Supporting Information:

Substrate profiling and high resolution co-complex crystal structure of a secreted C11 protease conserved across commensal bacteria

Emily J. Roncase[1§], Clara Moon[1§], Sandip Chatterjee[1], Gonzalo E. González-Páez[1], Charles S. Craik[2], Anthony J. O'Donoghue[3], Dennis W. Wolan[1*]

[1]Department of Molecular Medicine, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037

[2]Department of Pharmaceutical Chemistry, University of California, San Francisco, 600 16th Street, San Francisco, CA 94158

[3]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093

[§]Co-contributors

[*]Corresponding Author e-mail: wolan@scripps.edu

Table of Contents

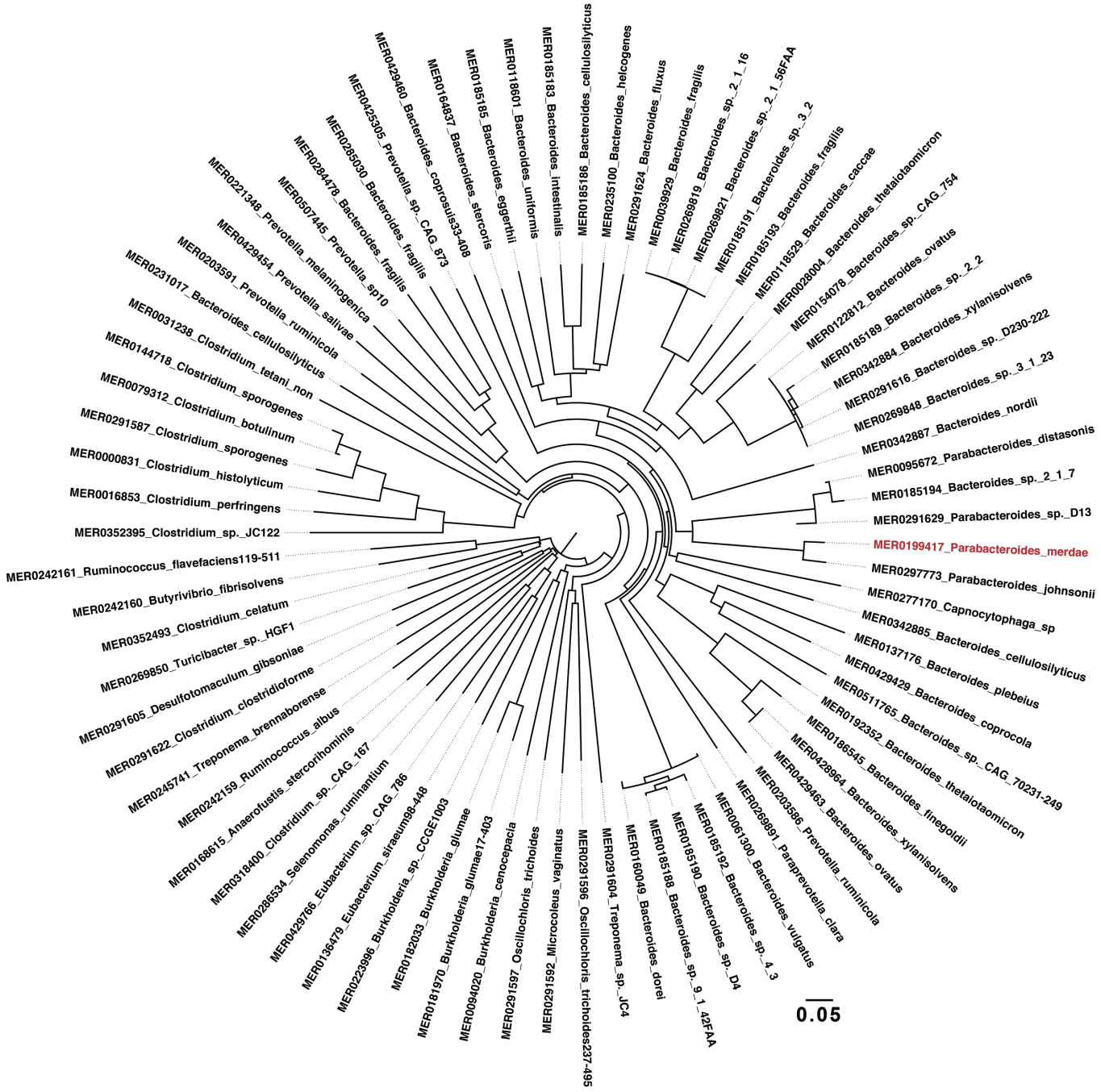

