THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 39, pp. 26557–26568, September 25, 2009

Structural and Molecular Mechanism for Autoprocessing of MARTX Toxin of *Vibrio cholerae* at Multiple Sites*[S]

Received for publication, May 29, 2009, and in revised form, July 9, 2009 Published, JBC Papers in Press, July 20, 2009, DOI 10.1074/jbc.M109.025510

Katerina Prochazkova[‡1], Ludmilla A. Shuvalova[§¶], George Minasov[§¶], Zdeněk Voburka[‖], Wayne F. Anderson[§¶], and Karla J. F. Satchell[‡1,2]

From the [‡]Department of Microbiology-Immunology, [§]Center for Structural Genomics of Infectious Diseases, and [¶]Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611 and the [‖]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague 161-10 Praha 6, Czech Republic

The multifunctional autoprocessing repeats-in-toxin (MARTX) toxin of *Vibrio cholerae* causes destruction of the actin cytoskeleton by covalent cross-linking of actin and inactivation of Rho GTPases. The effector domains responsible for these activities are here shown to be independent proteins released from the large toxin by autoproteolysis catalyzed by an embedded cysteine protease domain (CPD). The CPD is activated upon binding inositol hexakisphosphate ($InsP_6$). In this study, we demonstrated that $InsP_6$ is not simply an allosteric cofactor, but rather binding of $InsP_6$ stabilized the CPD structure, facilitating formation of the enzyme-substrate complex. The 1.95-Å crystal structure of this $InsP_6$-bound unprocessed form of CPD was determined and revealed the scissile bond Leu^{3428}–Ala^{3429} captured in the catalytic site. Upon processing at this site, CPD was converted to a form with 500-fold reduced affinity for $InsP_6$, but was reactivated for high affinity binding of $InsP_6$ by cooperative binding of both a new substrate and $InsP_6$. Reactivation of CPD allowed cleavage of the MARTX toxin at other sites, specifically at leucine residues between the effector domains. Processed CPD also cleaved other proteins in *trans*, including the leucine-rich protein YopM, demonstrating that it is a promiscuous leucine-specific protease.

Multifunctional-autoprocessing repeats-in-toxin (MARTX)[3] toxins are a family of large bacterial protein toxins with conserved repeat regions at the N and C termini that are predicted to transfer effector domains located between the repeats across the eukaryotic cell plasma membrane (1). The best characterized MARTX is the >450-kDa secreted virulence-associated MARTX of *Vibrio cholerae*. This toxin causes disassembly of the actin cytoskeleton and enhances *V. cholerae* colonization of the small intestine, possibly by facilitating evasion of phagocytic cells (2, 3). The central region of the *V. cholerae* MARTX toxin contains four discrete domains: the actin cross-linking domain (ACD) that introduces lysine-glutamate cross-links between actin protomers (4, 5), the Rho-inactivating domain (RID) that disables small Rho GTPases (6), an $\alpha\beta$ hydrolase of unknown function (1), and an autoprocessing cysteine protease domain (CPD) (7, 8).

The CPD is a 25-kDa domain found in all MARTX toxins located just before the start of the C-terminal repeats (7, 8). This domain is activated for autoproteolysis upon binding inositol hexakisphosphate ($InsP_6$) (7), a molecule ubiquitously present in eukaryotic cell cytosol (9–11), but absent in extracellular spaces and bacteria. Thus, autocatalytic processing would not occur until after translocation of the CPD and effector domains is completed. In the context of the holotoxin, catalytic residue Cys^{3568} was found to be essential for the toxin to induce efficient actin cross-linking by the ACD and Rho inactivation by the RID, demonstrating that autoprocessing is essential for MARTX to induce cell rounding (8).

While it is clear that $InsP_6$ activates the CPD and that autoprocessing is essential for MARTX function (7), the mechanism by which $InsP_6$ activates CPD is not well understood. Furthermore, only one processing site at Leu^{3428}–Ala^{3429} has been identified, although multiple processing events would be required to release each effector independently. In fact, after autoprocessing at Leu^{3428}–Ala^{3429}, CPD is reported to adopt a conformation with reduced affinity for $InsP_6$ (7), raising questions as to how the protease might process MARTX at other sites.

We present here the structure of the pre-processed form of the *V. cholerae* MARTX CPD bound to $InsP_6$. Our results dem-

* This work was supported, in whole or in part, by National Institutes of Health Grant RO1AI051490 (to K. J. F. S.) from NIAID. This work was also supported by the Dept. of Health and Human Services under Contract HHSN272200700058C (to W. F. A.). Use of the Advanced Photon Source was supported by the U.S. Dept. of Energy, Office of Science, Office of Basic Energy Sciences, under Contract DE-AC02-06CH11357. Use of the Life Science Collaborative Access Team Sector 21 was supported in part by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor Grant 085P1000817. FT-MS services were provided by the Chicago Biomedical Consortium-University of Illinois, Chicago Research Resources Center, which was established by funds from The Searle Funds at the Chicago Community Trust to the Chicago Biomedical Consortium.

The atomic coordinates and structure factors (code 3FZY) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Table 1 and Figs. 1–6.

[1] Supported in part by a development project from National Institute of Health Grant U54 AI057153 to Great Lakes (Region V) Regional Center for Excellence in Biodefense and Emerging Infectious Diseases Research.

[2] Supported in part by a Burroughs Wellcome Fund Investigators in Pathogenesis of Infectious Disease Award. To whom correspondence should be addressed: 320 E. Superior Ave., Searle 3–525, MC S213, Chicago, IL 60611. Tel.: 312-503-2162; Fax: 312-503-1339; E-mail: k-satchell@northwestern.edu.

[3] The abbreviations used are: MARTX, multifunctional-autoprocessing repeats in toxin; CPD, cysteine protease domain; ACD, actin cross-linking domain; RID, RhoGTPase inactivation domain; NEM, *N*-ethylmaleimide; CK, chloromethyl ketone; TEV, tobacco etch virus; ITC, isothermal titration calorimetry; FT-MS, Fourier transform mass spectrometry; $InsP_6$, inositol hexakisphosphate; rCPD, recombinant CPD.

onstrate that autoprocessing is activated by rearrangement of a β-hairpin loop upon $InsP_6$ binding that locks the N terminus of the CPD in the active site, facilitating hydrolysis of the Leu^{3428}–Ala^{3429} peptide bond. After autoprocessing, CPD adopts a post-processing form that has poor affinity for $InsP_6$ and thus must be cooperatively reactivated for high affinity binding of $InsP_6$ by association of a new substrate. As a consequence, we are able to demonstrate how CPD cleaves MARTX toxin between effector domains and releases them from the large toxin resulting in increased catalytic activity of the effectors.

EXPERIMENTAL PROCEDURES

Reagents and Supplies—Chromatography columns for the ÄKTA Purifier FPLC system were purchased from Amersham Biosciences; common reagents from Fisher Biotech; chloromethyl ketone (CK) inhibitors, $InsP_6$ (phytic acid), and G-actin from Sigma; *N*-ethylmaleimide (NEM) from Pierce; beef pancreas trypsin (3× crystallized) and chicken egg white trypsin inhibitor from MP Biomedical; and SYPRO® Orange dye from Invitrogen. Enzymes for recombinant DNA experiments were from New England Biolabs and Invitrogen. Crystallization trays were purchased from Fisher and crystallization screens from Qiagen.

Cloning and Protein Purification—pHisCPD, pHisCPD/C-S, and pMCSG7-*cpdΔ51* for overexpression of rCPD, rCPD/C-S, and pro-CPD, respectively, have been described (7, 8) and are diagrammed in supplemental Fig. 1. Using the Stratagene QuikChange II XL mutagenesis kit, a C3568S substitution was introduced into pMCSG7-*cpdΔ51* for overexpression of pro-CPD/C-S as previously described (8). Other amino acid substitutions were introduced into pHisCPD (or pMCSG7-*cpdΔ51*) using primers listed in supplemental Table 1.

rCPD and mutant variants were purified as previously described (7). pro-CPD and pro-CPD/C-S were expressed in *Escherichia coli* BL21(λDE3), and soluble protein extracts were prepared by sonication in Buffer A (20 mM Tris, 500 mM NaCl (pH 8.0)) containing 5 mM imidazole as previously described (8). Proteins were loaded onto a 5-ml HisTrap HP column, washed with Buffer A containing 10 mM imidazole, and eluted with Buffer A containing 250 mM imidazole. Eluted protein was diluted with 20 mM Tris (pH 8.0) to reduce NaCl to 100 mM and loaded onto a 5-ml DEAE FF column. The column was washed with 10 mM Tris, 100 mM NaCl (pH 8.0), eluted with 10 mM Tris, 300 mM NaCl (pH 8.0), and then protein was dialyzed against 10 mM Tris, 300 mM NaCl (pH 7.4). Protein concentration was determined using a NanoDrop ND-1000 Spectrophotometer. Unless used for crystallization, protein was adjusted to 10% glycerol and stored at −80 °C.

For post-CPD purification, pro-CPD was incubated with 1 μM $InsP_6$ overnight at 25 °C. To remove $InsP_6$, protein was dialyzed against 10 mM Tris, 300 mM NaCl, 1 mM EDTA (pH 7.4) and then against two changes of buffer without EDTA. Protein was applied to a 5-ml HisTrap HP column and collected in the flow through or separated using a Superdex 200 10/30 column.

For cloning of YopM, primers as listed in supplemental Table 1 were used to amplify *yopM* from *Yersinia pseudotuberculosis* strain 32777 (12) and then ligation independent cloning was used to transfer the product to pMCSG7 according to published protocols (13). Protein was expressed in *E. coli* BL21(λDE3) and then purified from inclusion bodies as described above for pro-CPD except 8 M urea was added to buffers.

