

# Simplified, Enhanced Protein Purification Using an Inducible, Autoprocessing Enzyme Tag

Aimee Shen<sup>1</sup>, Patrick J. Lupardus<sup>2</sup>, Montse Morell<sup>1</sup>, Elizabeth L. Ponder<sup>1</sup>, A. Masoud Sadaghiani<sup>3</sup>, K. Christopher Garcia<sup>2,4</sup>, Matthew Bogyo<sup>1\*</sup>

**1** Department of Pathology, Stanford School of Medicine, Stanford, California, United States of America, **2** Department of Molecular and Cellular Physiology, Stanford School of Medicine, Stanford, California, United States of America, **3** Department of Systems and Chemical Biology, Stanford School of Medicine, Stanford, California, United States of America, **4** Howard Hughes Institute, Stanford School of Medicine, Stanford, California, United States of America

## Abstract

We introduce a new method for purifying recombinant proteins expressed in bacteria using a highly specific, inducible, self-cleaving protease tag. This tag is comprised of the *Vibrio cholerae* MARTX toxin cysteine protease domain (CPD), an autoprocessing enzyme that cleaves exclusively after a leucine residue within the target protein-CPD junction. Importantly, *V. cholerae* CPD is specifically activated by inositol hexakisphosphate (InsP<sub>6</sub>), a eukaryotic-specific small molecule that is absent from the bacterial cytosol. As a result, when His<sub>6</sub>-tagged CPD is fused to the C-terminus of target proteins and expressed in *Escherichia coli*, the full-length fusion protein can be purified from bacterial lysates using metal ion affinity chromatography. Subsequent addition of InsP<sub>6</sub> to the immobilized fusion protein induces CPD-mediated cleavage at the target protein-CPD junction, releasing untagged target protein into the supernatant. This method condenses affinity chromatography and fusion tag cleavage into a single step, obviating the need for exogenous protease addition to remove the fusion tag(s) and increasing the efficiency of tag separation. Furthermore, in addition to being timesaving, versatile, and inexpensive, our results indicate that the CPD purification system can enhance the expression, integrity, and solubility of intractable proteins from diverse organisms.

**Citation:** Shen A, Lupardus PJ, Morell M, Ponder EL, Sadaghiani AM, et al. (2009) Simplified, Enhanced Protein Purification Using an Inducible, Autoprocessing Enzyme Tag. PLoS ONE 4(12): e8119. doi:10.1371/journal.pone.0008119

**Editor:** Wenqing Xu, University of Washington, United States of America

**Received:** October 8, 2009; **Accepted:** November 5, 2009; **Published:** December 2, 2009

**Copyright:** © 2009 Shen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by a Burroughs Wellcome Foundation grant and NIH grants R01 AI078947 and R01 EB005011 to M.B., the Damon Runyon Cancer Research Fellowship to P.J.L., a Keck Foundation and Howard Hughes Medical Institute grant to K.C.G., a Stanford Dean's Fellowship to A.S., and a Beatriu de Pinós of Agaur Fellowship to M.M. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** A.S., P.J.L., K.C.G., and M.B. are listed as inventors on a provisional patent application describing the CPD purification system technology. This patent will not alter the authors' adherence to PLoS ONE policies on sharing data and materials. Materials and information associated with the authors' publication will be freely available to those as requested for the purpose of academic, non-commercial research.

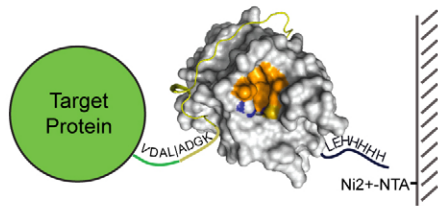
\* E-mail: mbogyo@stanford.edu

## Introduction

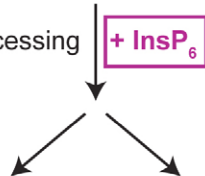
The availability of simple, reliable, and cost-effective methods for recombinant protein purification is critical for the work of high throughput structural and proteomic centers and many individual researchers alike. While the addition of affinity tags such as poly-His and glutathione transferase (GST) to target proteins has greatly simplified purification strategies, it is often difficult to obtain soluble recombinant protein [1]. As a result, intractable affinity-tagged target proteins are often fused to small proteins such as NusA and SUMO to improve their solubility, expression, and stability [2].