Figure S1. Conservation of *P. merdae* PmC11 cysteine protease (red) across a small sampling of commensal bacterial strains. MEROPS[1] entry numbers for each C11 protease are shown.

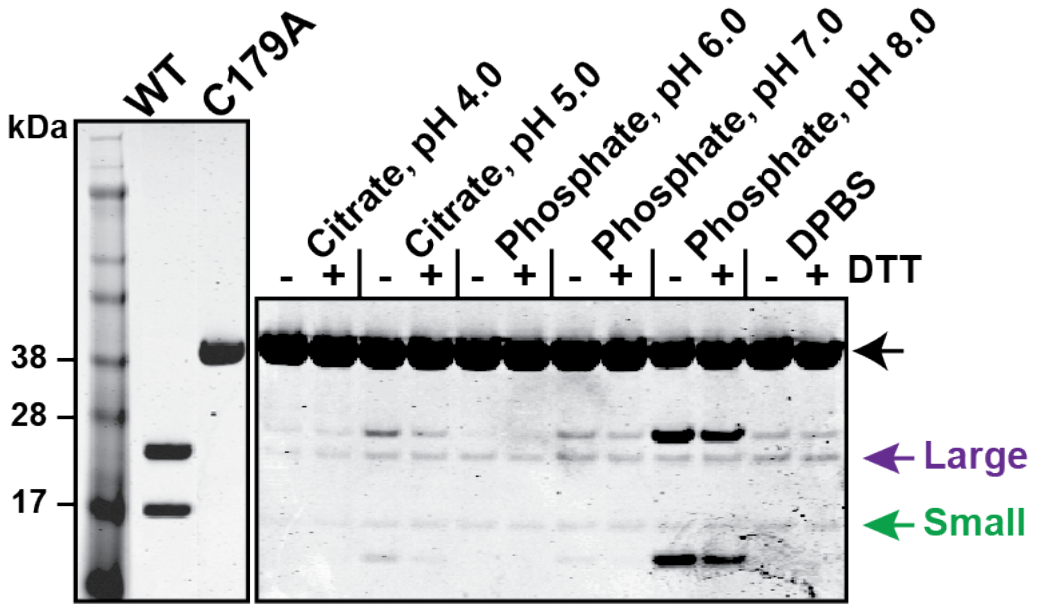

Figure S2. Incubation of 5 μM inactive zymogen C179A PmC11 and 50 nM active WT PmC11 for 3 hr at 37 °C demonstrates proteolytic activity of the exogenously purified PmC11. The protease is most active at pH 8.0 and addition of 10 mM DTT to the activity buffer does not affect catalytic efficiency. Activity buffers consist of 10 mM citrate or phosphate, 50 mM NaCl, 0.1% CHAPS with and without 10 mM DTT. Large and small domains of the WT PmC11 are denoted relative to the large and small cleavage products of the C179A mutant.