For $RtxA_{1580-3909}$ purification, primers as listed in supplemental Table 1 were used to amplify the DNA of the *rtxA* gene corresponding to amino acids 2635–3909 from N16961 genomic DNA. The resulting product was cloned into EcoRI-HindIII sites of pET28a (Novagen) creating pRTX2635. The DNA corresponding to amino acids 1580–2635 with flanking NcoI sites was then amplified. The resulting product was cloned into NcoI-EcoRI sites of pRTX2635 creating pRTX1580. $RtxA_{1580-3909}$ protein with a C-terminal His_6 tag (diagrammed in supplemental Fig. 1) was expressed in *E. coli* BL21(λDE3). A soluble extract was prepared by sonication as described above for pro-CPD and loaded onto a 1-ml HisTrap HP column. Column was washed with Buffer A containing 40 mM imidazole and eluted with Buffer A containing 250 mM imidazole. Protein was then further purified using a Superdex 200 10/30 column in 20 mM Tris, 300 mM NaCl (pH 7.4). Protein was adjusted to 5% glycerol and stored at −80 °C.

Structure Determination and Refinement—For crystallization, pro-CPD/C-S was concentrated to 7 mg/ml (274 μM) in 10 mM Tris, 300 mM NaCl (pH 7.4), and $InsP_6$ was added to 1 mM. To allow the possibility of cleavage of the His_6 tag during crystallization, tobacco etch virus (TEV) protease, purified as described by Kapust and Waugh (14), was added at a 1:200 molar ratio. Crystallization screens were set up in 96-well sitting drop plates (1:1 protein/reservoir) at room temperature. Diffraction quality crystals appeared after 1 week in a condition containing 32% polyethylene glycol 4000, 0.8 M lithium chloride, and 0.1 M Tris-HCl (pH 8.5) (PEGs II Screen, Qiagen).

Crystals for data collection were flash-cooled in liquid nitrogen. Diffraction data (λ = 0.97872 Å) were collected at 100 K from a single crystal at experimental station 21ID-F of the Life Science Collaborative Access Team at the Advance Photon Source, Argonne, IL. Intensities were indexed, integrated, and scaled using the HKL2000 suite (15). The structure was solved with Phaser (16), using post-processed CPD (PDB code 3EEB) with 28 residues omitted at N terminus as a search model. After rigid group refinement and cycles of positional refinement in Refmac (17) missing residues as well as residues of the purification tag were placed in the electron density maps during manual corrections of the model using Coot (18). The quality of the final refined model was verified with Procheck in the CCP4 suite (19, 20). The final model consists of 206 residues of CPD, 9 and 4 residues of the purification tag for chain A and B, respectively, and 2 molecules of $InsP_6$. The data quality and the quality of the model are shown in Table 1. Figures were prepared with MacPyMol (Delano Scientific, Palo Alto, CA).

Isothermal Titration Calorimetry—ITC was performed using a MicroCal MSC-ITC calorimeter as previously described (7) except that, for some studies, post-CPD was preincubated with 100 μM CK inhibitors for 60 min at 25 °C prior to use, and the dissociation constant of post-CPD was determined using 194 μM post-CPD titrated with 7 mM $InsP_6$.

Analysis of Protease Activity and Cleavage Products—Autoprocessing of 2 μg of rCPD, pro-CPD, or mutant variants or 10

TABLE 1
Data collection and refinement statistics

	Crystal 1
Data collection	
Space group	$P2_12_12_1$
Cell dimensions, *a*, *b*, *c* (Å)	46.06, 66.37, 137.96
Resolution (Å)	25.00-1.95 (1.98-1.95)[a]
R_{merge}	0.076 (0.448)
$\langle I\rangle/\langle\sigma(I)\rangle$	18.4 (3.1)
Completeness (%)	99.4 (95.3)
Redundancy	6.7 (3.9)
Refinement	
Resolution (Å)	23.67-1.95
No. reflections	29,932
R_{work}/R_{free}	0.169/0.216
No. atoms	
Protein	3,484
Ligand/ion	85
Water	313
B-factors	
Protein, Chain A/Chain B	39.5/41.2
Ligand-$InsP_6$, Chain A/Chain B	41.5/52.0
Water	45.8
Root mean square deviations	
Bond length (Å)	0.013
Bond angle (°)	1.7
Ramachandran statistics[b]	
Most favored	89.7%
Additional allowed	9.8%
Generously allowed	0.5%

[a] Values in parentheses are for highest resolution shell.
[b] As defined in Ref. 19.

μg of $RtxA_{1580-3909}$ in 20 mM Tris, 150 mM NaCl (pH 7.4) was initiated by addition of $InsP_6$ in 20-μl reaction volumes. When indicated, proteins were incubated with 100 μM NEM or CK inhibitors except for inhibition of $RtxA_{1580-3909}$, which was performed with 1 mM inhibitors. For some studies, proteins were dialyzed against 20 mM Tris, 150 mM NaCl (pH 7.4) to remove excess inhibitor from final reactions. After addition of $InsP_6$ at concentrations indicated in the figure legends, reactions were incubated at 37 °C, stopped by boiling in sample buffer, separated by SDS-PAGE, and stained with Coomassie Blue R250. Fourier transform mass spectrometry (FT-MS) analysis of cleavage reaction mixes was performed as previously described (8).

Cleavage in *trans* of rCPD by post-CPD was performed as described above except that post-CPD was first mixed with 2 μg of rCPD mutant protein and processing was initiated by addition of 100 μM $InsP_6$. Cleavage of other proteins was performed similarly with 10 μg of $RtxA_{1580-3909}$ or YopM.

Thermal Shift Assays—25-μl solutions containing 0.5 mg/ml protein in 20 mM Tris, 150 mM NaCl (pH 7.4), SYPRO® Orange dye diluted 1:625, and $InsP_6$ from 0 to 5 mM were set up in 96-well trays. Samples were heated from 25 to 75 °C in 0.5 °C/12-s steps. Fluorescence intensity was measured at excitation/emission wavelengths of 495/519 nm using a Bio-Rad iCycler iQ5. T_m was calculated by fitting a Boltzmann sigmoidal nonlinear regression curve using Prism4 for Mac from GraphPad Software (La Jolla, CA).

Limited Trypsin Digestions—300 μl of 0.5 mg/ml protein in 20 mM Tris, 150 mM NaCl (pH 7.4) were incubated with or without $InsP_6$ for 1 h at 25 °C. Trypsin was then added to 20-μl aliquots for 30 min at 37 °C followed by addition of 100 μg/ml trypsin inhibitor. For some studies, post-CPD was preincubated with 100 μM CK inhibitors for 30 min at 25 °C. Proteins were analyzed either by SDS-PAGE or FT-MS as previously described (8). Cleavage fragments were identified by comparing mass peaks to fragments predicted by FindPept (21, 22).

N-terminal Sequencing—Autoprocessed $RtxA_{1580-3909}$ fragments were separated by SDS-PAGE and electroblotted onto polyvinylidene difluoride membranes. N-terminal sequences were determined by Edman degradation according to standard protocols using the automated protein sequencing system Procise (Applied Biosystems).

Actin Cross-linking—*In vitro* actin cross-linking assay was performed at 25 °C for 25 min as described previously (23) with 3 μM actin and 0.09 μM $RtxA_{1580-3909}$.

Accession Code—The coordinates and the structure factors for pre-processed CPD have been deposited in PDB with the accession code 3FZY.

RESULTS

Crystal Structure of the Pre-processed Form of MARTX CPD—For this study, the uncleaved form of *V. cholerae* MARTX CPD (residues 3428–3637 according to the original *rtxA* gene annotation by Lin *et al.* (24)) with a catalytically inactive C3568S mutation (pro-CPD/C-S) was purified. This protein has a His_6 tag and TEV protease recognition site upstream of the processing site with no additional sequences on the C terminus (see diagram in supplemental Fig. 1). The binding affinity of pro-CPD/C-S for $InsP_6$ was determined as 0.18 ± 0.02 μM (supplemental Fig. 2). This enhanced binding affinity compared with previously determined dissociation constants likely reflects the absence of long N- or C-terminal extensions present on previously studied CPD proteins (7, 25).

To define the conformation that is adopted upon $InsP_6$ binding, we determined a 1.95-Å resolution crystal structure of pro-CPD/C-S bound with $InsP_6$, capturing the scissile bond in the active site prior to autoproteolysis of the N terminus. The core (Fig. 1*A*, *green*) formed by residues Phe^{3458}–Ser^{3600} is nearly identical to the previously determined structure of the post-processing form of CPD, here referred to as post-CPD (25), and consists of 7 β-strands and 3 α-helices. A C-terminal subdomain (Fig. 1*A*, *magenta*) from Ser^{3601} to the end of the protein forms a 5-stranded β-structure, previously designated the β-flap (25). The N terminus of CPD was predicted by the algorithm DisoPred2 (26) to be disordered and, consistent with this prediction, the N terminus (Fig. 1*A*, blue) is a loose strand wrapped around the outside of the protease. This strand is anchored to the core the P1 residue Leu^{3428} inserted in the S1 site and by Trp^{3442} and Phe^{3458} embedded into the core (Fig. 1*A*, *orange*). Ile^{3445} and Val^{3447} also contact the core through hydrophobic interactions. Altogether, these interactions create a close association of the N terminus from amino acids Trp^{3442} to Phe^{3458} with the protease core.