Since these tags can alter the biological activity of target proteins and interfere with protein crystallization studies, many biological and biomedical applications require that the tag be removed from the target protein. Most commonly used methods involve the addition of exogenous site-specific proteases to cleave the affinity tag off the target protein at engineered sites [2]. Unfortunately, high levels of endoprotease must often be applied for extended periods of time, and this can result in undesirable cleavages within the target protein. Furthermore, these endoproteases are costly, often exhibit poor solubility, and require the inclusion of additional chromatography steps to remove the exogenous protease.

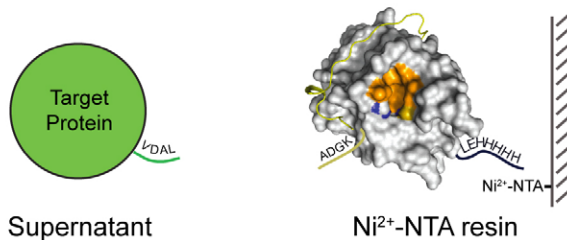
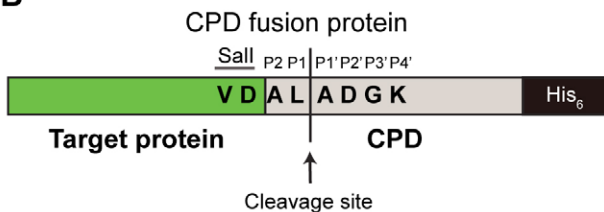
To circumvent these disadvantages, we have developed an on-bead cleavage purification system in which a site-specific affinity-tagged protease is fused directly to the target protein. This approach condenses affinity purification, cleavage, and tag separation into a single step, simplifying protein purification procedures and increasing purification yields. The key element of this purification method is the *Vibrio cholerae* MARTX toxin cysteine protease domain (CPD) [3]. The CPD exhibits several properties that facilitate its development into an inducible, autocleaving protease tag. First, the CPD is a highly specific protease that cleaves exclusively after Leu residues [4]. Second, the CPD is inducible, as it is specifically activated by the eukaryotic-specific small molecule inositol hexakisphosphate (InsP<sub>6</sub>) [5,6]. Since InsP<sub>6</sub> is absent from bacterial cells [7,8], full-length CPD-His<sub>6</sub> fusion proteins can be purified from bacterial lysates in a protease-inactive form using imidazole affinity chromatography (IMAC). Addition of InsP<sub>6</sub> to an immobilized, C-terminally His<sub>6</sub>-tagged fusion protein induces autoprocessing at the P1 Leu cleavage site (P1 refers to the residue N-terminal to the scissile bond), which is located at the target protein-CPD junction (Figure 1). This processing event releases the untagged target protein into the supernatant, while the C-terminally His<sub>6</sub>-tagged CPD remains immobilized on the Ni<sup>2+</sup>-NTA resin. Third, as an

**A**1. Bind CPD fusion protein to  $\text{Ni}^{2+}$ -NTA resin

2. Induce autoprocessing



3. Isolate untagged protein in supernatant

**B**

**Figure 1. CPD fusion protein purification system.** (A) Schematic of target protein purification using the CPD, described in detail in the text. (B) Schematic of CPD fusion protein. The P4 and P3 residues Val and Asp, respectively, encoded by the SalI site, and the remaining P2-P4' residues contained within the CPD are shown. Prime positions refer to residues C-terminal to the autocleavage site, which is demarcated as a black vertical line. The composition of residues appended to the C-terminus of target proteins following autoprocessing can vary between one and four residues as described in Figure 2. At present, the CPD system functions as a C-terminal fusion to target proteins and thus complements existing methods in which the affinity tag can only be applied as an N-terminal fusion [18].  
doi:10.1371/journal.pone.0008119.g001

autoprocessing enzyme, the CPD exhibits poor transcleavage efficiency [4,5]. This property should limit fusion protein cleavage to the CPD-target protein junction and permit the high fidelity removal of the His<sub>6</sub>-CPD tag from the target protein.

In this report, we demonstrate using a variety of target proteins that this novel purification system combines the simplicity of one-step purification systems [9,10] with many of the advantages of affinity tags [2] in that it can increase the expression, solubility, and integrity of target proteins. Thus, this method facilitates the rapid purification of both soluble and intractable, recombinant, untagged proteins, suggesting that it will have widespread utility in

individual research labs and high-throughput structural and proteomic centers.

