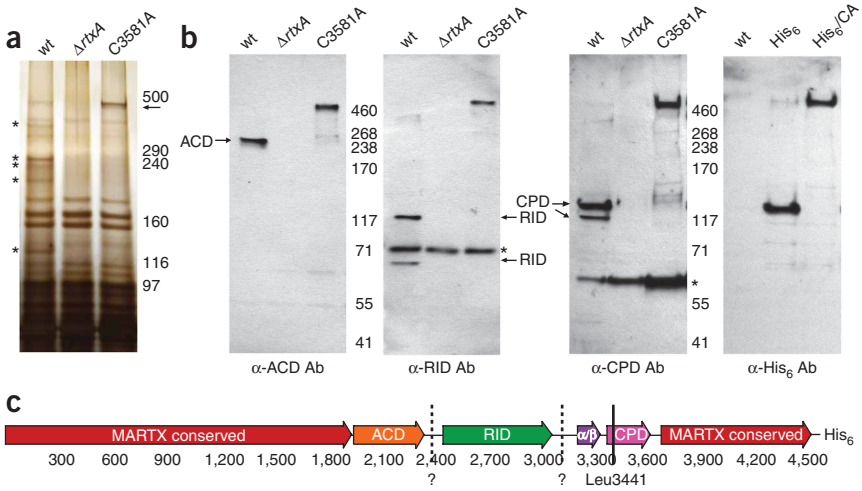

Figure 4 $MARTX_{Vc}$ is processed in a CPD-dependent manner. (**a**) Silver stain of culture supernatants harvested from *V. cholerae* strains harboring either an intact *rtxA* gene (wt), a null mutation in *rtxA* (Δ*rtxA*) or a point mutation in the CPD catalytic cysteine (C3581A). Unprocessed $MARTX_{Vc}$ (predicted size 460 kDa) is indicated with an arrow, while the asterisks demarcate $MARTX_{Vc}$-specific bands. (**b**) Western blot analysis of *V. cholerae* culture supernatants used in **a** and of *V. cholerae* culture supernatants derived from strains expressing $MARTX_{Vc}$ with a C-terminal His_6 tag in either the wt (His_6) or C3581A background (His_6/CA) (far right panel). Antibodies were raised against His_6-tagged ACD (amino acids 1964–2375), His_6-tagged RID (amino acids 2552–3099) and His_6-tagged CPD (amino acids 3391–3650). Background bands are indicated with asterisks. (**c**) Schematic of $MARTX_{Vc}$ toxin. Conserved glycine-rich repeat regions in the N and C termini of MARTX toxins (MARTX conserved, red); actin crosslinking domain (ACD, orange); Rho-inactivating domain (RID, green); α/β hydrolase domain (α/β, purple); cysteine protease domain (CPD, pink). Amino acid numbering is given below. The C-terminal His_6 tag encoded in strains His_6 and His_6/CA is shown. The known cleavage site within the CPD at Leu3441 is shown as a solid line. Putative cleavage sites inferred from western blot analyses are shown as dashed lines and labeled with question marks.

which indicates that MARTX CPDs exhibit a shared mechanism of activation. Analysis of the exact mass of the *in vitro* cleavage products by Fourier transform mass spectroscopy (FT-MS) revealed that all five MARTX CPDs were autoprocessed after leucine (**Fig. 3a**,**c**). When the P1 leucine residues of *V. cholerae* N16961 and *Photorhabdus luminescens* Plu3217 MARTX CPD were mutated to alanine, autocleavage occurred at a previously disfavored upstream leucine residue and before a P1′ aspartate (**Fig. 3a**,**c**). Notably, the CPD cleavage site mutant of *V. cholerae* did not cleave at the first available leucine residue (Leu3432), potentially due to the presence of a more bulky glutamine residue in the P1′ position.

To generate a consensus cleavage site sequence for MARTX CPDs, we analyzed the identified cleavage sites using WebLogo (http://weblogo.berkeley.edu/) (**Fig. 3d**). Although the training set is limited, these analyses suggest that, in addition to the strict requirement for leucine in the P1 position, small residues may be preferred in the P2, P1′ and P2′ positions. A small P2 residue, however, is not essential, given that substrates with leucine and tryptophan in the P2 position were still recognized by the CPD (**Figs. 1c** and **3c**). No conservation was observed in the P3 position, thus confirming that the P3 position contributes little to substrate specificity. Taken together, these analyses demonstrate that MARTX CPDs, like other clan CD proteases, are highly selective for the P1 residue; in the case of MARTX CPDs, the P1 residue recognized is a leucine.

The CPD processes $MARTX_{Vc}$ toxin at multiple sites

The universal conservation of CPDs in MARTX toxins suggests that these proteases play a critical role in regulating toxin function. To interrogate the role of the CPD in activating MARTX toxin, we examined the number and location of CPD-dependent processing sites within *V. cholerae* MARTX toxin. We first compared the secreted protein profiles of wild-type *V. cholerae*, an *rtxA* deletion strain (Δ*rtxA*; *rtxA* encodes $MARTX_{Vc}$) and a CPD catalytically dead strain (C3581A) to assess whether $MARTX_{Vc}$ is cleaved at multiple sites in a CPD-dependent manner. Whereas multiple $MARTX_{Vc}$-specific protein bands were detected in culture supernatants of wild-type *V. cholerae* (**Fig. 4a**, asterisks), a single, predominant MARTX-specific protein the predicted size of unprocessed $MARTX_{Vc}$ toxin (~460 kDa) was detected in culture supernatants of the C3581A mutant (**Fig. 4a**, arrow).