Structure of the Catalytic and Substrate Binding Sites—At P3′ residue Gly^{3431} (Fig. 1*A*, *orange*), the N terminus turns toward the catalytic site. The enzyme is clearly autocatalytic with its own N terminus in the active site. The catalytic dyad residues, His^{3519} and Ser-substituted Cys^{3568}, are found at the ends of parallel strands, β5 and β6 (Fig. 1*B*). The catalytic residues are separated by ~6 Å with the scissile bond inserted between them. This arrangement is similar to that in the

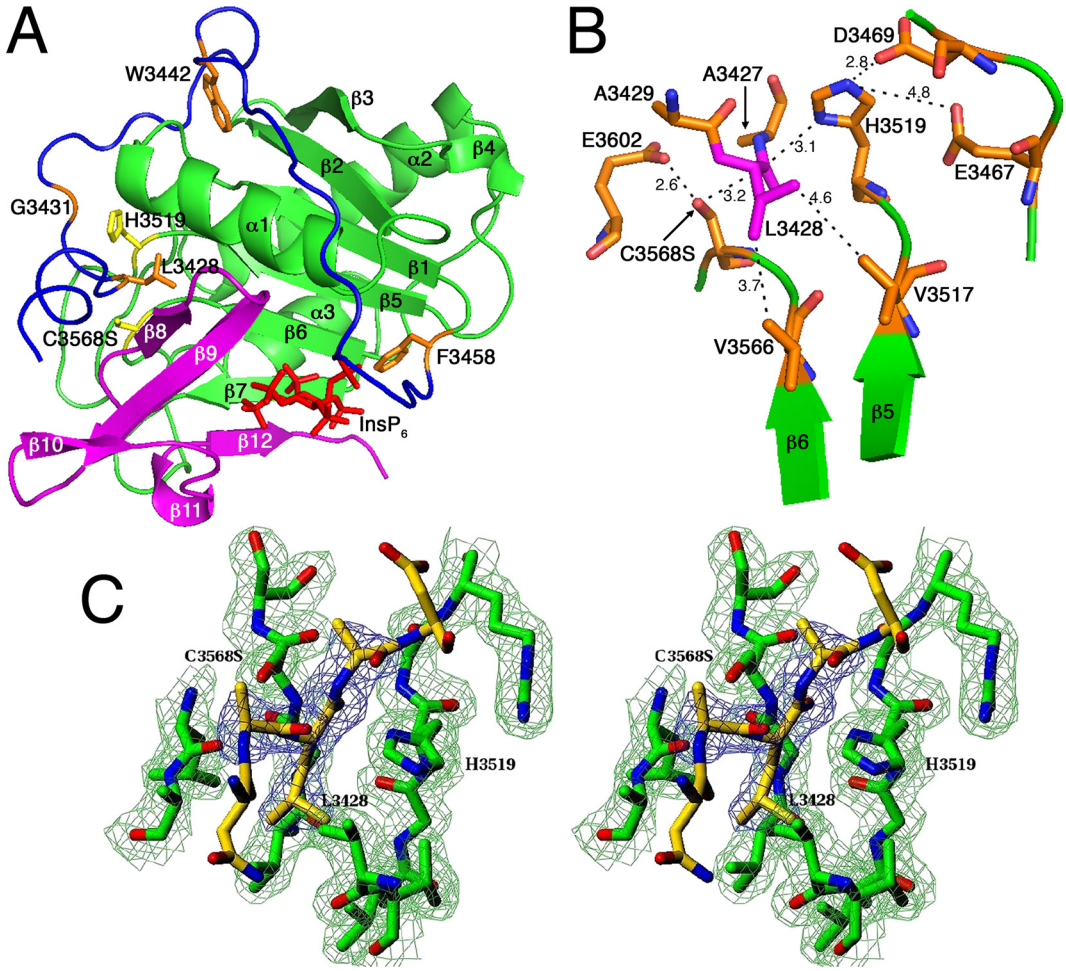

FIGURE 1. **Structural model of pro-CPD/C-S reveals enzyme-substrate complex.** *A*, pro-CPD/C-S with N terminus (*blue*), protease core (*green*), β-flap (*magenta*), and $InsP_6$ (*red*). Key residues (*orange*) and catalytic residues (*yellow*) are labeled according to annotation of Lin *et al.* (24). *B*, schematic representation of the Clan CD fold catalytic site with P1 Leu^{3428} (*magenta*) inserted into S1 site. Distances (in angstroms) of key bonds are shown as *dashed lines*. *C*, stereo view of the active site of pro-CPD/C-S as a *stick model* with surrounding $2F_o - F_c$ map contoured at 1 sigma (*green*) and the N terminus residues, surrounded with omit $F_o - F_c$ map contoured at 4 sigma level (*blue*; omitted residues are Ala-Leu-Ala). For *B* and *C*, carbon of the active site, carbon of the substrate, oxygen and nitrogen atoms are colored in *green*, *yellow*, *red*, and *blue*, respectively.

enzyme-substrate complexes of Clan CD peptidases, including caspases and gingipains (27, 28). Consistent with classification in this clan (29, 30), it is proposed that Cys^{3568} is not activated by His^{3519}, but rather is substrate-activated by close alignment of the scissile bond, as previously proposed for the caspase interleukin-1β converting enzyme (28). This initial step of the processing reaction would then be followed by protonation of the leaving group by His^{3519}.

Concerning other residues involved in catalysis, we previously determined that mutation of either Glu^{3467} or Asp^{3469} does not affect autoprocessing, but a double Glu^{3467}/Asp^{3469} mutant is partially defective (7). Assessment of the structure (Fig. 1*B*) shows that both residues function to properly align the catalytic residue His^{3519}. On the other side of the catalytic site, Glu^{3602} is reoriented in our structure compared with the previous determined structure (25), and in this location appears to correctly position the Ser-substituted catalytic cysteine (Fig. 1*B*). However, the presence of Glu^{3602} is not critical, because mutation of Glu^{3602} to Ala did not affect proteolysis (data not shown).

The most significant difference in this pre-processing structure compared with the previous structure of post-CPD is that the P3, P2, P1, and P1′ residues of the N terminus are rigidly aligned in a binding cleft (Fig. 1*C*). The P1 Leu^{3428} is inserted in the hydrophobic S1 site where it interacts directly with residues Val^{3517} and Val^{3566} on the parallel β5 and β6 strands of the Clan CD peptidase fold. Other residues that form the hydrophobic pocket are Val^{3472}, Ala^{3475}, and Leu^{3479} on α1 of the protein core and Leu^{3603} and Val^{3605} of β8 of the β-flap. Interestingly, the P3 Asn residue that is part of the TEV protease recognition site is not accessible, accounting for the presence of the N-terminal tag in the structure despite attempts to remove it during crystallization (see "Experimental Procedures").

Autoprocessing Is Specific for Leu-Xaa Bond—All Clan CD proteases have specificity for cleavage at a single recognition residue (29, 30). Specificity for Leu as the P1 residue was demonstrated by altering Leu^{3428} to Ala in rCPD, a recombinant variant of CPD with a longer 75 residue extension facilitating easy detection of autoproteolysis (7) (diagrammed in supplemental Fig. 1). This protein was autoprocessed, but the site of hydrolysis shifted primarily to the P10 Leu residue (Fig. 2, *A* and *C*). At lower concentrations of $InsP_6$, other sites were also cleaved resulting in a banding pattern that likely corresponds to

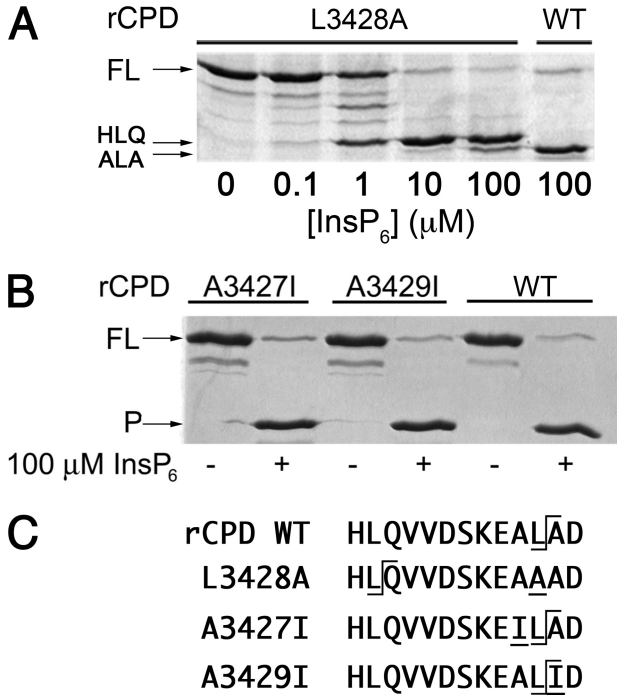

FIGURE 2. **rCPD cleaves a Leu-Xaa bond.** *A*, Coomassie-stained gel of autoprocessing of rCPD L3428A (*FL*) shows processing at indicated concentrations of $InsP_6$ after 2-h incubation resulted in a shift of autocleavage site to generate a slightly longer fragment (marked *HLQ*) than the normal processing site (marked *ALA*). *B*, the size of the processed product (*P*) is not affected by alteration of P2 or P1′ Ala residues in full-length rCPD (*FL*). *C*, the predominant processing sites in various point mutants (*double underlined*) are shown by *line bracket*.

processing at any of five other leucine residues in the extension on rCPD.

Changing the Ala residues in either the P1′ or the P2 position of rCPD to Ile had no effect on processing (Fig. 2, *B* and *C*). Consistent with this finding, these side chains are directed to the solvent (Fig. 1*C*). Furthermore, in the rCPD L3428A mutant, the new P10 processing site is flanked by histidine and glutamine (Fig. 2*C*), also confirming that the P2 and P1′ residues do not contribute to substrate recognition. Within the structure of activated CPD, the Asp residue in position P2′ is pointed toward solvent, and mutation of this residue did not affect autoprocessing (data not shown). At the P3 position, the normally present Glu residue is replaced by Asn from the TEV protease recognition site in pro-CPD, and thus, there is no specificity at this position. Overall, recognition of the CPD cleavage site depends only on Leu^{3428}.