## Results

### Development of the One-Step CPD Purification System

In order to produce CPD fusion proteins, we first constructed CPD expression vectors (pET-CPD expression vectors) using the pET expression vector backbone. DNA encoding the CPD was cloned into the SalI restriction site (Figure 2) such that the fusion protein produced upon IPTG induction of *E. coli* harboring the pET-CPD<sub>SalI</sub> vector carries the P2-P1 residues of the native CPD (Ala-Leu, respectively) and the P4-P3 residues encoded by the SalI site (Val-Asp, respectively) (Figures 1 and 2). The P1 residue refers to the amino acid N-terminal to the scissile bond, while the residue N-terminally adjacent to the P1 residue is termed P2, and so on. When InsP<sub>6</sub> is added to induce CPD-mediated autocleavage of the fusion protein, the untagged target protein is released from the resin and carries four additional C-terminal residues (Val-Asp-Ala-Leu); the His<sub>6</sub>-tagged CPD remains bound to the resin (Figure 1). The Val-Asp-Ala-Leu C-terminal addition can be reduced to two amino acids (Glu-Leu) by cloning into the SacI site, or to a single amino acid (Leu) by cloning into the BamHI site and adding a Leu codon to the 3' cloning primer (Figure 2).

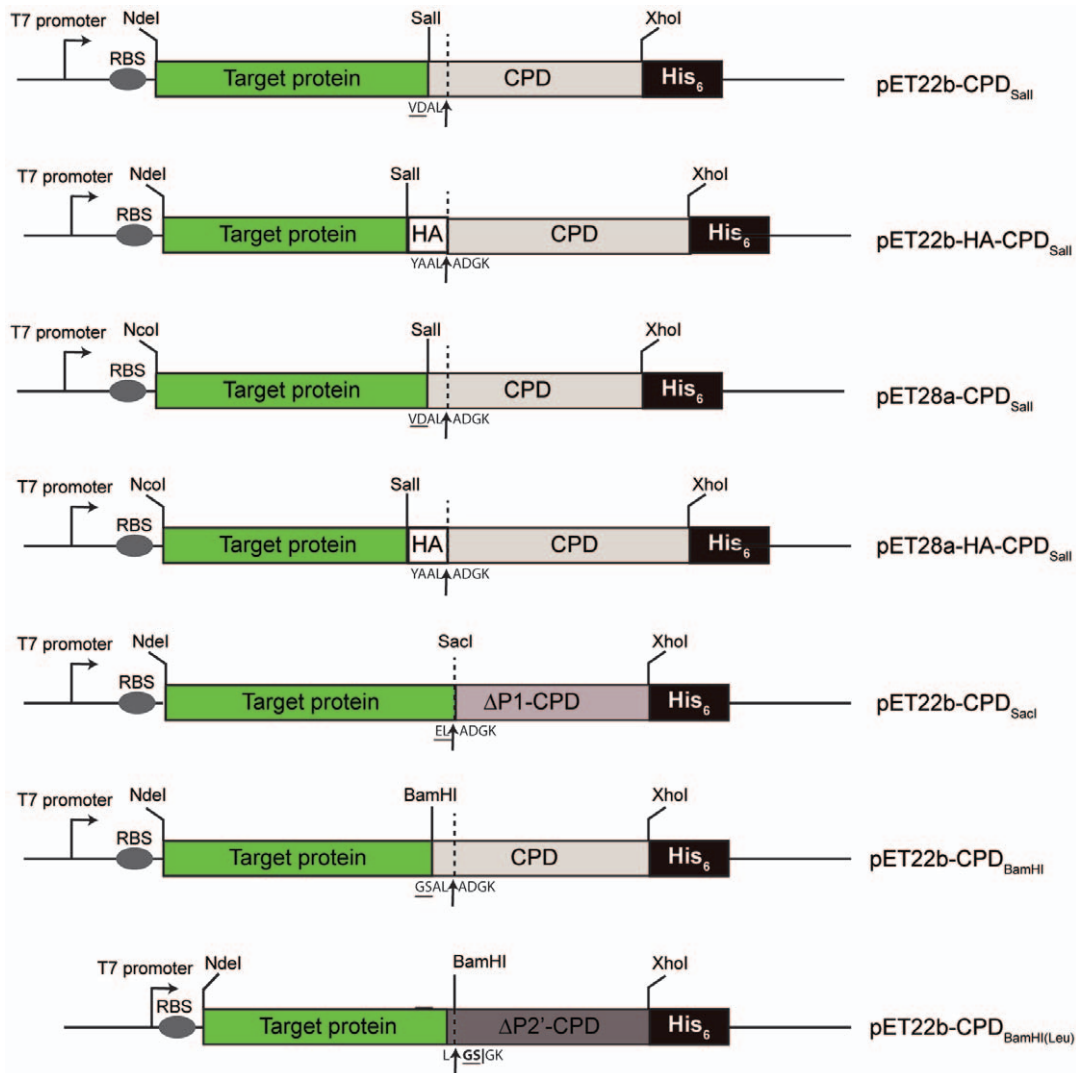
To demonstrate the feasibility of this system, we first expressed and purified green fluorescent protein (GFP) as a fusion to CPD-His<sub>6</sub> using IMAC. As anticipated, addition of increasing amounts of InsP<sub>6</sub> stimulated the release of GFP from the Ni<sup>2+</sup>-NTA resin in a dose-dependent manner (Figures 3A and B), while the His<sub>6</sub>-tagged CPD remained bound to the Ni<sup>2+</sup>-NTA agarose beads (bead eluate, Figure 3A).