We next analyzed the domain composition of $MARTX_{Vc}$ cleavage products in culture supernatants by western blot analysis using antibodies specific for the ACD, RID and CPD domains of $MARTX_{Vc}$, respectively. Each antibody produced a distinct western blot profile in wild-type culture supernatants: the anti-ACD antibody primarily detected a single protein fragment (~250 kDa); the anti-RID antibody detected two $MARTX_{Vc}$ fragments (~110 kDa and ~65 kDa); and the anti-CPD antibody recognized two $MARTX_{Vc}$ fragments (~120 kDa and ~110 kDa) (**Fig. 4b**). The ~110 kDa protein fragment was detected by both the anti-CPD and anti-RID specific antibodies, which indicates that it harbors portions of both domains. The ~120 kDa fragment recognized by the anti-CPD antibody was also detected by a His_6-specific antibody in culture supernatants of a *V. cholerae* strain harboring a $MARTX_{Vc}$ with a C-terminal His_6 tag. This result indicates that the ~120 kDa protein contains the extreme C terminus of the CPD (**Fig. 4b**). In contrast, all four antibodies recognized a single ~460 kDa protein in the C3581A culture supernatant (**Fig. 4b**), which is consistent with the silver stain analysis (**Fig. 4a**). From these analyses, we can infer that, in addition to the previously mapped Leu3441 site within the CPD, two additional CPD-dependent processing sites are present in $MARTX_{Vc}$ between the ACD-RID and RID-α/β junctions (**Fig. 4c**).

Identification of $MARTX_{Vc}$ toxin cleavage sites

$MARTX_{vc}$ processing may result from direct cleavage by the CPD or from CPD-mediated activation of a second protease that sequentially cleaves $MARTX_{Vc}$; this latter scenario is frequently observed in viral polyprotein processing systems[24,25]. To distinguish between these possibilities, we tested whether $InsP_6$-activated CPD could cleave $MARTX_{Vc}$-derived polypeptides *in vitro*. Transcleavage of recombinant ACD-α/β, ACD-RID and RID-pGap1 polypeptides by the CPD all produced an ~71 kDa fragment; CPD-mediated cleavage of the ACD-α/β and ACD-RID polypeptides produced an ~55 kDa fragment; and transcleavage of the ACD-α/β and RID-α/β fragments liberated an ~36 kDa protein (**Fig. 5a**). By deduction, the ~71 kDa, ~55 kDa and ~36 kDa fragments comprise the ACD, RID and α/β

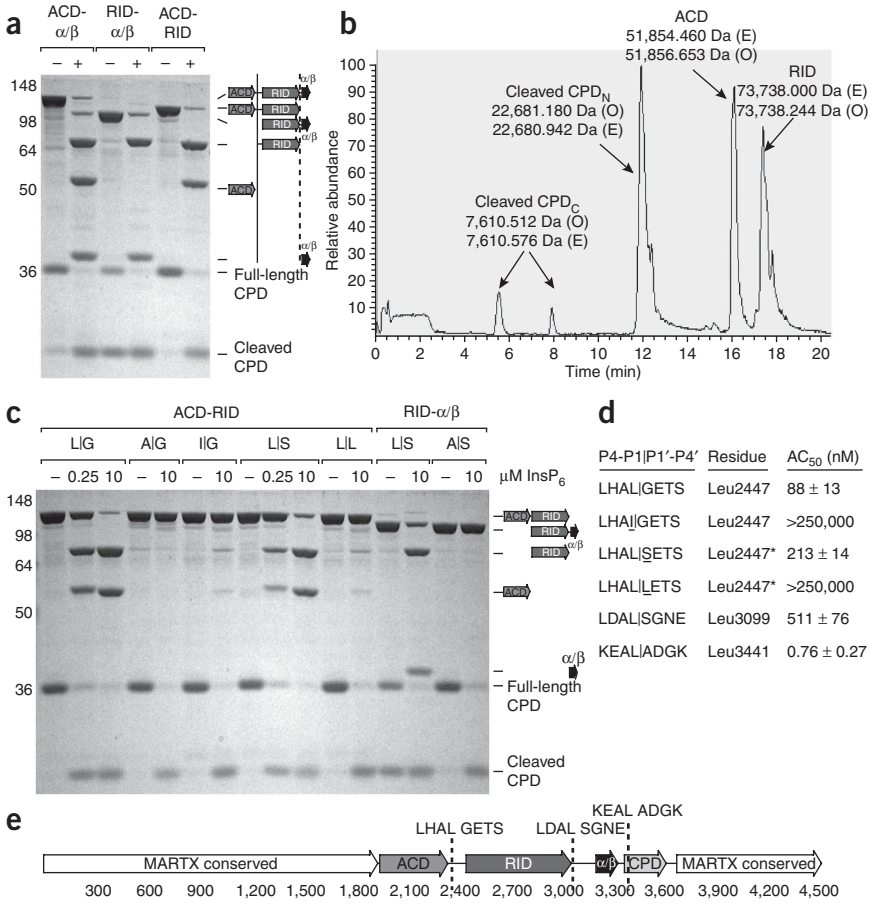

Figure 5 Identification of MARTX$_{Vc}$ toxin cleavage sites *in vitro*. (**a**) CPD-mediated transcleavage of MARTX$_{Vc}$ polypeptides *in vitro*. Recombinant CPD (amino acids 3391–3650) and the indicated MARTX$_{Vc}$ polypeptides were incubated ± InsP$_6$, and cleavage reactions were resolved by SDS-PAGE and visualized by SDS-PAGE (schematic shown to the right). (**b**) Reverse-phase HPLC chromatogram of CPD-mediated transcleavage of recombinant ACD-RID. The observed masses (O) of the polypeptides detected within a given peak are indicated and were determined by FT-MS analysis; the expected masses (E) of polypeptide fragments are also shown. (**c**) InsP$_6$-induced CPD-mediated transcleavage of mutant MARTX$_{Vc}$ polypeptides. The P1 and P1′ residues of the wild-type and mutant cleavage sites are given as P1|P1′. LIG is the wild-type sequence for the ACD-RID cleavage site (Leu2447); LIS is the wild-type sequence for the RID-α/β cleavage site (Leu3099). (**d**) Comparison of CPD-mediated cleavage at various processing sites. The P4 to P4′ residues are shown for each cleavage site. The concentration of InsP$_6$ at which 50% cleavage of the indicated polypeptides occurred (AC$_{50}$) is shown (± s.d.). In the presence of InsP$_6$, the ACD-RID and RID-α/β polypeptides were subjected to transcleavage with recombinant CPD, while the α/β-CPD polypeptide was subjected to autocleavage. Mutations introduced into the P1 and P1′ sites of the ACD-RID cleavage site are underlined. (**e**) Schematic of MARTX$_{Vc}$ toxin. Amino acid numbering is given below, and cleavage site sequences (P4-P4′) are given.

domains, respectively, which indicates that the CPD directly cleaves MARTX$_{Vc}$ between the ACD and RID domains and between the RID and α/β hydrolase domains.