Binding of $InsP_6$ Increases CPD T_m and Trypsin Resistance—Twelve residues previously identified to contact $InsP_6$ (7, 25) are dispersed between the N-terminal leader, the core, and the β-flap, but come together in the tertiary structure (Fig. 3*A*). This global requirement for binding suggests a primary role of $InsP_6$ may be to stabilize CPD after translocation as an unfolded or partially folded protein. Consistent with this proposal, ITC experiments revealed a high negative enthalpy and entropy upon ligand binding that would be consistent with the protein becoming more ordered ($\Delta H = -59$ kcal/mol; $T\Delta S = -50$ kcal/mol at 300 K, supplemental Fig. 2). Yet, no change in quaternary structure occurred upon $InsP_6$ binding as determined by dynamic light scattering, chemical cross-linking studies, and native gel electrophoresis (data not shown). Thus, the entropy change upon $InsP_6$ binding is most consistent with the monomeric protein becoming more ordered, which should result in a more stable structure.

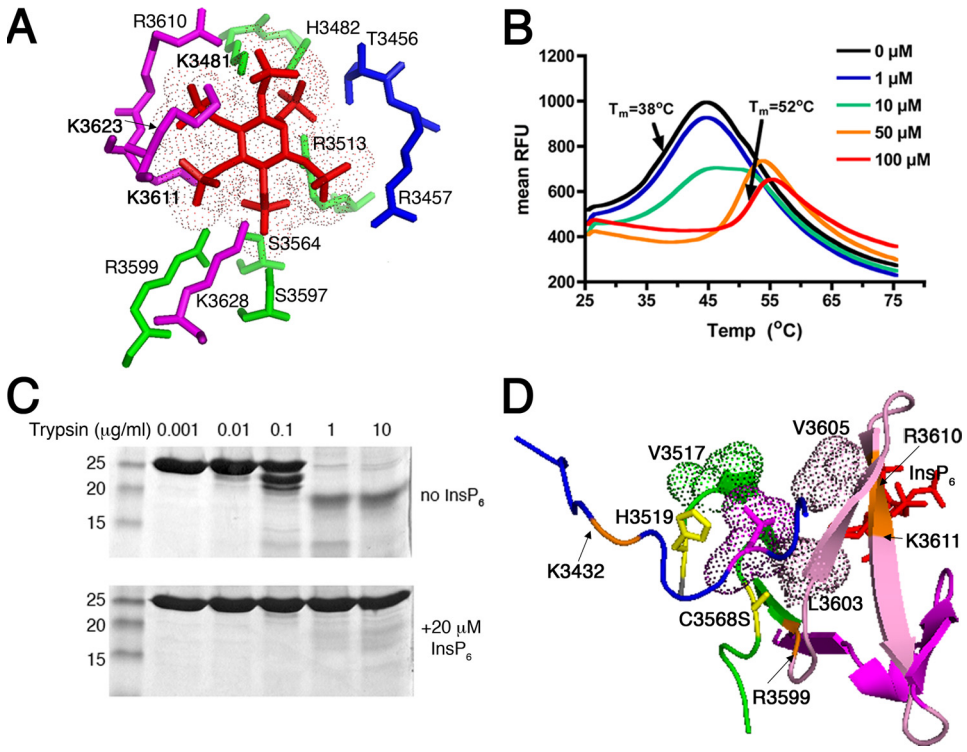

FIGURE 3. **Binding of $InsP_6$ increases pro-CPD/C-S T_m, and the protein becomes trypsin-resistant.** *A*, close-up view of $InsP_6$ binding pocket shows 12 residues known to contact $InsP_6$ (*red* with *space-filling dots*) derive from the N terminus (*blue*), the protease core (*green*), and the β-flap (*magenta*). *B*, SYPRO® Orange melting curves of pro-CPD/C-S at different concentrations of $InsP_6$. *C*, Coomassie-stained gel of limited proteolysis of pro-CPD/C-S at varying concentrations of trypsin. Locations of trypsin cleavage in the absence of $InsP_6$ as determine by FT-MS (supplemental Fig. 4) are shown in *orange* in *D* with color scheme used in *A* except antiparallel $\beta 8\beta 9$ are highlighted *pink* and S1 hydrophobic residues are *space-filling dots*.

To determine to what extent protein stability changes, T_m was determined by SYPRO® Orange thermal shift assays (31–33). In the absence of $InsP_6$, half-maximal denaturation of pro-CPD/C-S occurred at 38 °C, but increased by 14 °C with addition of $InsP_6$ (Fig. 3*B*). As further evidence of increased protein stability, trypsin-sensitive pro-CPD/C-S became resistant to trypsin digestion after $InsP_6$ was added (Fig. 3*C*).

It is possible that the low T_m of pro-CPD/C-S is due to electrostatic repulsion when the positively charged binding site is not occupied by $InsP_6$. Consistent with this, high concentrations of phosphate and citrate also increased T_m. However, pro-CPD in phosphate had only

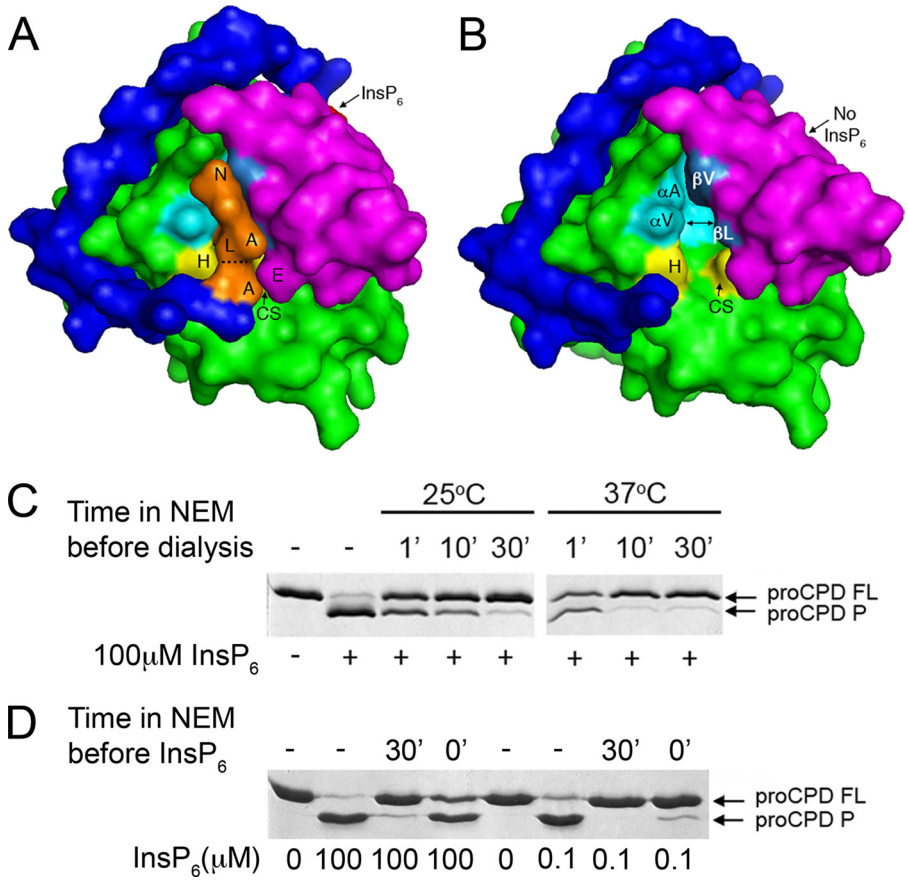

FIGURE 4. **Susceptibility of pro-CPD to inhibition by NEM depends on position of the N terminus.** *A* and *B*, space-filling models of pro-CPD/C-S with N terminus (*blue*), protease core (*green*), β-flap (*magenta*), and $InsP_6$ (*red*). Models are based on the B molecule. In *A*, P1 Leu^{3428} (*L*) is buried in the hydrophobic S1 site (formed by the protease core (*aqua*) and the β-flap (*medium blue*)). The catalytic cysteine (*CS*) is covered by the N terminus at the scissile bond (*dashed line*) and by Glu^{3602} (*E*). In *B*, removal of the N terminus to the P3′ residue exposes the catalytic cysteine. Other labeled residues are the catalytic His^{3519} (*H*), S1 residues from the α1 helix (Val^{3472} (*αV*) and Ala^{3475} (*αA*)) or the β8 strand (Leu^{3603} (*βL*) and Val^{3605} (*βV*)) and the N terminus residues in the binding cleft identified by *single letter code*. *C*, pro-CPD with intact cysteine was incubated without (−) or with 100 μM NEM for time indicated at temperature indicated after which the protein was dialyzed at 4 °C to remove excess inhibitor. Processing was then initiated by addition of 100 μM $InsP_6$ and processing of full-length protein (*FL*) to the processed form (*P*) was assessed by SDS-PAGE. *D*, pro-CPD was incubated without NEM (−) or with 100 μM NEM for 30 min (*30′*) at 25 °C after which $InsP_6$ as indicated was added. In *lanes* marked *0′*, NEM and $InsP_6$ were first mixed and then added simultaneously (0 min preincubation). Reactions were monitored after 1 h at 37 °C by SDS-PAGE.

50% processing capability within 3 h, and citrate stimulated no processing above controls, even after 24 h (supplemental Fig. 3). Thus, $InsP_6$ is required both to stabilize the protein and to generate an active conformation for efficient cleavage.

Arg^{3599}, Arg^{3610}, and Lys^{3611} Are Exposed to Trypsin in pro-CPD/C-S in Absence of $InsP_6$—To define the region that is structurally altered upon $InsP_6$ binding, we used FT-MS to identify peptides after limited proteolysis in 0.1 μg/ml trypsin (supplemental Fig. 4). Arg^{3599}, Arg^{3610}, and Lys^{3611} were identified as exposed residues by detection of peptides originating from the site of cleavage to both the N and C terminus, indicating they were generated by a single cleavage event. Each of these residues contact $InsP_6$; moreover, Lys^{3611} is known to be absolutely essential for $InsP_6$ binding (7). By contrast, other Arg and Lys residues that form the $InsP_6$ binding pocket were protected, indicating that there is not complete exposure of the $InsP_6$ pocket.