### The CPD Cleaves Exclusively at the Fusion Protein Junction

We have previously shown that *V. cholerae* CPD is positioned to undergo autocleavage at a proximal N-terminal leucine and that it exhibits significantly reduced transcleavage efficiency [4,5], which should limit its ability to cleave target proteins at heterologous sites. Indeed, mutation of the P1 Leu to an Ile residue is sufficient to prevent CPD-mediated transcleavage, a finding that is explained by the observation that the P1 Leu residue fits snugly into the S1 substrate binding pocket in the crystal structure of the P1 Leu azapoxide inhibitor modified *V. cholerae* CPD [4]. Nevertheless, since other site-specific proteases used to remove fusion tags have been observed to cleave target proteins at secondary sites [2], we examined whether the CPD would spuriously cleave target proteins. Specifically, we tested whether the CPD would cleave an intrinsically disordered protein after Leu residues within the target protein. We used the intracellular domain (ICD) of the cytokine receptor gp130 as a test substrate, since it is unstructured in solution by NMR [11] and contains multiple Leu residues that might serve as cleavage substrates [12]. The ICD-CPD-His<sub>6</sub> fusion protein was expressed and purified from *E. coli* lysates using IMAC, and CPD-mediated cleavage of the immobilized fusion protein was activated by InsP<sub>6</sub> addition. As shown in Figure 4, autoprocessing occurred exclusively at the ICD-CPD interdomain junction, with a single protein equivalent to the size of His<sub>6</sub>-tagged ICD being released into the supernatant fraction. These results strongly suggest that the CPD will not promiscuously cleave target proteins.

### Fusion of Target Proteins to the CPD Can Increase Their Expression and Purity

We noticed that the expression of the ICD-CPD-His<sub>6</sub> fusion protein was at least three-fold higher than the ICD-His<sub>6</sub> protein in



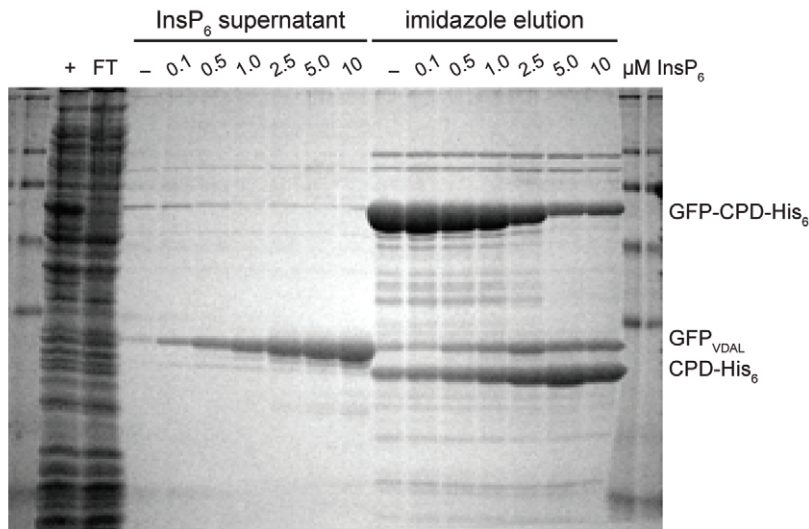
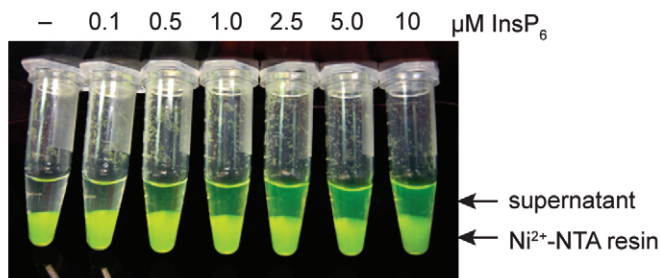
**Figure 2. Schematic of pET-CPD expression vectors.** Bent arrow, T7 promoter, Oval (RBS), ribosome binding site, green rectangle, target protein, grey rectangle, CPD, *V. cholerae* MARTX (aa. 3440–3650), darker grey rectangle,  $\Delta$ P1-CPD, *V. cholerae* MARTX (aa. 3442–3650), darkest grey rectangle,  $\Delta$ P2'-CPD, *V. cholerae* MARTX (aa. 3444–3650), black rectangle, His<sub>6</sub>-tag, white rectangle, HA-tag. The dotted vertical line and arrow indicate the CPD cleavage site. Residues added onto the C-terminus of the target protein following CPD-mediated cleavage, and the relevant restriction sites are shown (residues encoded by the restriction sites that are appended to the C-terminus of target proteins are underlined). The composition of the residues added to the C-terminus of the target protein can be varied depending on the cloning site and pET-CPD vector used. It should be noted that the P1 Leu shown for pET22b-CPD<sub>BamHI-Leu</sub> must be encoded in the 3' cloning primer of the target gene (i.e. add a Leu codon to the end of the target insert). Both pET22b and pET28a vector backbones were used to construct the CPD expression vectors.