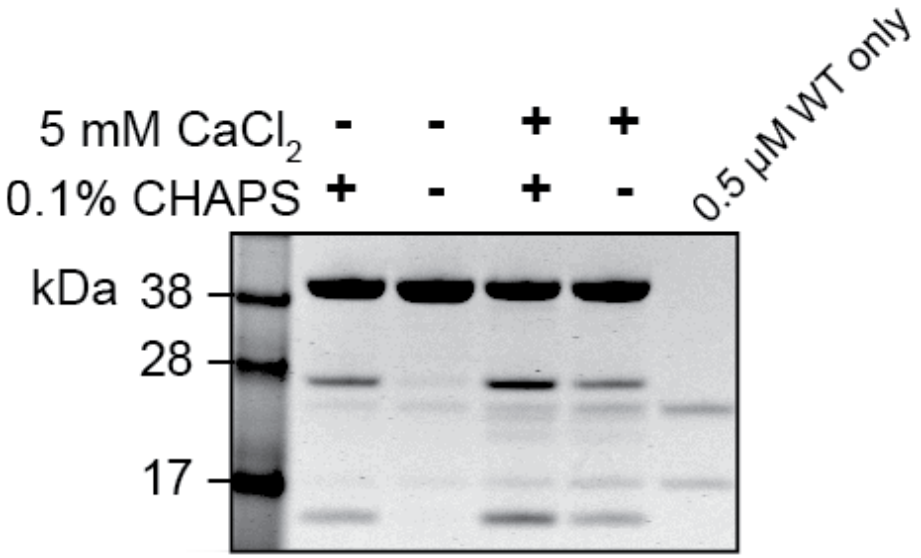

Figure S3. Similar to Figure S1, incubation of 5 μM inactive zymogen C179A PmC11 and 50 nM active WT PmC11 demonstrates proteolytic activity is improved by the presence of $CaCl_2$ and 0.1% CHAPS to the buffer consisting of 10 mM phosphate, pH 8.0, 50 mM NaCl, and 2 mM DTT.