The location of these residues within the structure revealed that the antiparallel β8β9 hairpin structure of the β-flap that directly contacts both $InsP_6$ and P1 Leu^{3428} (Fig. 3*D*) could be improperly folded or have an altered arrangement in the absence of $InsP_6$. In particular, access of trypsin to Arg^{3599} would require the absence of an interaction of the loop between strands β7 and β8 with the loop that contains the catalytic cysteine. Furthermore, changes in the structure of β8 just after Arg^{3599} would affect the S1 site. This is consistent with $InsP_6$ contributing to proper structure of β8β9 and consequently the active site.

N Terminus Is Tightly Bound in the Active Site of pro-CPD upon Binding of $InsP_6$—Because a structural rearrangement upon $InsP_6$ binding may alter the S1 pocket, it was considered whether the N terminus would require $InsP_6$ to enter the binding cleft. As seen if Fig. 4*A*, when Leu^{3428} occupies the active site, cysteine proteases inhibitor NEM would only have access to Cys^{3568} when the N terminus transiently departs from the active site. By contrast, if the N terminus is not in the S1 site, Cys^{3568} is exposed to solution and should be rapidly inactivated. When pro-CPD was preincubated with NEM at 25 °C followed by dialysis at 4 °C to remove the inhibitor prior to addition of $InsP_6$, long preincubation times (up to 30 min) were required for 100% inactivation (Fig. 4*C*). If NEM incubation was performed at 37 °C, inactivation was more rapid. These data indicated that the N terminus normally occupies the S1 site independent of $InsP_6$ binding, but may transiently exit the site such that NEM has access to the catalytic cysteine over time.

Further evidence that the N terminus is not tightly bound within $InsP_6$-free pro-CPD was obtained from the analysis of fragments after limited trypsin proteolysis of pro-CPD/C-S (supplemental Fig. 4). A 3240.50 mass unit fragment corresponding to trypsin cleavage after Lys^{3432} in pro-CPD/C-S indicated that the N-terminal loop between the Trp^{3442} anchor and the P1 Leu^{3428} is accessible to trypsin. However, because the $InsP_6$-bound form was trypsin-resistant (Fig. 3*C*), the N terminus becomes more tightly bound, or locked, in the active site by $InsP_6$. If this is correct, NEM should not inhibit autoprocessing if added simultaneously with 100 μM $InsP_6$, because rapidly locking the N terminus in a high concentration of $InsP_6$ should

block access of inhibitors to the cysteine. Indeed, when added simultaneously, NEM poorly inhibited $InsP_6$-induced autoprocessing (Fig. 4*D*). By contrast, at an $InsP_6$ concentration below the K_d, when <50% of pro-CPD would be bound by $InsP_6$ immediately upon addition, inhibition was >90% (Fig. 4*D*).

Therefore, NEM and trypsin accessibility data support the view that binding of $InsP_6$ to pro-CPD causes $\beta8\beta9$ to adopt an appropriate structure for proper orientation of the scissile bond relative to the catalytic cysteine. Hydrophobic residues on $\beta8$ would then fully contribute to the S1 site. The resulting rigid structure would then be amenable to substrate-activated autoprocessing initiated by Cys^{3568}.

After Processing, CPD Loses Its High Affinity for $InsP_6$—A previous study suggested that post-CPD has reduced affinity for $InsP_6$ (7). Determination of the binding affinity of $InsP_6$ for purified post-CPD revealed that it does indeed have reduced affinity (K_d = 100 ± 20 μM, n = 6). Notably, this dissociation constant is above the physiological concentration of $InsP_6$ (40–60 μM (9–11)) indicating that only a fraction of processed CPD would bind $InsP_6$ *in vivo*. This result suggests that, upon processing, CPD adopts a conformation distinct from the preprocesssing form. However, a structure similar to pro-CPD can be adopted upon incubation in very high concentrations of $InsP_6$, as was done to determine the structure of post-CPD (25).

Post-CPD Is Biochemically Distinct from pro-CPD—In the absence of an $InsP_6$-free post-CPD structure, the structural alteration resulting in reduced $InsP_6$ binding affinity was investigated using biochemical assays. Limited proteolysis of post-CPD revealed trypsin cleavage at Lys^{3611}, further indicating poor binding of $InsP_6$ to post-CPD (Fig. 6*A* and supplemental Fig. 5). However, in contrast to the pre-processed form, access of trypsin to Arg^{3599} likely occurred only after prior cleavage at Arg^{3610} or Lys^{3611} (supplemental Fig. 5). In addition, thermal shift assays showed that post-CPD has a T_m of 47 °C independent of $InsP_6$ concentration (supplemental Fig. 6). Altogether, these data indicate that $\beta8\beta9$ remained folded after processing and post-CPD did not simply revert to the less stable, pre-processed form.

Reoccupation of the S1 Site Reactivates CPD for High Affinity Binding of $InsP_6$—To model the effect of N-terminal processing, the pro-CPD/C-S structural model was re-examined with the N terminus removed (Fig. 4*B*). Departure of Leu^{3428} from the S1 site would expose the S1 site creating an 8–9 Å hydrophobic crevice in the protein. We postulated that the exposed hydrophobic S1 site could be partially covered by a shift of $\beta8$ toward $\alpha1$. A simultaneous shift of the antiparallel $\beta9$ would weaken contact of $\beta9$ residues Arg^{3610}, and particularly Lys^{3611}, with $InsP_6$, releasing $InsP_6$ from the binding pocket. Lys^{3611} would then become susceptible to trypsin as observed (Fig. 5*A*). If this model is correct, then re-occupation of the S1 site could reactivate CPD by shifting $\beta8\beta9$ back toward the $InsP_6$ pocket, thereby facilitating high affinity binding of $InsP_6$ and formation of a new active enzyme-substrate complex.

To test this model, CK inhibitors were used as a surrogate for a new substrate. CK inhibitors are known to inactivate Clan CD proteases by alkylating the catalytic cysteine when the attached amino acid or peptide binds within the S1 site (34). When post-CPD was pre-treated with inhibitor L-LeuCK or *z*-Gly-Leu-Phe-CK, $InsP_6$ did bind with K_d = 0.9 ± 0.3 μM and K_d = 1.1 ± 0.3 μM, respectively (Fig. 5*B*). These data indicate that the CKs successfully mimic a new substrate and restimulate post-CPD for $InsP_6$ binding. Furthermore, these proteins regain resistance to trypsin, indicating reformation of a structure resembling an active enzyme-substrate complex (Fig. 4*A*), which was not reached even by addition of 2 mM $InsP_6$ without CKs.

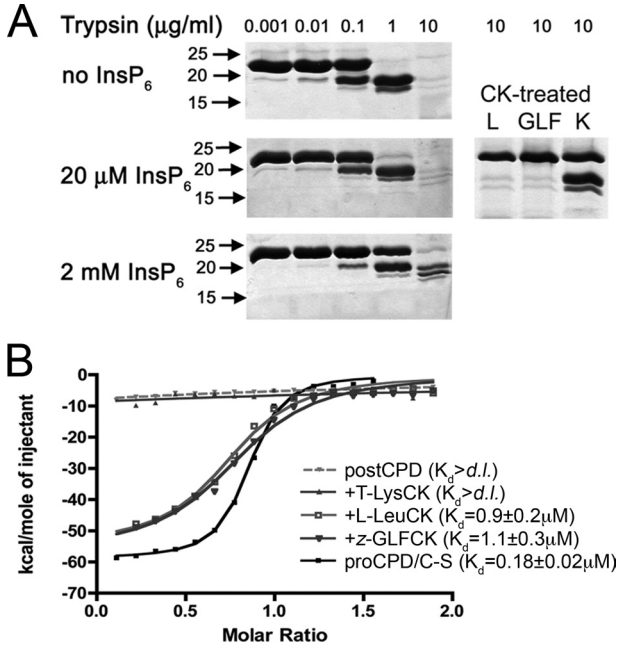

FIGURE 5. **Post-CPD can be reactivated for $InsP_6$ binding and trypsin resistance by CK inhibitors.** *A*, purified post-CPD was treated with trypsin in the presence of the indicated concentration of $InsP_6$ and proteolysis was determined by SDS-PAGE. The first cleavage band arising in the 0.1-μg *lane* corresponds to processing at Arg^{3610} or Lys^{3611}, with the second band corresponding to processing at Arg^{3599}. Preincubation of post-CPD with 100 μM L-LeuCK (*L*) or *z*-Gly-Leu-Phe-CK (*GLF*) protects the protein from trypsin while T-LysCK (*K*) is only partially protected. *B*, ITC curves performed with 15 μM post-CPD as described under "Experimental Procedures" either without or with preincubation of 100 μM CK inhibitors as noted. The pro-CPD/C-S control is shown in greater detail in supplemental Fig. 2.

T-LysCK did not restore binding of $InsP_6$ to a detectable range under standard conditions indicating that reoccupation of the S1 site by Leu is specifically required for reactivation of post-CPD. These data thus support a cooperative model for reactivation of CPD requiring both a new substrate and $InsP_6$.