doi:10.1371/journal.pone.0008119.g002

*E. coli* lysates (Figure 4, compare + lanes). This result suggested that the CPD might generally enhance target protein expression and/or solubility levels. To test this hypothesis, we compared the expression and solubility of CPD fusions to several other target proteins carrying either a His<sub>6</sub>-tag and/or GST-fusion tag (Figures 5–8 and Table 1). In all cases, the presence of the CPD-His<sub>6</sub> fusion tag increased the expression and solubility of target proteins. For example, fusion of the CPD-His<sub>6</sub> tag to biotin ligase (BirA) from *E. coli* (BirA-CPD-His<sub>6</sub>) raised BirA expression levels by three-fold over the GST-BirA construct [13] (Figure 5 and Table 1).

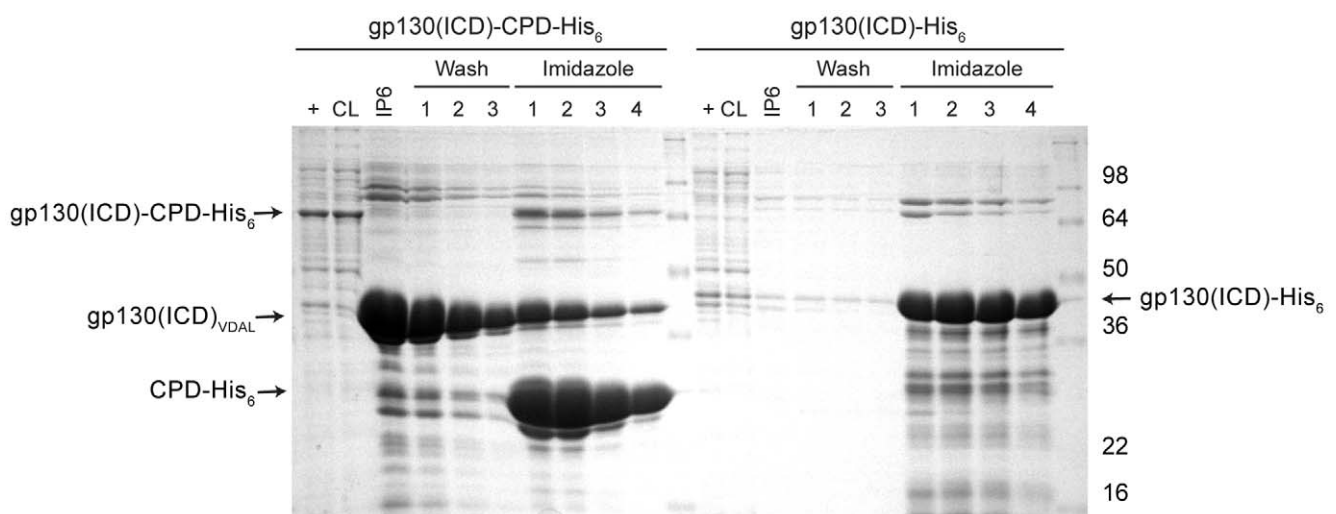
The CPD purification system also enhanced the expression and purity of a previously uncharacterized SUMO/Sentrin-specific peptidase 1 (SEN1) from the parasitic pathogen *Plasmodium falciparum*, the causative agent of malaria (Figure 6) [14]. Although