In order to map specific CPD cleavage sites within recombinant MARTX$_{Vc}$ fragments, we measured the exact mass of *in vitro* transcleavage products by FT-MS. While these analyses confirmed the Leu3441 CPD autoprocessing site (**Fig. 5b**)[12], the resolution of the mass spectrometer for the larger fragments was insufficient to unequivocally identify the cleavage sites. However, given that CPD-mediated processing of MARTX$_{Vc}$ likely occurs after leucine, we were able to identify Leu2447 and Leu3099 as putative cleavage sites. To validate these sites, the effect on CPD-mediated transcleavage upon mutation of Leu2447 and Leu3099 to alanine was examined. Mutation of both residues to alanine abrogated processing of recombinant ACD-RID and RID-α/β polypeptides, respectively (**Fig. 5c**), while mutation of Leu2447 to its isomer isoleucine severely reduced CPD transcleavage of the ACD-RID polypeptide (**Fig. 5c**). These data confirm our predicted cleavage sites and support the conclusion that the CPD requires a P1 leucine residue for substrate recognition.

We next examined whether the CPD exhibited differential affinity toward MARTX$_{Vc}$ processing sites. To this end, we measured the concentration of InsP$_6$ required to half-maximally activate CPD cleavage (AC$_{50}$) at Leu2447, Leu3099 and Leu3441 (**Fig. 5d** and **Supplementary Fig. 5** online). The AC$_{50}$ for the α/β-CPD junction (Leu3441) was 0.76 ± 0.27 nM, which is consistent with the previously measured AC$_{50}$ for a recombinant CPD fragment lacking the N-terminal α/β domain (0.91 ± 0.10 nM)[13]. The AC$_{50}$ for the ACD-RID junction (Leu2447) was 88 ± 13 nM, and the AC$_{50}$ for the RID-α/β junction (Leu3099) was 511 ± 77 nM (**Fig. 5d**). These results demonstrate that the CPD recognizes MARTX$_{Vc}$ cleavage sites with differential affinity. The substantially lower AC$_{50}$ for Leu3441 likely reflects the positioning of Leu3441 close to the active site[13] such that the primary sequence around the Leu3441 cleavage site should not affect the CPD's affinity for this site. In contrast, slight variations in the primary sequence around Leu2447 and Leu3099 could account for the ~sixfold difference in AC$_{50}$ for these sites. To explore this possibility, we exchanged the P1′ glycine (Gly2448) of the Leu2447 cut site for the P1′ serine (Ser3100) of the Leu3099 cleavage site. This alteration increased the AC$_{50}$ of the Leu2447 cut site by ~2.5-fold (**Fig. 5d**), which suggests that the CPD prefers small, neutral residues to polar residues in the P1′ position. Large residues in the P1′ position are poorly tolerated by MARTX$_{Vc}$ CPD, as mutation of Gly2448 to leucine largely abrogated CPD-mediated cleavage at Leu2447 (**Fig. 5c**). These substrate preferences are consistent with the observation that the S1′ subsite is flat and nonpolar in the inhibitor-bound crystal structure (**Fig. 2a**).

CPD-mediated processing optimally activates MARTX$_{Vc}$ function

Having identified multiple MARTX$_{Vc}$ processing sites *in vitro*, we sought to determine whether these cleavage sites were relevant *in vivo*. Thus, we introduced mutations of Leu2447, Leu3099 and Leu3441 to alanine either singly, doubly or triply into the genome of *V. cholerae* and assessed their effect on MARTX$_{Vc}$ processing in culture supernatants by western blot analysis. Cleavage of MARTX$_{Vc}$ at these three sites theoretically should liberate eight polypeptides that can be detected by the anti-ACD, anti-RID and anti-CPD antibodies