Reactivated post-CPD Can Process rCPD Mutant Proteins in Trans—Because post-CPD can be reactivated for $InsP_6$ binding, it should be able to process other substrates. Previously, cleavage of other proteins was discounted, because the Leu^{3428}–Ala^{3429} bond in the rCPD/C-S protein used as a substrate was not processed during co-incubation with pro-CPD (7). However, our data presented here indicate that, in the presence of $InsP_6$, Leu^{3428} in the rCPD/C-S mutant would be locked within the S1 site and thereby protected. To determine if cleavage of rCPD/C-S by post-CPD occurs when the N terminus is freely available, rCPD/C-S with a mutation of hydrophobic S1 pocket residue Leu^{3479} preventing insertion of the N terminus was used as a substrate. Indeed, rCPD L3479D/C3568S was cleaved by post-CPD in an $InsP_6$-dependent manner, demonstrating in *trans* processing. rCPD/C-S, which can bind $InsP_6$, was fully protected, while mutant proteins that cannot bind

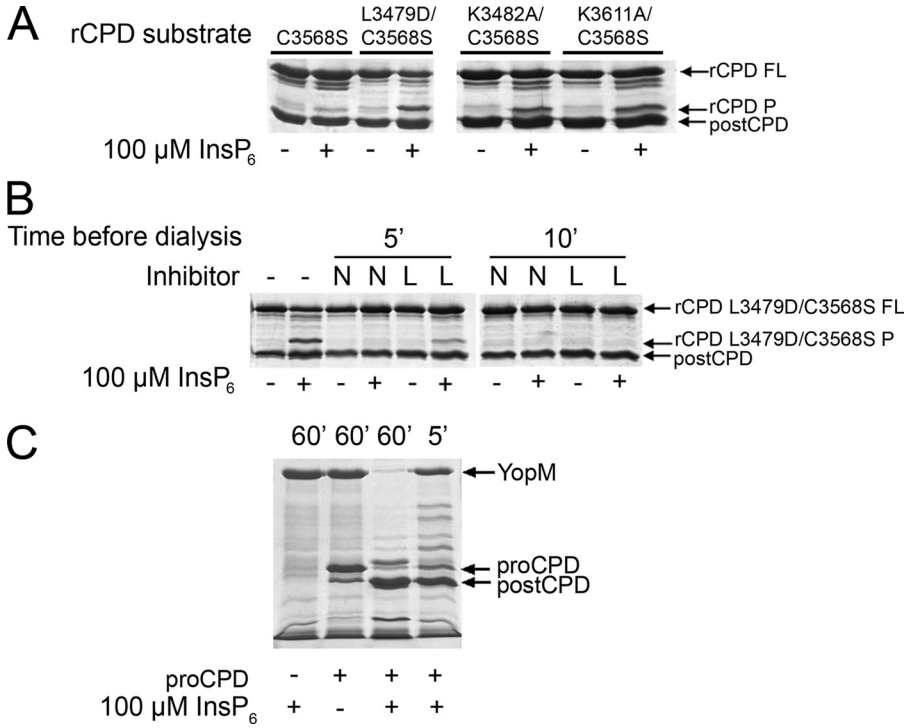

FIGURE 6. **Post-CPD can process rCPD/C-S and unrelated proteins in *trans*.** *A*, Coomassie-stained gels of full-length mutant rCPD/C-S proteins (rCPD FL) that cannot autoprocess. If the S1 site (L3479D) or $InsP_6$ binding pocket (K3611A and K3482A) are modified to prevent Leu from occupying its own active site, mutants can be processed (*P*) by post-CPD dependent upon $InsP_6$. *B*, full-length proteins (*FL*) were incubated without inhibitor (−) or with 100 μM NEM (N) or L-LeuCK (L) for time specified at 25 °C after which proteins were dialyzed to remove excess inhibitor, and 100 μM $InsP_6$ was added when indicated to initiate processing of rCPD L3479D/C3568S by post-CPD. Processing for both assays was determined by SDS-PAGE after 2-h incubation at 37 °C. For *C* purified YopM was incubated with pro-CPD in 0.4 M urea and with or without 100 μM $InsP_6$ for the times indicated.

$InsP_6$ were cleaved by post-CPD, confirming that binding of $InsP_6$ locks the N terminus in the S1 site (Fig. 6*A*). Post-CPD processing of rCPD L3479D/C3568S was rapidly inhibited by NEM and L-LeuCK, thereby demonstrating that the cysteine in post-CPD is freely exposed (Fig. 6*B*). Together, these data demonstrate that cooperativity through binding of a new substrate along with a new molecule of $InsP_6$ can reactivate CPD to process other proteins. These data also show that the cysteine is freely accessible in the absence of the N terminus, and thus activation of CPD does not involve an allosteric structural conversion to expose a previously buried cysteine to substrate, as suggested by others (25).

Post-CPD Degrades the Leucine-rich Protein YopM—To test if CPD is a specific protease able to cleave only the Leu^{3428}–Ala^{3429} site of MARTX or if it is able to promiscuously cleave other sites, YopM of *Yersinia pseudotuberculosis* was used as a substrate for in *trans* processing. YopM is a member of the leucine-rich repeats family of type III secretion effectors and is composed of 21% leucine (35). To expose the normally buried Leu residues (36), YopM was purified from inclusion bodies and 0.4 M urea was included in processing reaction mix. pro-CPD without $InsP_6$ or $InsP_6$ alone did not affect YopM. When pro-CPD and $InsP_6$ were added to YopM together, pro-CPD first autoprocessed then extensively cleaved YopM within 5 min, and YopM was completely degraded by 1 h (Fig. 6*C*). Therefore, CPD is a promiscuous enzyme able to cleave non-MARTX substrates.

A Longer Form of MARTX Can Be Both Autoprocessed and Processed in Trans—Because the CPD can be reactivated and can recognize apparently any exposed leucine residue, it was considered likely that CPD can process MARTX at additional sites. To test this, recombinant protein $RtxA_{1580-3909}$ was purified. This protein contains all of the effector domains and part of the protein upstream and downstream without the repeat regions.

As the first experiment (Fig. 7*A*, *lanes 1–6*), $RtxA_{1580-3909}$ was treated with NEM for 30 min to inactivate its CPD, and then this protein was dialyzed and used as a substrate for in *trans* processing by pro-CPD. After $InsP_6$ was added, pro-CPD cleaved itself generating post-CPD, and then processed $RtxA_{1580-3909}$ (*lane 4*). The processing pattern differed over time, but within 1 h, full-length $RtxA_{1580-3909}$ was absent, and stable processing products were generated (*lanes 4–6*). To test for autoprocessing, $InsP_6$ was added to $RtxA_{1580-3909}$ alone. This protein was also processed to stable fragments (*lanes 7–9*).

A distinct difference in the band patterns between in *trans* processing (Fig. 7*A*, *lane 6*) and autoprocessing (*lane 9*) of $RtxA_{1580-3909}$ was noted. In particular, an in *trans* processing product (marked by an *asterisk*) is absent from the autoprocessing reaction. This fragment corresponds to the αβ hydrolase (see below), indicating that Leu^{3428}–Ala^{3429} is poorly cleaved during in *trans* processing but readily cleaved by autoprocessing. This presumably occurred because Leu^{3428}–Ala^{3429} is locked in the active site of the inactivated $RtxA_{1580-3909}$ when $InsP_6$ is added and is thus not accessible for in *trans* cleavage, in agreement with our observations with rCPD/C-S as a substrate.

$RtxA_{1580-3909}$ preincubated with NEM was not capable of autoprocessing, indicating that autoprocessing depends on the catalytic cysteine of CPD (Fig. 7*B*). By contrast, in the presence of L-LeuCK, autoprocessing of $RtxA_{1580-3909}$ was only partially inhibited. Notably, protein that was processed was predominantly cleaved into two fragments of an estimated 59 and 185 kDa, a processing event that corresponds to cleavage at Leu^{3428}–Ala^{3429}, but inhibition of further processing (Fig. 7*B*). These data indicated that processing occurs first at Leu^{3428}–Ala^{3429} followed by processing at other sites. In the presence of L-LeuCK, the cysteine is readily inhibited once Leu^{3428}–Ala^{3429} departs the active site, confirming results with post-CPD.

Autoprocessing of $RtxA_{1580-3909}$ Results in Releasing of Individual Domains—To define the processing sites within $RtxA_{1580-3909}$, Edman degradation was performed to identify the five predominant fragments that arose after 1-min autopro-