PfSEN1 carrying an N-terminal His<sub>6</sub>-tagged can be readily expressed and purified from *E. coli*, a number of contaminating bands are present, and the N-terminal His<sub>6</sub>-tag must be removed by the addition of thrombin followed by multiple chromatography steps (Table 2). In contrast, when PfSEN1 is expressed as a fusion to CPD-His<sub>6</sub> and released as untagged PfSEN1 upon InsP<sub>6</sub> addition, only one minor contaminant co-purifies with PfSEN1 (Figure 6B). This variant is easily removed using gel filtration chromatography (Figure 6C), and the untagged PfSEN1 is of sufficient purity that we have used it to obtain diffraction-quality crystals (E. Ponder, unpublished results). Notably, although the heterologous expression of *P. falciparum* proteins in *E. coli* is typically challenging [15], we have observed that this system can enhance the expression and purification of other parasite proteins from *P. falciparum* and a related apicomplexan parasite *Toxoplasma gondii*.

**A****B**

**Figure 3. Purification of GFP using the CPD-His<sub>6</sub> tag.** (A) SDS-PAGE analysis of GFP purification using Coomassie stain. GFP-CPD-His<sub>6</sub> bound to Ni<sup>2+</sup>-NTA resin was incubated with increasing amounts of InsP<sub>6</sub> for 2 hrs at 4°C. GFP released into the supernatant was collected (InsP<sub>6</sub> supernatant); Ni<sup>2+</sup>-bound proteins were then eluted from the resin by the addition of 200 mM imidazole (Imidazole elution). Collected fractions were analyzed by SDS-PAGE. (B) Visual analysis of GFP released into the supernatant fraction upon InsP<sub>6</sub> addition to immobilized GFP-CPD-His<sub>6</sub> fusion protein.

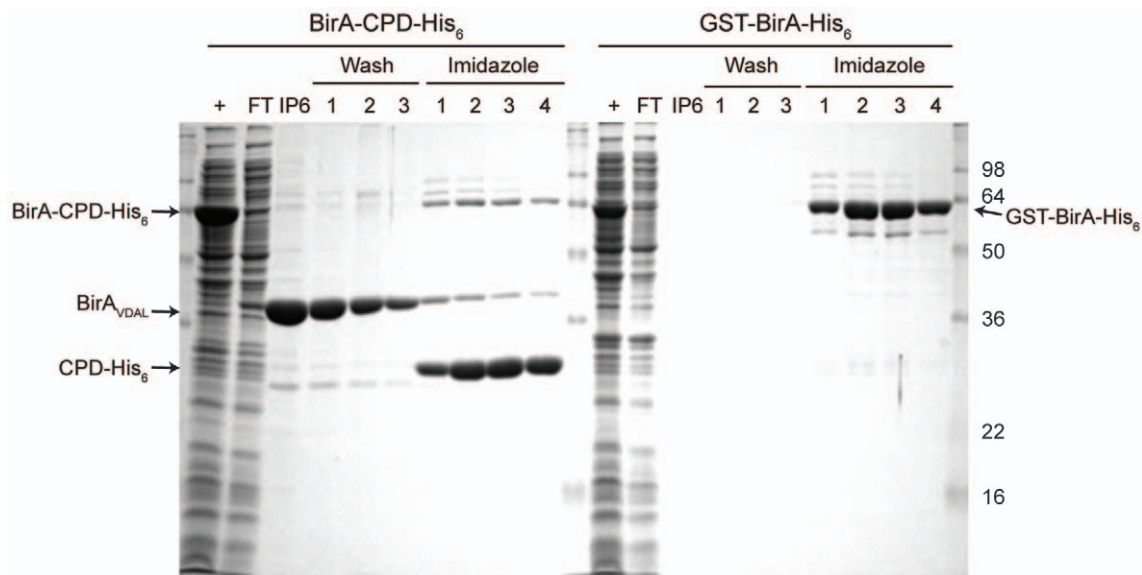
doi:10.1371/journal.pone.0008119.g003



**Figure 4. The CPD does not cleave within an intrinsically unstructured protein.** gp130 intracellular domain (ICD)-CPD-His<sub>6</sub> or gp130(ICD)-His<sub>6</sub> bound to Ni<sup>2+</sup>-NTA resin was incubated with 100 μM InsP<sub>6</sub> for 2 hr at room temperature; the resin was washed four times, followed by elution of Ni<sup>2+</sup>-bound proteins by 200 mM imidazole. Purification fractions were analyzed by SDS-PAGE followed by Coomassie staining. CL, cleared lysate, FT, flowthrough, IP<sub>6</sub>, elution from InsP<sub>6</sub> incubation.