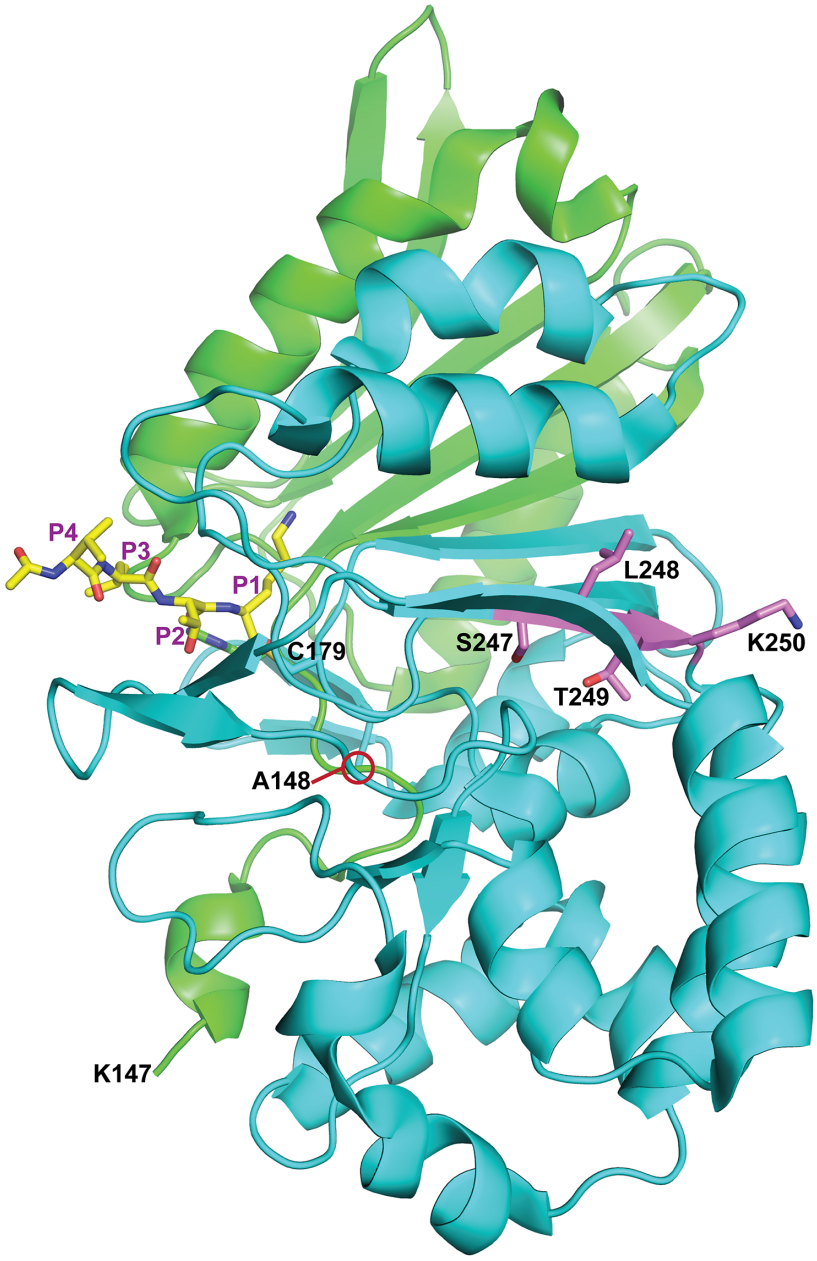

Figure S4. Proposed cleavage site of C179A protein by WT PmC11 with sequence SLTK^{250} in magenta is on the opposite side of the protein relative to the active site. The mature cleavage site between Lys147 and Ala148 are labeled.

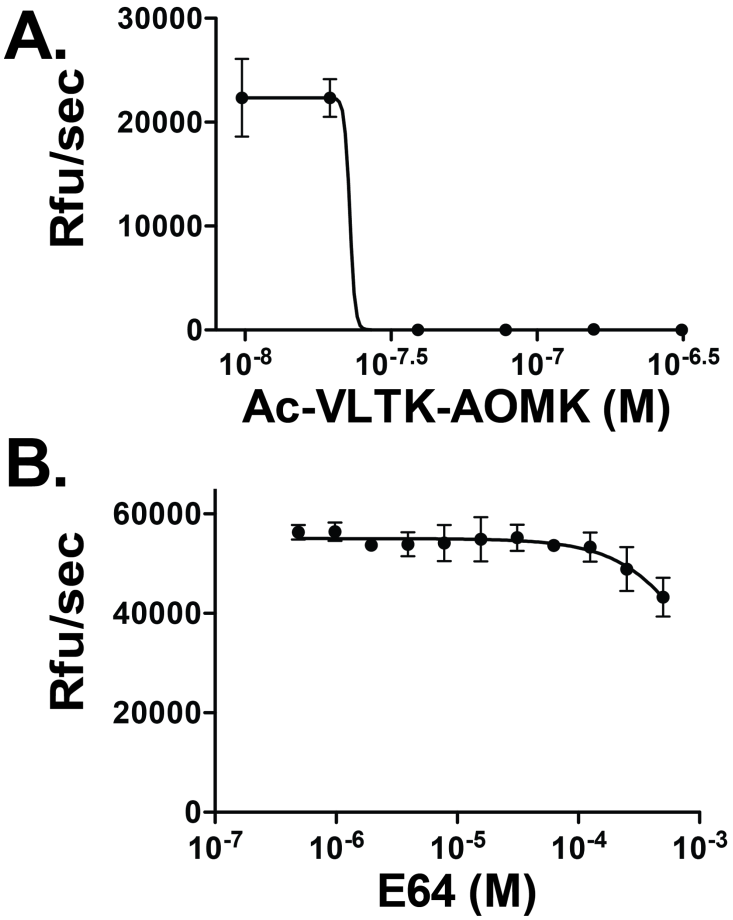

Figure S5. A. Inhibition of 25 nM PmC11 by Ac-VLTK-*AOMK* (**A**) and E64 (**B**) shows the strong potency of the inhibitor with an IC50 <25 nM and only ~30% inhibition, respectively.