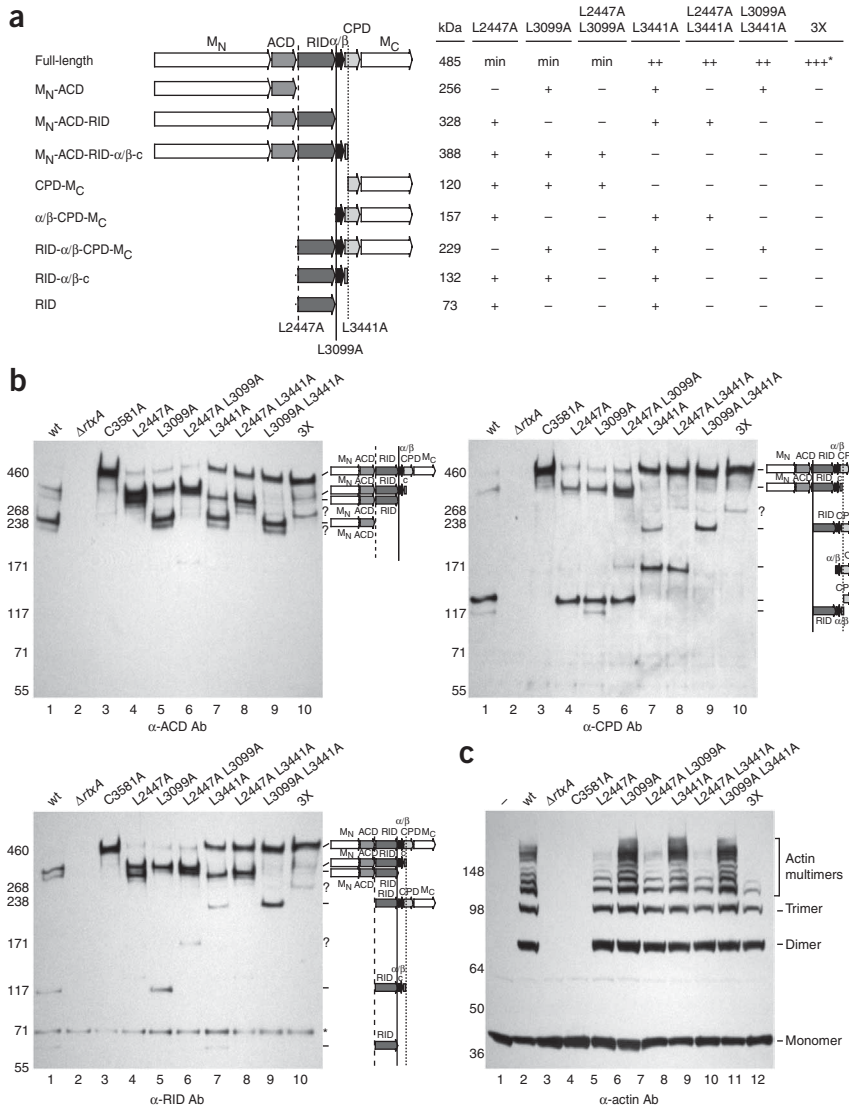

Figure 6 Effect of cleavage site mutations on $MARTX_{Vc}$ processing and function. (**a**) Summary of western blot analyses of MARTX toxin in *V. cholerae* cleavage site mutant culture supernatants. M_N, N-terminal MARTX conserved region; M_C, C-terminal MARTX conserved region. The predicted MWs of $MARTX_{Vc}$ fragments are given. + indicates the presence of a given polypeptide band in culture supernatants by western blot analysis; ++ indicates increased levels of full-length $MARTX_{Vc}$ in mutant culture supernatants relative to the minimal amounts observed in wild-type supernatants (min). The triple mutant L2447A L3099A L3441A is designated as 3X. The majority of $MARTX_{Vc}$ secreted by the 3X mutant is unprocessed, although small amounts of aberrantly processed toxin are observed (+++*). (**b**) Western blot analysis of *V. cholerae* cleavage site mutant culture supernatants using antibodies specific for discrete regions of $MARTX_{Vc}$. Unidentifiable bands are noted with a question mark; background bands are indicated with an asterisk. Cleavage sites that affect detection of $MARTX_{Vc}$ fragments for a given antibody are shown. (**c**) Actin crosslinking activity of *V. cholerae* cleavage site mutants. Culture supernatants harvested from strains used in **b** were incubated with HFF cells for 2 h; HFFs were lysed, and lysates were resolved by SDS-PAGE. Actin crosslinking was visualized by western blotting using an anti-actin antibody. The crosslinked forms of actin are labeled to the right.

(**Figs. 5** and **6a**). Indeed, all eight fragments were detected in wild-type culture supernatants (**Fig. 6a**,**b**). Mutation of cleavage site leucine residues to alanine prevented cleavage at these sites (**Fig. 6a**,**b**). For example, Leu2447A resulted in the disappearance of two polypeptides containing either a C-terminal ACD domain or N-terminal RID domain (**Fig. 6a**,**b**, lane 4). Likewise, culture supernatants of the L3099A mutant lacked $MARTX_{Vc}$ fragments with either a C-terminal RID domain or an N-terminal α/β domain (**Fig. 6a**,**b**, lane 5). Conversely, mutation of Leu3441 caused the loss of fragments with either a C-terminal α/β domain or N-terminal CPD domain from culture supernatants (**Fig. 6a**,**b**, lane 7), as well as the accumulation of unprocessed $MARTX_{Vc}$ relative to the L2447A and L3099A mutants. In fact, the Leu3441 mutation was epistatic to other cleavage site mutations: unprocessed $MARTX_{Vc}$ was the most prominent species detected in culture supernatants of any strain carrying the Leu3441 mutation (**Fig. 6b**). This result suggests that processing at Leu3441 stimulates the transcleavage activity of the CPD protease. Mutation of all three cleavage sites rendered $MARTX_{Vc}$ largely resistant to CPD-mediated processing (**Fig. 6a**,**b**, lane 10), since only a small amount of processing at an alternative site was observed (**Fig. 6a**,**b**, question marks). Given that these alternative cleavages were highly inefficient, these analyses indicate that the primary $MARTX_{Vc}$ cleavage sites *in vivo* are Leu2447, Leu3099 and Leu3441.

To assess the role of $MARTX_{Vc}$ processing on toxin activation, we examined whether $MARTX_{Vc}$ cleavage site mutants exhibited reduced actin crosslinking in human foreskin fibroblast cells (HFFs). *V. cholerae* culture supernatants harvested from wild-type, Δ*rtxA*, C3581A (CPD catalytic mutant) and cleavage site mutants were incubated with HFFs, and western blot analysis was used to visualize $MARTX_{Vc}$-induced actin crosslinking in HFF lysates. Although all supernatants from cleavage site mutants induced actin crosslinking (**Fig. 6c**), supernatants from strains carrying the L2447A mutation exhibited lower amounts of actin crosslinking relative to wild type, with the triple mutant being the most attenuated. These results suggest that optimal ACD enzymatic function requires processing between the ACD-RID junction (Leu2447), although a single cleavage of $MARTX_{Vc}$ can activate the ACD.

Chemical inhibition of $MARTX_{Vc}$ toxin function

Lastly, we examined whether our small-molecule CPD inhibitors could effectively disrupt $MARTX_{Vc}$ processing *in vivo*. To this end, we grew wild-type *V. cholerae* cultures in the presence of increasing concentrations of CPD inhibitors and then measured $MARTX_{Vc}$ processing in culture supernatants by western blot analysis. Treatment of cultures with each of the three most potent inhibitors (AS01, VEA223 and JCP598) resulted in the accumulation of unprocessed $MARTX_{Vc}$ toxin relative to untreated wild-type culture supernatants (**Fig. 7a**). Although the compounds only partially blocked processing