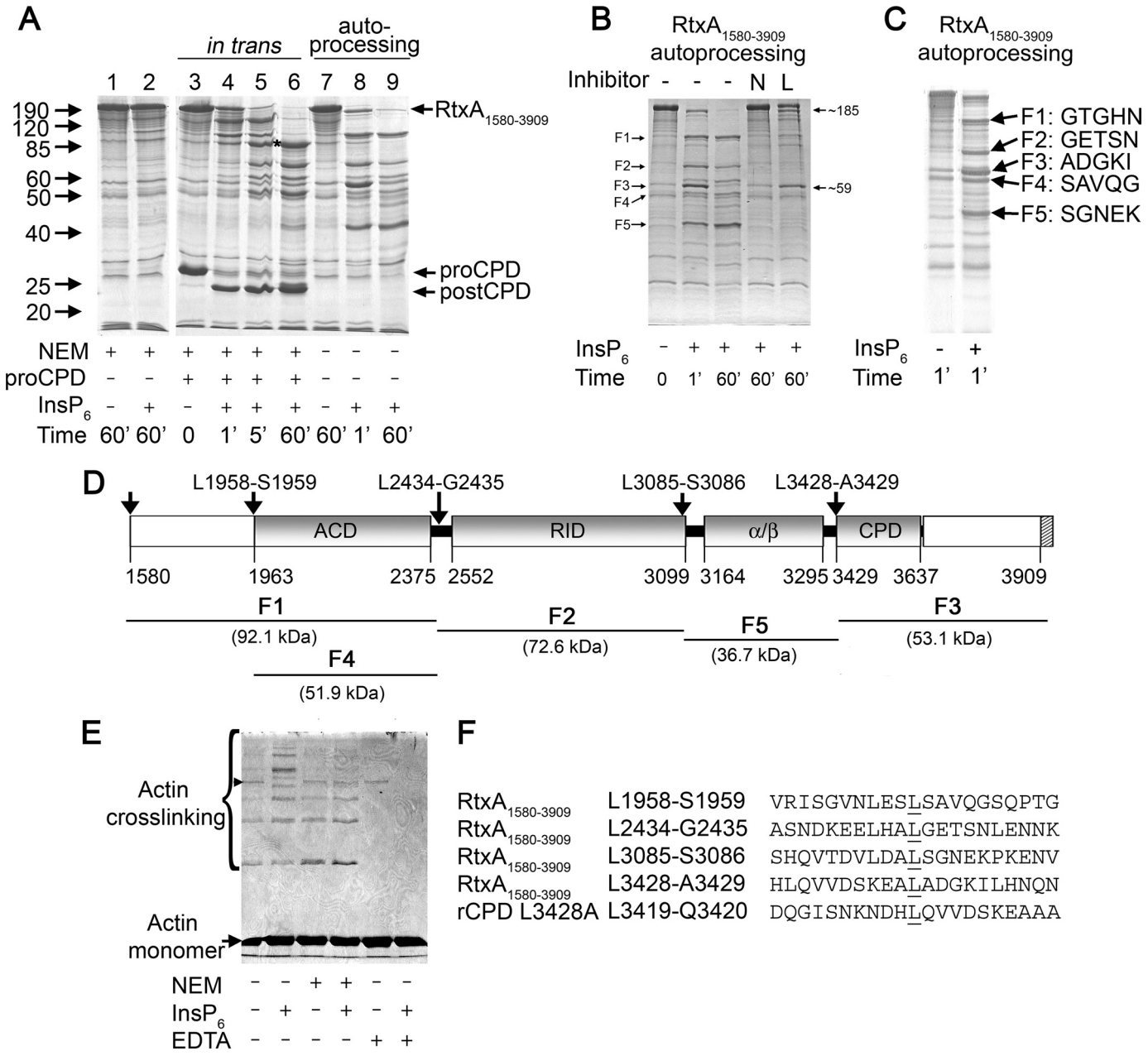

FIGURE 7. **CPD domain autoprocesses RtxA$_{1580-3909}$ in multiple sites.** *A* and *B*, NEM-treated RtxA$_{1580-3909}$ can serve as a substrate for CPD (*lanes 1–6*), but also untreated RtxA$_{1580-3909}$ can undergo autoprocessing (*lanes 5–7*). For *lanes 1–6*, RtxA$_{1580-3909}$ was first incubated with 1 mM NEM for 30 min at 25 °C, after which protein was dialyzed to remove excess of inhibitor. Then, pro-CPD was mixed with NEM-pretreated RtxA$_{1580-3909}$ and incubated for 5 min, after which processing was initiated by addition of 100 μM InsP$_6$ for the time indicated at 37 °C. The *asterisk* in *lane 6* marks a fragment absent after autoprocessing (*lane 9*). For *lanes 7–9* and *B*, autoprocessing reactions were initiated by addition of 100 μM InsP$_6$ and terminated at the times indicated. In *B*, protein was incubated with 1 mM of the indicated inhibitor for 30 min at 25 °C prior to addition of InsP$_6$. *C* and *D*, N termini of marked fragments (*F1–F5*) arising from autoprocessing for 1 min were identified by Edman degradation and are marked with the first five amino acids (*C*) and correspond to fragments diagrammed in *D*. *E*, actin cross-linking reactions were performed with commercial G-actin in the presence of InsP$_6$ and/or 5 mM EDTA as indicated. The *arrowhead* marks RtxA$_{1580-3909}$. *F*, alignment of all processing sites identified in this study.

cessing (Fig. 7*C*). The N-terminal sequences indicated processing at four sites: Leu^{1958}–Ser^{1959}, Leu^{2434}–Gly^{2435}, Leu^{3085}–Ser^{3086}, and Leu^{3428}–Ala^{3429} (Fig. 7, *C* and *D*).

The largest fragment F1 (92 kDa) corresponded to the N terminus of RtxA$_{1580-3909}$, through the ACD ending at Leu^{2434} (Fig. 7*D*). The smaller 52 kDa fragment (F4) corresponding to Ser^{1959}–Leu^{2434} would also contain the ACD. These results indicate the naturally occurring ACD is a 480-amino acid protein, larger than the functionally active 412-amino acid protein (5). It is also notable that the pattern of RtxA$_{1580-3909}$ processing sometimes varied between experiments. In some cases, ACD was found predominantly as the 92-kDa F1 fragment with Leu^{1958}–Ser^{1959} uncleaved, whereas in other experiments, it was mostly present as the F4 52-kDa fragment. This variability indicated that the region upstream of ACD may be not be consistently properly folded and interferes with the processing reaction. However, in the context of the holotoxin, this structure would presumably be

properly folded and thus consistently recognized for cleavage by CPD.

The N-terminal sequence of the 75-kDa F2 fragment revealed a processing site at Leu^{2434}–Gly^{2435}, and its size would suggest it corresponds to the RID domain extending to Leu^{3085}. Thus, the naturally occurring size of the RID is 651 amino acids, larger than the 547-amino acid functional protein (6). The N-terminal sequence of the 36-kDa F5 fragment indicated the processing at Leu^{3085}–Ser^{3086} encompassing the $\alpha\beta$ hydrolase to the known CPD processing site at Leu^{3428}–Ala^{3429}.

Concerning the CPD itself, at 1 min after initiation of the autoprocessing reaction (Fig. 7*A*, *lane 8*), the CPD from $RtxA_{1580-3909}$ was identified as the 59-kDa F3 fragment, which corresponds to the theoretical size of 53.1 kDa, from the Leu^{3428}–Ala^{3429} cleavage site to the C terminus of $RtxA_{1580-3909}$, including the His_6 tag. As an independent verification, the presence of a His_6 tag on the F3 fragment was demonstrated by Western blotting using anti-His antibody (data not shown). This fragment disappears within 60 min (*lane 9*), and Western blotting with anti-CPD antibody indicated the CPD was C-terminally processed to shorter fragments suggesting processing at any of the 19 leucine residues downstream of the CPD in $RtxA_{1580-3909}$ (data not shown). The variability of cleavage at these sites indicates this represents nonspecific processing events in the partially unstructured ends of a recombinant protein.

Altogether, from these analyses, it was found that autoprocessing of MARTX results in processing of at least four sites resulting in release of the three effector domains. If autoprocessing is left unchecked, the promiscuous CPD may also apparently cleave other exposed leucine residues.

Releasing of ACD Increases Its Activity—If the effectors are normally bound in a larger structure, then releasing them may be essential for these proteins to become catalytically functional. To investigate this, *in vitro* actin cross-linking catalyzed by the ACD within $RtxA_{1580-3909}$ was performed with and without addition of $InsP_6$ (Fig. 7*E*). $RtxA_{1580-3909}$ in the absence of $InsP_6$ induced some actin cross-linking, but the activity was increased by the addition of $InsP_6$. The cross-linking of actin was inhibited by NEM, demonstrating the CPD was essential, and by EDTA, demonstrating the cross-links are due to the Mg^{2+}-dependent ACD.

DISCUSSION

In this study, we determined the crystal structure of the enzyme-substrate complex of *V. cholerae* MARTX CPD and used the structural information to suggest biochemical experiments that probe the site and mechanism for $InsP_6$-induced autoproteolysis both at Leu^{3428}–Ala^{3429} and at other sites of the MARTX toxin. The results support a model of controlled activation in which $InsP_6$ is important for stabilization of the proenzyme prior to autoprocessing specifically at leucine residues, and secondly, functioning as a cooperative factor along with a new substrate for reactivation to process other sites. The subsequent autoprocessing would release the effector domains ACD, RID, and $\alpha\beta$ hydrolase to independently access different targets. The final result will be actin destruction due to cross-linking of cytosol localized monomeric G-actin by the ACD and membrane localized Rho GTPases by the RID. Consistent with the important role of autoprocessing, both *in vitro* (Fig. 7) and *in vivo* (8), the ability of MARTX to efficiently cross-link actin was dependent upon it first being autoprocessed to release the ACD as an independent domain.

In this study, we sought to more clearly understand the mechanism of $InsP_6$-induced autoprocessing. The key structure for coordination of the interdependence between the peptide to be processed and $InsP_6$ was identified as a β-hairpin structure ($\beta8\beta9$) that contacts the target peptide through interaction of hydrophobic side chains on $\beta8$ with P1 Leu^{3428} and with $InsP_6$ through arginine and lysine side chains on $\beta9$. Prior to $InsP_6$ binding, the protein exists in a conformation that is trypsin-sensitive, most notably in a portion that may coordinate proper alignment of the catalytic cysteine. Upon $InsP_6$ binding, the β-flap undergoes a structural alteration that locks the substrate in the S1 pocket in a rigid structure amenable to substrate-activated processing. Subsequent to autoprocessing, the β-flap may shift resulting in release of $InsP_6$, although cooperative binding of a new substrate and $InsP_6$ results in reactivation.

Previously, a role for the β-flap in coupling binding with activation was postulated and Asp^{3606} and Trp^{3620} were identified as residues within the β-flap that would affect folding and thereby establish contact with the protein core (25). In this work, the key structure that is altered is identified as $\beta8\beta9$ by partial trypsin digestion.