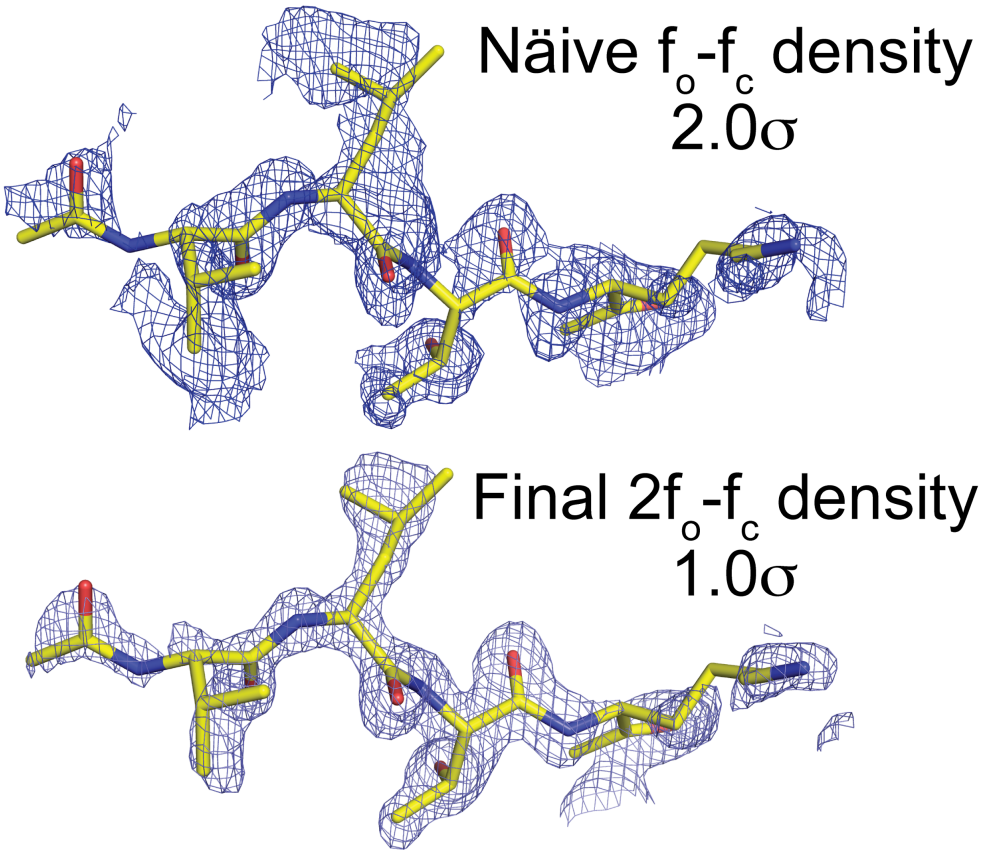

Figure S6. Naïve density map for Ac-VLTK after molecular replacement with model PDB 3UWS. Final density map is shown below.