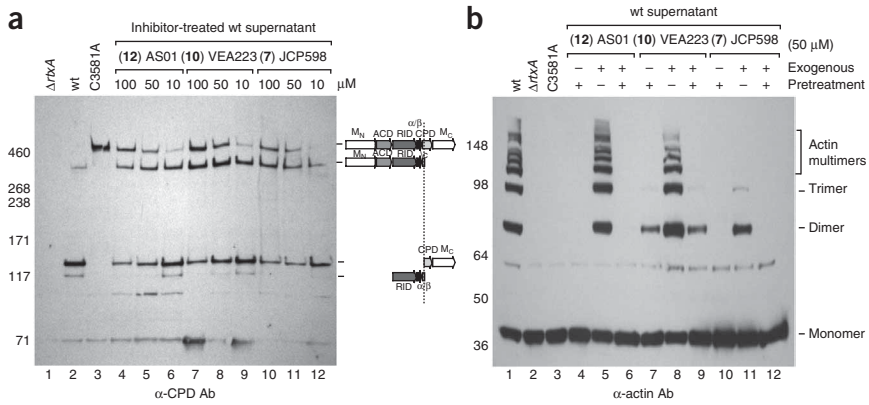

Figure 7 Chemical inhibition of $MARTX_{Vc}$ processing and toxin function. (**a**) Dose-dependent reduction in $MARTX_{Vc}$ processing by CPD inhibitors. Western blot analysis of culture supernatants harvested from wild-type cultures grown in the presence or absence of inhibitor until mid-log phase using a CPD-specific antibody. (**b**) Effect of CPD inhibitors on $MARTX_{Vc}$ actin crosslinking in HFFs. Pretreatment indicates that *V. cholerae* culture supernatants were pretreated with 50 μM of inhibitor for 15 min. For exogenous treatment, the inhibitor was added at 50 μM to HFFs (in DMEM medium) before addition of *V. cholerae* culture supernatants. HFF cells were exposed to *V. cholerae* culture supernatants to stimulate actin crosslinking, after which HFFs were lysed, and lysates were resolved by SDS-PAGE. Western blot analysis using an anti-actin antibody was used to visualize crosslinked actin species, which are indicated to the right.

at Leu3441 even at the highest concentration of 100 μM, they completely inhibited processing at Leu3099 at concentrations greater than 50 μM. This latter result is consistent with the *in vitro* observation that the CPD exhibits reduced affinity for the Leu3099 cleavage site relative to the Leu3441 autoprocessing site (**Fig. 5d**).

The failure of CPD inhibitors to completely prevent $MARTX_{Vc}$ processing during growth of *V. cholerae* in LB medium could reflect their relative instability in these growth conditions. Thus, we evaluated the ability of CPD inhibitors to prevent $MARTX_{Vc}$ toxin activation in host cells. All three compounds, AS01, VEA223 and JCP598, completely blocked the actin crosslinking activity of wild-type *V. cholerae* supernatants when added at concentrations of 50 μM (**Fig. 7b**, lanes 4, 7 and 10). In contrast, addition of AS01 and VEA223 exogenously to the medium of HFF cells immediately before adding untreated wild-type culture supernatants resulted in little inhibition of toxin function (**Fig. 7b**, lanes 5 and 8). Conversely, addition of JCP598 to the medium of HFF cells substantially reduced $MARTX_{Vc}$-induced actin crosslinking (**Fig. 7b**, lane 11). We suspect that the difference in inhibitor potencies can be attributed to differences in membrane permeability of the inhibitors. The negative charges on AS01 and VEA223 likely reduce their ability to cross host cell membranes, whereas the hydrophobic JCP598 readily passes through cell membranes to inhibit $MARTX_{Vc}$ function even after toxin translocation. Taken together, our results validate the CPD as a target for small molecules designed to block MARTX toxin activation.

DISCUSSION

The MARTX toxin of *V. cholerae* is autoproteolytically activated by an internal cysteine protease domain. Here, we demonstrated that $MARTX_{Vc}$ CPD is a leucine-specific protease whose activity can be chemically inhibited to prevent $MARTX_{Vc}$ activation. We further determined that $MARTX_{Vc}$ processing at interdomain regions optimally activates effector domain function, since the actin crosslinking activity of the ACD was most efficient when cleavage occurred between the ACD and RID domains (Leu2447, **Fig. 6c**), and optimal CPD activity depended upon processing at the α/β-CPD junction (Leu3441, **Fig. 6b**). Based on the results of our cleavage site mapping, we propose the following model. Processing of $MARTX_{Vc}$ by the CPD at Leu2447, Leu3099 and Leu3441 liberates the RID and α/β hydrolase domains, respectively. This cleavage profile leaves the ACD and CPD domains tethered to the membrane-bound N- and C-terminal MARTX conserved regions, respectively (**Supplementary Fig. 6** online). Membrane localization of the CPD likely ensures that the protease can efficiently access its transcleavage substrates and may account for conservation in position of the CPD in all MARTX toxins[1]. Furthermore, processing at Leu3441 may additionally be required for the CPD to optimally bind its transcleavage substrates, while processing at Leu2447 may enhance ACD activity by liberating its C-terminal end.

Our study identified the first chemical inhibitors (to our knowledge) of $MARTX_{Vc}$ CPD (**Fig. 1** and **Table 1**) and demonstrated their utility in preventing MARTX toxin activation (**Fig. 7**). These inhibitors appear to be selective, since the most potent inhibitor in our assays, JCP598, exhibits little reactivity against a wide variety of proteases *in vitro*[21]. Furthermore, these inhibitors may prevent the proteolytic activation of other MARTX toxins, since related MARTX CPDs were leucine-specific and $InsP_6$-inducible (**Fig. 3**). Blocking CPD function likely represents the most effective strategy for preventing host cell intoxication by MARTX family members, which are multifunctional and heterogeneous in composition[1]. Additionally, the inhibitors identified in our study may block the cysteine protease activity of *Clostridium* sp. cytotoxins[14–16], since the CPD of *C. difficile* toxin B, the primary virulence factor of this nosocomial pathogen, exhibits a similar substrate specificity[16].