This model for activation of CPD by formation of a stable enzyme-substrate complex is quite distinct from a recently proposed model for activation by an allosteric structural conversion to expose a previously buried catalytic cysteine to substrate (25). That allosteric mechanism was suggested when it was observed that a 1-min exposure of pro-CPD to NEM followed by addition of $InsP_6$, poorly inhibited autoprocessing, initially suggesting the catalytic cysteine is buried, a result we confirm in Fig. 4*C*. However, we found that a longer incubation of pro-CPD or $RtxA_{1580-3909}$ with NEM fully inhibited processing indicating transient exposure of the cysteine (Figs. 5 and 7). Inspection of the pre-processed structure suggests the N terminus frequently, if not predominantly, occupies the S1 site, thus, inhibition by NEM occurs only over time, targeting the cysteine only when the N terminus transiently departs the active site. This time dependence was reduced by increasing temperature to promote more dynamic motion of the N terminus and, most definitively, by complete removal of the N terminus by autoprocessing. The free accessibility of the cysteine in the processed form of CPD is also strongly supported by the observation that a fluorescent maleimide reacts with the cysteine of post-CPD, but not the pre-processed form, when it is added to a CPD autoprocessing assay (25).

The previously proposed allosteric mechanism was further suggested by the observation that autoprocessing of pro-CPD was inhibited when exposed to NEM and $InsP_6$ simultaneously (25). These data directly conflict with our results that simultaneous addition of NEM with 100 μM $InsP_6$ resulted in a protein that was ~80% processed (Fig. 4*D*). It is notable that the experiment cited as supporting a structural transition was carried out with pro-CPD that had previously been exposed to NEM for

1 min, a treatment expected to inhibit ~50% of CPD (Fig. 4*C*). Thus, the pro-CPD in that experiment was probably inactivated prior to NEM and $InsP_6$ co-injection, preventing autoprocessing. Thus, all data support a cooperative mechanism for activation, as opposed to the allosteric structural transition model, for pro-enzyme activation and reactivation both in the context of small recombinant CPD proteins (Fig. 6) and the larger recombinant protein $RtxA_{1580-3909}$ (Fig. 7).

In the context of the MARTX holotoxin, we propose that translocation of the ~185-kDa central region of MARTX to the cytosol likely involves at least partial unfolding of the toxin, because translocation of a fully folded protein would require a substantial pore and it has been demonstrated that there is no leakage of cytosol contents due to MARTX (37). Upon CPD translocation, Leu^{3428} moves into the CPD S1 site facilitating high affinity binding of $InsP_6$. Upon binding, $\beta 8\beta 9$ adopts the proper conformation locking the scissile bond into the active site where it is cleaved. Reactivation of CPD then occurs by reoccupation of the S1 site by a new substrate and binding of $InsP_6$. This reactivation results in cleavage at other leucine residues located between the effector domains releasing them from the larger toxin. Secondary structure predictions have shown that these leucines are present in unstructured regions that flank the effectors, explaining the specificity of preferred cleavage sites despite the fact that CPD is apparently a promiscuous protease able to process at almost any exposed leucine. Interestingly, autoprocessing of $RtxA_{1580-3909}$ by its own CPD was more efficient than in *trans* processing by pro-CPD (Fig. 7*A*). This result suggests that close proximity to other processing sites may create the necessary high local concentration of substrate facilitating immediate reoccupation of the active site after departure of Leu^{3428}–Ala^{3429}. As an extension, this observation might indicate that the effector domains are associated with CPD until processing is completed.

The specificity of this enzyme for leucine is of keen interest. All Clan CD proteases are recognized for having high specificity for processing: for example, eukaryotic caspases at Asp residues, the gingipain K and R proteases of *Porphyromonas gingivalis* at Arg and Lys, respectively, and mammalian legumain at Asn (29, 30). The identification of specificity for Leu of the MARTX CPD indicates that this and similar proteases represent a unique family of proteases, although the structure of the catalytic site clearly places it in Clan CD. All the processing sites defined in this study are shown in Fig. 7*F* in the context of 10 amino acids on each side of the scissile bond. By analyzing these peptides, no conservation of residues other than Leu in P1 site is evident, and mutation analysis indicated any of the two neighboring residues can be altered. Yet, among the preferred processing sites, there is a preference for a small amino acid residue on each side of the Leu (Fig. 7*F*). Extension of this analysis to the Ala-Leu-Gly or Ser-Leu-Gly recognition sites for *Clostridium difficile* Toxin B further indicates preference for small-Leu-small suggesting a conserved preference among this family of proteases (38).

As a conclusion, we have described a new model for delivery of multiple toxin effectors where, upon translocation of the central region into eukaryotic cells, CPD binds $InsP_6$, cleaves itself and then other sites in MARTX, releasing the functional domains that then reach their targets and contribute to pathogenesis of cholera.

Acknowledgments—We thank R. Tungekar, B. Geissler, and C.-H. Luan for their input. We thankfully acknowledge the Keck Biophysics Facility at Northwestern University for use of ITC.

REFERENCES

1. Satchell, K. J. (2007) *Infect. Immun.* **75,** 5079–5084
2. Olivier, V., Haines, G. K., 3rd, Tan, Y., and Satchell, K. J. (2007) *Infect. Immun.* **75,** 5035–5042
3. Olivier, V., Salzman, N. H., and Satchell, K. J. (2007) *Infect. Immun.* **75,** 5043–5051
4. Kudryashov, D. S., Durer, Z. A., Ytterberg, A. J., Sawaya, M. R., Pashkov, I., Prochazkova, K., Yeates, T. O., Loo, R. R., Loo, J. A., Satchell, K. J., and Reisler, E. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 18537–18542
5. Sheahan, K. L., Cordero, C. L., and Satchell, K. J. (2004) *Proc. Natl. Acad. Sci. U.S.A.* **101,** 9798–9803
6. Sheahan, K. L., and Satchell, K. J. (2007) *Cell. Microbiol.* **9,** 1324–1335
7. Prochazkova, K., and Satchell, K. J. (2008) *J. Biol. Chem.* **283,** 23656–23664
8. Sheahan, K. L., Cordero, C. L., and Satchell, K. J. (2007) *EMBO J.* **26,** 2552–2561
9. French, P. J., Bunce, C. M., Stephens, L. R., Lord, J. M., McConnell, F. M., Brown, G., Creba, J. A., and Michell, R. H. (1991) *Proc. Biol. Sci.* **245,** 193–201
10. Jackson, T. R., Hallam, T. J., Downes, C. P., and Hanley, M. R. (1987) *EMBO J.* **6,** 49–54
11. Vallejo, M., Jackson, T., Lightman, S., and Hanley, M. R. (1987) *Nature* **330,** 656–658
12. Collyn, F., Léty, M. A., Nair, S., Escuyer, V., Ben Younes, A., Simonet, M., and Marceau, M. (2002) *Infect. Immun.* **70,** 6196–6205
13. Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R., and Donnelly, M. I. (2002) *Protein Expr. Purif.* **25,** 8–15
14. Kapust, R. B., and Waugh, D. S. (2000) *Protein Expr. Purif.* **19,** 312–318
15. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
16. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) *Acta Crystallogr. D Biol. Crystallogr.* **61,** 458–464
17. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) *Acta Crystallogr. D. Biol. Crystallogr.* **55,** 247–255
18. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2126–2132
19. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
20. Winn, M. D., Ashton, A. W., Briggs, P. J., Ballard, C. C., and Patel, P. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1929–1936
21. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) in *The Proteomics Protocols Handbook* (Walker, J. M., ed) pp. 571–607, Humana Press, Totowa, NJ
22. Gattiker, A., Bienvenut, W. V., Bairoch, A., and Gasteiger, E. (2002) *Proteomics* **2,** 1435–1444
23. Cordero, C. L., Kudryashov, D. S., Reisler, E., and Satchell, K. J. (2006) *J. Biol. Chem.* **281,** 32366–32374
24. Lin, W., Fullner, K. J., Clayton, R., Sexton, J. A., Rogers, M. B., Calia, K. E., Calderwood, S. B., Fraser, C., and Mekalanos, J. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96,** 1071–1076
25. Lupardus, P. J., Shen, A., Bogyo, M., and Garcia, K. C. (2008) *Science* **322,** 265–268
26. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F., and Jones, D. T. (2004) *J. Mol. Biol.* **337,** 635–645
27. Eichinger, A., Beisel, H. G., Jacob, U., Huber, R., Medrano, F. J., Banbula, A., Potempa, J., Travis, J., and Bode, W. (1999) *EMBO J.* **18,** 5453–5462
28. Wilson, K. P., Black, J. A., Thomson, J. A., Kim, E. E., Griffith, J. P., Navia, M. A., Murcko, M. A., Chambers, S. P., Aldape, R. A., Raybuck, S. A., and Livingston, D. J. (1994) *Nature* **370,** 270–275
29. Chen, J. M., Rawlings, N. D., Stevens, R. A., and Barrett, A. J. (1998) *FEBS Lett.* **441,** 361–365

30. Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J., and Barrett, A. J. (2008) *Nucleic Acids Res.* **36,** D320–D325
31. Lo, M. C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Ellestad, G. (2004) *Anal. Biochem.* **332,** 153–159
32. Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P., and Salemme, F. R. (2001) *J. Biomol. Screen* **6,** 429–440
33. Siebold, C., Berrow, N., Walter, T. S., Harlos, K., Owens, R. J., Stuart, D. I., Terman, J. R., Kolodkin, A. L., Pasterkamp, R. J., and Jones, E. Y. (2005) *Proc. Natl. Acad. Sci. U.S.A.* **102,** 16836–16841
34. Powers, J. C., Asgian, J. L., Ekici, O. D., and James, K. E. (2002) *Chem. Rev.* **102,** 4639–4750
35. Viboud, G. I., and Bliska, J. B. (2005) *Annu. Rev. Microbiol.* **59,** 69–89
36. Evdokimov, A. G., Anderson, D. E., Routzahn, K. M., and Waugh, D. S. (2001) *J. Mol. Biol.* **312,** 807–821
37. Fullner, K. J., and Mekalanos, J. J. (2000) *EMBO J.* **19,** 5315–5323
38. Rupnik, M., Pabst, S., Rupnik, M., von Eichel-Streiber, C., Urlaub, H., and Söoling, H. D. (2005) *Microbiology* **151,** 199–208