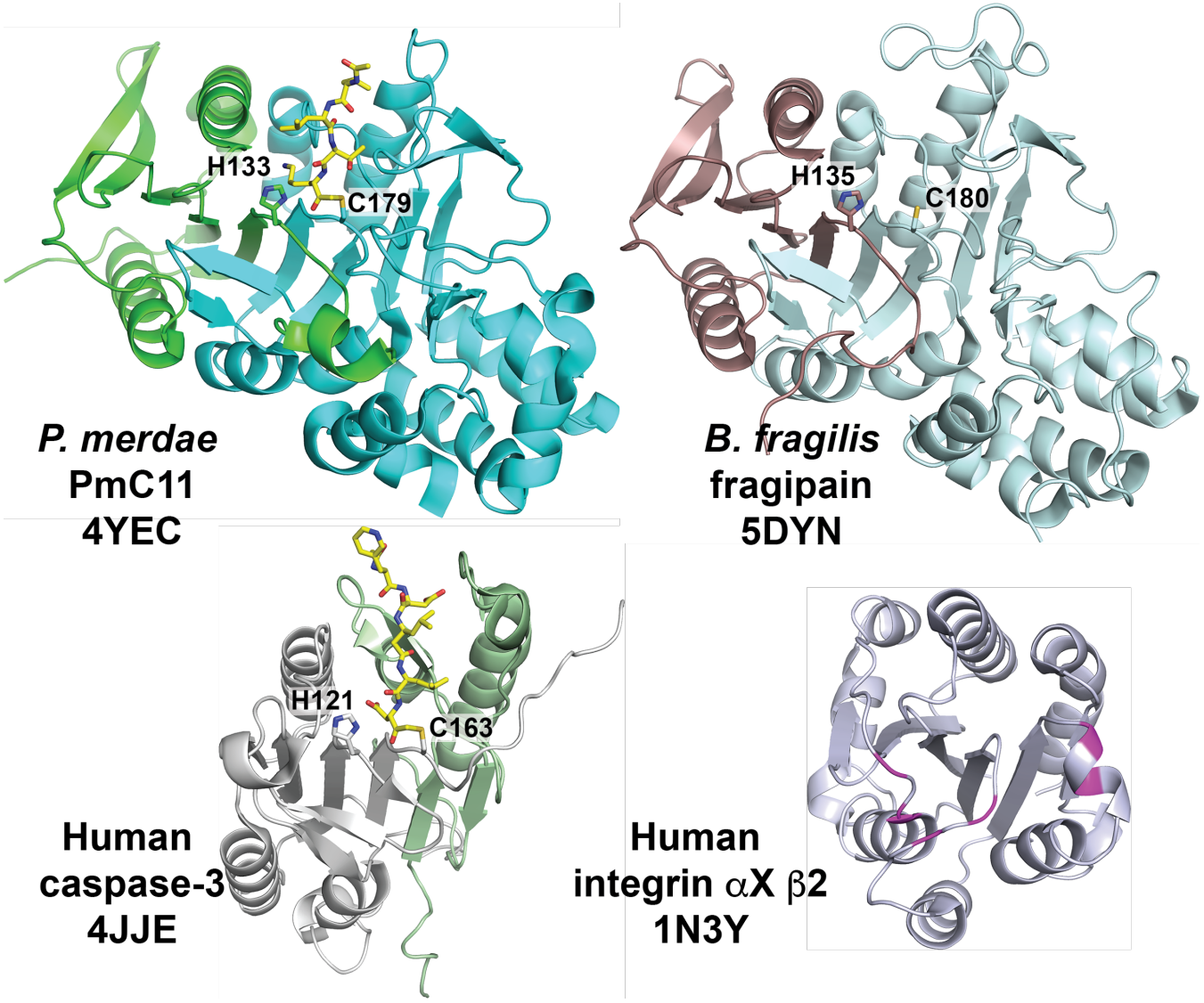

Figure S7. DALI search results demonstrate PmC11 has structural similarities with other proteases, including *B. fragilis* Cys protease fragipain (PDB ID: 5DYN)[2], human caspases (PDB ID: 4JJE)[3] as well as human integrin (PDB ID: 1N3Y)[4]. The commonalities of the structures consists of a mixed β-sheet protected by α-helices above and below the plane of the sheet. The active site His and Cys residues of the proteases superimpose and have highly conserved spatial orientation. All active site His and Cys residues are labeled and the putative binding site residues of human integrin, located on the same side of the β-sheet as the active sites of the cysteine proteases, are highlighted in magenta.