doi:10.1371/journal.pone.0008119.g004





**Figure 5. The CPD improves biotin ligase expression relative to a GST tag.** SDS-PAGE analysis of purifications of biotin ligase (BirA, 35 kDa) fused to either CPD-His<sub>6</sub> or GST-His<sub>6</sub> using Coomassie stain. His<sub>6</sub>-tagged proteins bound to the Ni<sup>2+</sup>-NTA resin were incubated with 50  $\mu$ M InsP<sub>6</sub> for 1 hr at room temperature, and the resin was washed three times, followed by elution of Ni<sup>2+</sup>-bound proteins by 200 mM imidazole. CL, cleared lysate, +, IPTG induced culture, FT, flowthrough, IP6, elution from InsP<sub>6</sub> incubation.  
doi:10.1371/journal.pone.0008119.g005

### Fusion of Target Proteins to the CPD Can Improve Their Stability and Solubility

In addition to augmenting the expression of target proteins, CPD-His<sub>6</sub> fusions protected target proteins from proteolytic degradation. This effect was observed when the CRAC-activation domain (CAD) of the ER calcium sensor STIM1 was fused to the CPD (Figure 7). CAD is a small 107 aa polypeptide that activates Ca<sup>2+</sup> release-activated Ca<sup>2+</sup> (CRAC) channels by binding to the CRAC channel protein Orail [16]. Until now, large-scale expression and purification of this important regulatory domain has proven difficult due to its apparent instability even when fused to GST (Figure 7). However, using the CPD system, we were able to obtain significant quantities of a CAD-containing polypeptide (CAD128), which has subsequently been used in high-throughput screens for Orail-CAD binding partners (A.M. Sadaghiani).

Finally, the CPD purification system also increased the solubility of difficult-to-express proteins. Fusion of the mouse macrophage metalloelastase (MMP12) to CPD-His<sub>6</sub> facilitated its purification from the soluble fraction of *E. coli* lysates, whereas His<sub>6</sub>-tagged MMP12 remained largely insoluble (Figure 8A). The currently used method for purification of His<sub>6</sub>-tagged MMP12 is a laborious procedure that requires solubilization of MMP12 inclusion bodies, refolding over multiple days, followed by anion and cation exchange chromatography [17]. The CPD purification system dramatically simplifies this purification procedure, allowing soluble, active MMP-12 to be isolated in approximately 7 hours (Figs. 8B and C, Table 3). We have used this improved purification protocol to rapidly express, purify and analyze MMP12 mutant proteins.