Combined with biochemical and structural studies, the inhibitor analyses revealed that the CPD exhibits exquisite selectivity for leucine in the P1 position. Only clan CD–specific compounds with a P1 leucine had inhibitory activity (**Fig. 1**); all MARTX CPDs examined cleaved exclusively after a P1 leucine (**Figs. 3** and **6**), and mutation of the P1 leucine to isoleucine abrogated CPD-mediated transcleavage (**Fig. 5c**). These mutational studies further indicated that the P1′ residue directs the substrate specificity of the CPD. All known MARTX CPD cleavage sites contain neutral, small residues in the P1′ position (**Fig. 3**), and mutation of this small P1′ residue to the bulkier leucine residue abrogated CPD-mediated processing (**Fig. 5c**), which indicates that the S1′ subsite likely does not tolerate large residues. Charged residues in the P1′ position also appeared to be disfavored, since cleavage between a P1 leucine (Leu3415) and P1′ aspartate (Asp3416) was only observed when Leu3441 of $MARTX_{Vc}$ CPD was mutated to an alanine (**Fig. 5c**). Lastly, the P3 and P2 positions were observed to contribute little to substrate selectivity. The P3 and P2 positions are poorly conserved among known $MARTX_{Vc}$ cleavage sites (**Fig. 3c**); VEA223 and JCP598 exhibited similar inhibitor potencies despite differing only in these positions (**Fig. 1** and **Table 1**), and the P2 and P3 residues minimally interact with the CPD subsites in the crystal structure (**Fig. 2b**).

The crystal structure of activated CPD bound to an aza-leucine epoxide inhibitor reveals, at a molecular level, the constituents determining CPD substrate specificity. A deep, hydrophobic S1 pocket perfectly accommodates the P1 leucine of the inhibitor exclusively (**Fig. 2**). In contrast, the S1′ subsite consists mainly of a flat nonpolar surface (**Fig. 2a**). The lack of recognition features in the S1′ region may explain the CPD's preference for small residues in the P1′ position, with glycine favored over serine (**Fig. 3c**), and its inability to accommodate a leucine residue in the P1′ position (**Fig. 5b**). Notably, caspases have been shown to exhibit a similar preference for the P1′ position, favoring glycine over serine and alanine[26,27].

This observation is one of many similarities that $MARTX_{Vc}$ CPDs and caspases share in substrate recognition. Like the caspases, the CPD is sensitive to inhibition by both aza-epoxides and AOMKs (**Table 1**). Furthermore, $MARTX_{Vc}$ CPD and caspase-3 exhibit the same stereoisomer preference around the epoxide group[21] and react with the epoxide in a similar manner[22]. The most striking example of the similarity between caspases and $MARTX_{Vc}$ CPDs, however, is the observation that the active site topologies of the CPD and caspase-3 are nearly identical (**Fig. 2c**). Despite their weak overall structural similarity and disparate mechanisms of activation[13], the catalytic residues are well aligned in a superposition of the central D and E β-strands (**Fig. 2c**). Furthermore, the S1 pocket of caspase-3 and $MARTX_{Vc}$ CPD occupies a similar position between the catalytic residues and is the primary substrate specificity determinant.

Notably, the distantly related clan CD protease gingipain-R shares a similar active site geometry and positioning of the S1 pocket with the caspases and $MARTX_{Vc}$ CPD (**Supplementary Fig. 7** online). This observation suggests that the mechanism of substrate recognition is broadly conserved among clan CD proteases. Specifically, all three proteases exhibit strict specificity for the P1 residue: $MARTX_{Vc}$ CPD, caspases and gingipain-R recognize hydrophobic leucine, acidic aspartate[23] and basic arginine[28], respectively and exclusively. Accordingly, the molecular surfaces around the active sites are highly evolved to recognize their respective P1 residue: the substrate binding cleft of the CPD, caspases and gingipain-R are neutral, basic and acidic, respectively. The S1 binding cleft is so selective that isoleucine fails to functionally substitute for a P1 leucine in the CPD (**Fig. 5c**), and glutamate fails to substitute for a P1 aspartate in caspases[27]. Thus, MARTX CPDs, caspases and gingipains appear to have evolved from a common structural scaffold; from this scaffold, the proteases have evolved distinct substrate recognition preferences and mechanisms of activation. These studies raise the possibility that residues in the S1 subsite of a given clan CD protease might be altered to engineer new substrate specificities for these highly specific enzymes.

METHODS

Bacterial and eukaryotic cell growth conditions and strain construction. For details see **Supplementary Methods** online.

Screen for inhibitors of $MARTX_{Vc}$ CPD autoprocessing. Autocleavage assays were performed in 50-μl volumes containing 1 μM N-terminally His_6-tagged $MARTX_{Vc}$ CPD (amino acids 3391–3650) in cleavage assay buffer (60 mM NaCl, 20 mM Tris pH 7.5, 250 mM sucrose) in 96-well plates. Inhibitors were added at a final concentration of 100 μM (1:100 final dilution from a 10 mM stock) and incubated with $MARTX_{Vc}$ CPD for 30 min at room temperature (22–25 °C). GTPγS (Sigma) was then added to give a final concentration of 200 μM (1:10 dilution). Cleavage reactions were incubated at 37 °C for 2 h, after which autocleavage was stopped by the addition of SDS-PAGE loading buffer. Samples were boiled for 3 min at 95 °C and resolved by SDS-PAGE on 15% gels. Cleavage reactions were visualized by Coomassie staining. The screen was performed in triplicate, and hits were confirmed in a secondary screen using the autocleavage assay.

AC_{50}:inhibitor ratios. Inhibitor potency was determined by measuring the concentration of $InsP_6$ required to induce half-maximal cleavage of $MARTX_{Vc}$ CPD in the presence of 10 μM inhibitor (AC_{50}(I)). 1 μM of recombinant CPD in 50 μl cleavage assay buffer was pretreated with 10 μM inhibitor (1:100 dilution) for 30 min at room temperature. $InsP_6$ (Calbiochem) was added at the indicated final concentrations (1:100 dilution), and autoprocessing was allowed to proceed for 1 h at 37 °C. Cleavage reactions were resolved by SDS-PAGE and visualized by Coomassie staining. Images were quantified using the publicly available program ImageJ (http://rsb.info.nih.gov/ij/) as previously described[13]. The amount of autocleaved protein relative to total protein was plotted versus concentration of $InsP_6$. The AC_{50}(I) was determined from these plots using the Michaelis-Menten function on KaleidaGraph (Synergy Software).