Table S1. PmC11:Ac-VLTK-*AOMK* co-complex X-ray data processing and structure refinement statistics

Structure	PmC11:Ac-VLTK-*AOMK*
PDB ID	4YEC
Space group	$P2_1$
Unit Cell Parameters (a,b,c) (Å)	38.89, 107.34, 40.94
Unit Cell Angles (z,y,z)	90.0, 116.59, 90.0
Data Processing	
Resolution range (Å) (outer shell)	36.6-1.12 (1.14-1.12)
Unique reflections	90,762 (846)
Completeness (%)	79.0 (14.6)
Redundancy	3.0 (1.2)
R_{meas} (%)[1]	12.6 (46.2)
R_{merge} (%)[2]	16.0 (29.4)
$R_{p.i.m.}$ (%)[3]	6.7 (32.7)
Average I/Average σ(I)	9.3 (2.1)
Refinement	
Resolution range (Å)	36.6-1.12 (1.15-1.12)
No. reflections[4] (test set)	90,707 (1,310)
R_{cryst} (%)[5]	15.8 (25.2)
R_{free} (%)[5]	19.1 (39.6)
Protein atoms / Waters	2890 / 284
CV[6] coordinate error (Å)	0.10
Rmsd bonds (Å) / angles (º)	0.011 / 1.39
B-values protein/waters/peptide ligand (Å^2)	22.6 / 38.1 / 39.5
Ramachandran Statistics (%)	
Most favored	96.4
Additional allowed	3.3
Generously allowed	0.3

[1] $R_{meas} = \{\Sigma_{hkl}[N/(N-1)]1/2\Sigma_i|I_{i(hkl)} - \langle I_{(hkl)}\rangle|\}/\Sigma_{hkl}\Sigma_i I_{i(hkl)}$, where $I_{i(hkl)}$ are the observed intensities, $\langle I_{(hkl)}\rangle$ are the average intensities and N is the multiplicity of reflection hkl. [2] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_{i(hkl)} - \langle I_{(hkl)}\rangle|/\Sigma_{hkl}\Sigma_i I_{i(hkl)}$ where $I_{i(hkl)}$ is the i^{th} measurement of reflection h and $\langle I_{(hkl)}\rangle$ is the average measurement value. [3] $R_{p.i.m.}$ (precision-indicating R_{merge}) = $\Sigma_{hkl}[1/(N_{hkl}-1)]^{1/2}\Sigma_i|I_{i(hkl)} - \langle I_{(hkl)}\rangle|/\Sigma_{hkl}\Sigma_i I_{i(hkl)}$. [4] Reflections with I > 0 were used for refinement. [5-7] [5] $R_{cryst} = \Sigma_h||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$, where F_{obs} and F_{calc} are the calculated and observed structure factor amplitudes, respectively. R_{free} is R_{cryst} with 5.0% test set structure factors. [6] Cross-validated (CV) Luzzati coordinate errors.

Supporting Information References

(1) Rawlings, N. D., Barrett, A. J., and Finn, R. (2016) Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res. 44*, D343–50.
(2) Choi, V. M., Herrou, J., Hecht, A. L., Teoh, W. P., Turner, J. R., Crosson, S., and Bubeck Wardenburg, J. (2016) Activation of Bacteroides fragilis toxin by a novel bacterial protease contributes to anaerobic sepsis in mice. *Nat. Med. 22*, 563–567.
(3) Vickers, C. J., González-Páez, G. E., and Wolan, D. W. (2013) Selective detection of caspase-3 versus caspase-7 using activity-based probes with key unnatural amino acids. *ACS Chem. Biol. 8*, 1558–1566.
(4) Vorup-Jensen, T., Ostermeier, C., Shimaoka, M., Hommel, U., and Springer, T. A. (2003) Structure and allosteric regulation of the alpha X beta 2 integrin I domain. *Proc. Nat. Acad. Sci. USA 100*, 1873–1878.
(5) Weiss, M. S., and Hilgenfeld, R. (1997) On the use of the merging R factor as a quality indicator for X-ray data. *J. Appl. Crystallogr. 30*, 203–205.
(6) Weiss, M. S. (2001) Global indicators of X-ray data quality. *J. Appl. Crystallogr. 34*, 130–135.
(7) Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality. *Science 336*, 1030–1033.