Synthesis of CPD-specific inhibitors. VEA223 was synthesized in solution phase using standard chemistries as described previously for the synthesis of JCP598 (ref. 21). AS01, AS02 and AS04 were generated using solid-phase synthesis as described previously[29].

***In vitro* CPD autocleavage and transcleavage assays.** Autocleavage assays were performed as described above for AC_{50}(I) determinations except that no inhibitor was used. Transcleavage assays were identical to autocleavage assays with the exception that recombinant $MARTX_{Vc}$ polypeptides were added to reactions at 1 μM. AC_{50} values were determined from triplicate assays as described above.

Protein expression and purification. Proteins for *in vitro* cleavage reactions and crystallization were purified as previously described[13].

Crystallization and data collection. For details see **Supplementary Methods**.

Structure determination and refinement. Initial phases were obtained by molecular replacement with PHASER[30], using the $MARTX_{Vc}$ CPD (PDB ID 3EEB) as a search model. The JCP598 inhibitor was constructed manually using COOT[31], and the structure was refined by iterative rounds of model adjustment followed by refinement with CNS[32]. The final model went through translation/libration/screw and restrained refinement with REFMAC5[33], resulting in final R and R_{free} values of 22.1% and 26.5%, respectively. Ramachandran analysis with MolProbity (http://molprobity.biochem.duke.edu/)[34] indicated that 95.0% of residues reside in the most favorable regions, with the remaining 5.0% in additionally allowed regions. Refinement statistics can be found in **Supplementary Table 1** online. Superposition of structures was performed with the program Superpose from the CCP4 program suite[35]. All structural figures were prepared with PyMOL[36]. The final model contains four copies of the $MARTX_{Vc}$ CPD in the asymmetric unit, each bound to one $InsP_6$ molecule, one sodium ion, and one JCP598 molecule. Chain A is used for all figures in the paper.

Silver staining of *V. cholerae* culture supernatants. *V. cholerae* culture supernatants were prepared as described previously[13] and resolved on a 3–8% Tris-acetate gel (Invitrogen). The gel was silver-stained using the SilverXpress Silver staining kit (Invitrogen).

Western blot analysis of MARTX toxin. Untreated *V. cholerae* supernatants were prepared and resolved as described previously[13]. For inhibitor-treated *V. cholerae* cultures, the indicated inhibitor was diluted 1:500 into 2 ml of LB medium containing a 1:1000 dilution of overnight wild-type *V. cholerae* culture. Diluted, inhibitor-treated cultures were grown until mid-log phase (OD_{600} ~ 0.5, ~2.25 h growth), and culture supernatants were trichloroacetic acid–precipitated and resolved as described[13]. Polyclonal MARTX-specific antibodies were raised against recombinant ACD (amino acids 1964–2375, CoCalico Biologicals), RID (amino acids 2552–3099, CoCalico Biologicals) and CPD[13], and western blot analyses were performed as previously described[13].

Fourier transform mass spectrometry. *In vitro* cleavage reactions were separated on a PLRPS 150 mm × 0.1 mm column (Varian, 5 μM particle size, 300 Å pore size) run at a flow rate of 700 nl min^{-1} in 0.1% trifluoroacetic

acid/water (A):0.1% trifluoroacetic acid/acetonitrile (B). The column was run on a gradient of 10% B to 60% B for 25 min, 60% B to 90% B for 2 min, held at 90% B for 2 min, then rapidly decreased to 10% B over 0.1 min, then run for 11 min at 10% B. Eluted samples were run on a Thermo LTQ-FT mass spectrometer (Thermo Fisher Scientific) using FTMS + p NSI full MS scanning mode, mass range 400.00–2000.00, FT resolution 100,000. The MW of eluted peptides was determined by deconvolution using Isopro 3.0 (MS/MS Software). The measured mass was compared to the predicted MW of possible peptide cleavage products determined from the primary sequence of recombinant polypeptides derived from $MARTX_{Vc}$ using ProtParam (http://ca.expasy.org/tools/protparam.html).

Actin crosslinking assay. For cleavage site mutant analyses, the actin crosslinking assay was performed as described previously[13] using culture supernatants. In the inhibitor actin crosslinking assays, 100 µl of *V. cholerae* culture supernatants were pretreated with the indicated inhibitor at a final concentration of 50 µM (1:200 dilution from a 10 mM stock). For exogenous addition, inhibitor was also added to HFFs (in 500 µl DMEM medium) at a final concentration of 50 µM (1:200 dilution from a 10 mM stock).

Accession codes. Protein Data Bank: Coordinates and structure factors have been deposited under accession number 3GCD. The $MARTX_{Vc}$ CPD was deposited as part of a previous study under accession number 3EEB; the D and E β-strands of caspase-3–aza-Asp epoxide were deposited as part of a previous study under accession number 2C1E.

Note: Supplementary information and chemical compound information is available on the Nature Chemical Biology website.

ACKNOWLEDGMENTS

We thank P. Gulig (University of Florida) for generously providing *Vibrio vulnificus* CMCP6 genomic DNA, E. Shank and R. Kolter (Harvard Medical School) for providing the *Photorhabdus luminescens* TTO1 strain, and M. Blokech and G. Schoolnik (Stanford School of Medicine) for help with *V. cholerae* strain construction and for providing *V. cholerae* genomic DNA. P.J.L. is supported by the Damon Runyon Cancer Research Foundation. K.C.G. is supported by the Keck Foundation and the Howard Hughes Medical Institute. M.B. is supported by the Burroughs Wellcome Foundation, the Searle Scholars Program, the US National Institutes of Health National Technology Center for Networks and Pathways (grant U54-RR020843) and the Human Frontier Science Program (grant RGP0024/2006-C).

AUTHOR CONTRIBUTIONS

The inhibitor screen, synthesis of AS01, AS02 and AS04, protein expression and purification, cleavage assays, FT-MS data analysis, *V. cholerae* strain construction, actin crosslinking assays, $MARTX_{Vc}$ silver staining and western blot analyses were performed by A.S. Crystallization of the MARTX CPD–$InsP_6$–JCP598 complex was performed by A.S. and P.J.L. P.J.L. collected the data, solved and analyzed the structure, and generated the figures of the inhibitor-bound CPD structure. J.C.P. provided the cysteine protease compound library. V.E.A. synthesized JCP598, AS01 and VEA223, guided A.S. in the synthesis of AS01, AS02 and AS04, and assessed the integrity of all compounds described in this paper. A.G. designed the conditions for running samples for FT-MS, ran the samples for FT-MS and provided advice in FT-MS analysis. Creative input and financial support for the project were provided by M.B. The manuscript was written by A.S. and M.B. with advice from P.J.L., V.E.A., K.C.G. and A.G.

Published online at http://www.nature.com/naturechemicalbiology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Satchell, K.J. MARTX, multifunctional autoprocessing repeats-in-toxin toxins. *Infect. Immun.* **75**, 5079–5084 (2007).
2. Li, L., Rock, J.L. & Nelson, D.R. Identification and characterization of a repeat-in-toxin gene cluster in *Vibrio anguillarum*. *Infect. Immun.* **76**, 2620–2632 (2008).
3. Lee, B.C. *et al. Vibrio vulnificus rtxE* is important for virulence, and its expression is induced by exposure to host cells. *Infect. Immun.* **76**, 1509–1517 (2008).
4. Lee, J.H. *et al.* Identification and characterization of the *Vibrio vulnificus rtxA* essential for cytotoxicity in vitro and virulence in mice. *J. Microbiol.* **45**, 146–152 (2007).
5. Liu, M., Alice, A.F., Naka, H. & Crosa, J.H. The HlyU protein is a positive regulator of *rtxA1*, a gene responsible for cytotoxicity and virulence in the human pathogen *Vibrio vulnificus*. *Infect. Immun.* **75**, 3282–3289 (2007).
6. Olivier, V., Haines III, G.K., Tan, Y. & Satchell, K.J. Hemolysin and the multifunctional autoprocessing RTX toxin are virulence factors during intestinal infection of mice with *Vibrio cholerae* El Tor O1 strains. *Infect. Immun.* **75**, 5035–5042 (2007).
7. Olivier, V., Salzman, N.H. & Satchell, K.J. Prolonged colonization of mice by *Vibrio cholerae* El Tor O1 depends on accessory toxins. *Infect. Immun.* **75**, 5043–5051 (2007).
8. Cordero, C.L., Sozhamannan, S. & Satchell, K.J. RTX toxin actin cross-linking activity in clinical and environmental isolates of *Vibrio cholerae*. *J. Clin. Microbiol.* **45**, 2289–2292 (2007).
9. Rahman, M.H. *et al.* Distribution of genes for virulence and ecological fitness among diverse *Vibrio cholerae* population in a cholera endemic area: tracking the evolution of pathogenic strains. *DNA Cell Biol.* **27**, 347–355 (2008).
10. Sheahan, K.L., Cordero, C.L. & Satchell, K.J. Identification of a domain within the multifunctional *Vibrio cholerae* RTX toxin that covalently cross-links actin. *Proc. Natl. Acad. Sci. USA* **101**, 9798–9803 (2004).
11. Sheahan, K.L. & Satchell, K.J. Inactivation of small Rho GTPases by the multifunctional RTX toxin from *Vibrio cholerae*. *Cell. Microbiol.* **9**, 1324–1335 (2007).
12. 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).
13. 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).
14. Egerer, M., Giesemann, T., Jank, T., Satchell, K.J. & Aktories, K. Auto-catalytic cleavage of *Clostridium difficile* toxins A and B depends on cysteine protease activity. *J. Biol. Chem.* **282**, 25314–25321 (2007).
15. Giesemann, T., Egerer, M., Jank, T. & Aktories, K. Processing of *Clostridium difficile* toxins. *J. Med. Microbiol.* **57**, 690–696 (2008).
16. Reineke, J. *et al.* Autocatalytic cleavage of *Clostridium difficile* toxin B. *Nature* **446**, 415–419 (2007).
17. Gordon, V.M. & Leppla, S.H. Proteolytic activation of bacterial toxins: role of bacterial and host cell proteases. *Infect. Immun.* **62**, 333–340 (1994).
18. Arastu-Kapur, S. *et al.* Identification of proteases that regulate erythrocyte rupture by the malaria parasite *Plasmodium falciparum*. *Nat. Chem. Biol.* **4**, 203–213 (2008).
19. Powers, J.C., Asgian, J.L., Ekici, O.D. & James, K.E. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem. Rev.* **102**, 4639–4750 (2002).
20. Prochazkova, K. & Satchell, K.J. Structure-function analysis of inositol hexakisphosphate-induced autoprocessing of the *Vibrio cholerae* multifunctional autoprocessing RTX toxin. *J. Biol. Chem.* **283**, 23656–23664 (2008).
21. Asgian, J.L. *et al.* Aza-peptide epoxides: a new class of inhibitors selective for clan CD cysteine proteases. *J. Med. Chem.* **45**, 4958–4960 (2002).
22. 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).
23. Barrett, A.J. & Rawlings, N.D. Evolutionary lines of cysteine peptidases. *Biol. Chem.* **382**, 727–733 (2001).
24. Bedard, K.M. & Semler, B.L. Regulation of picornavirus gene expression. *Microbes Infect.* **6**, 702–713 (2004).
25. Reed, K.E. & Rice, C.M. Overview of hepatitis C virus genome structure, polyprotein processing, and protein properties. *Curr. Top. Microbiol. Immunol.* **242**, 55–84 (2000).
26. Schilling, O. & Overall, C.M. Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. *Nat. Biotechnol.* **26**, 685–694 (2008).
27. Stennicke, H.R., Renatus, M., Meldal, M. & Salvesen, G.S. Internally quenched fluorescent peptide substrates disclose the subsite preferences of human caspases 1, 3, 6, 7 and 8. *Biochem. J.* **350**, 563–568 (2000).
28. 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).
29. Kato, D. *et al.* Activity-based probes that target diverse cysteine protease families. *Nat. Chem. Biol.* **1**, 33–38 (2005).
30. McCoy, A.J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
31. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
32. Brünger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
33. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
34. Davis, I.W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375–W383 (2007).
35. 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).
36. DeLano, W.L. *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, California, USA, 2002).