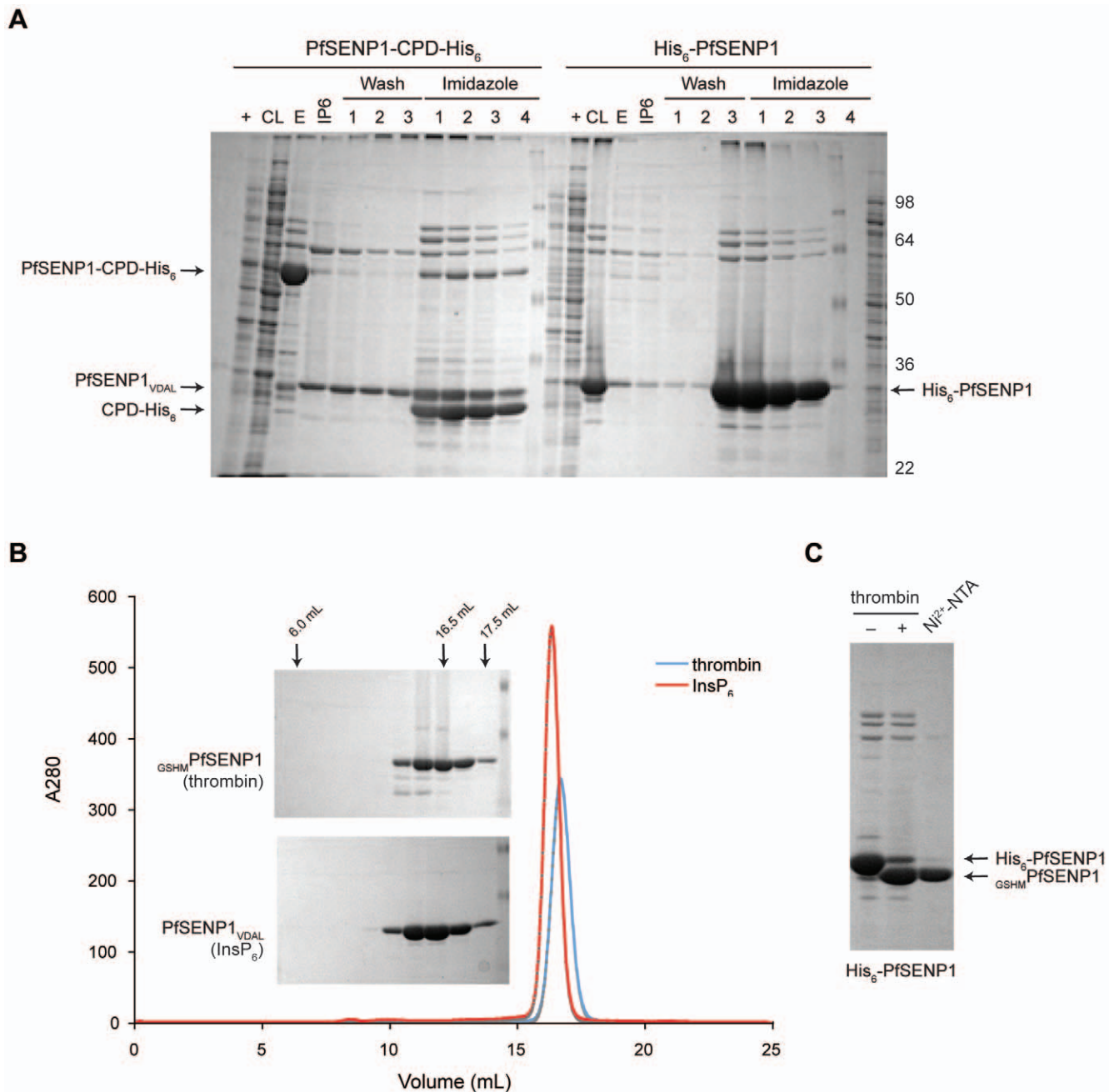
### Discussion

We have developed a novel one-step purification system that accelerates untagged recombinant protein purification from bacterial systems. By directly fusing an affinity-tagged, site-specific protease to a target protein, the CPD system ensures rapid and efficient removal of the fusion tag in a cost-effective manner. As a result, the CPD system overcomes many of the disadvantages

associated with the exogenous addition of site-specific proteases, like thrombin and TEV protease, to remove fusion tags. These disadvantages can include their expense, generally low activity [2,18], sensitivity to buffer conditions, and cleavage of target proteins at spurious sites [2]. In contrast, the CPD rapidly completes tag removal within two hours of addition (Figures 3–8), since the CPD is present at a 1:1 ratio to the target protein and poised to undergo the autocleavage reaction [5]. Furthermore, the responsiveness of the protease specifically to InsP<sub>6</sub> provides the user with complete control over the timing and conditions of fusion tag removal, while the autoprocessing nature of the CPD confers a high degree of specificity to fusion tag removal [4,5]. Specifically, the protease is poised to undergo autocleavage upon InsP<sub>6</sub> addition and exhibits poor transcleavage efficiency, as evidenced by the lack of CPD-mediated cleavage within any of the target proteins tested (Figures 3–8), including an intrinsically unstructured protein (Figure 4).

While purification systems based on fusing a protease to target proteins have previously been developed [9,10], our demonstration that the CPD can enhance the expression, solubility, and stability of target proteins (Figures 4–8) suggests that the CPD system likely represents an improvement over existing methods like the intein-chitin-binding-domain (CBD) [9,10] and sortase-His<sub>6</sub> one-step purification systems [9]. Although these self-cleaving systems simplify the purification of well-expressed proteins, the large size of the intein-CBD fusion tag can decrease target protein solubility [2], while sortase-His<sub>6</sub> fusion tags do not increase target protein solubility [9]. Furthermore, unlike self-cleaving elastin-like polypeptide (ELP) tags [19], CPD fusion proteins do not need to be subjected to temperature cycles, pH shifts, or high salt concentrations, a feature that is critical for the purification of intractable proteins. Based on the properties reported here, the CPD could replace the intein-tag in the self-cleaving-ELP system and potentially improve the solubility of ELP-tagged proteins while retaining their self-cleavability [19].

Indeed, a considerable strength of this method is that the CPD remains active over a wide range of conditions. CPD-mediated

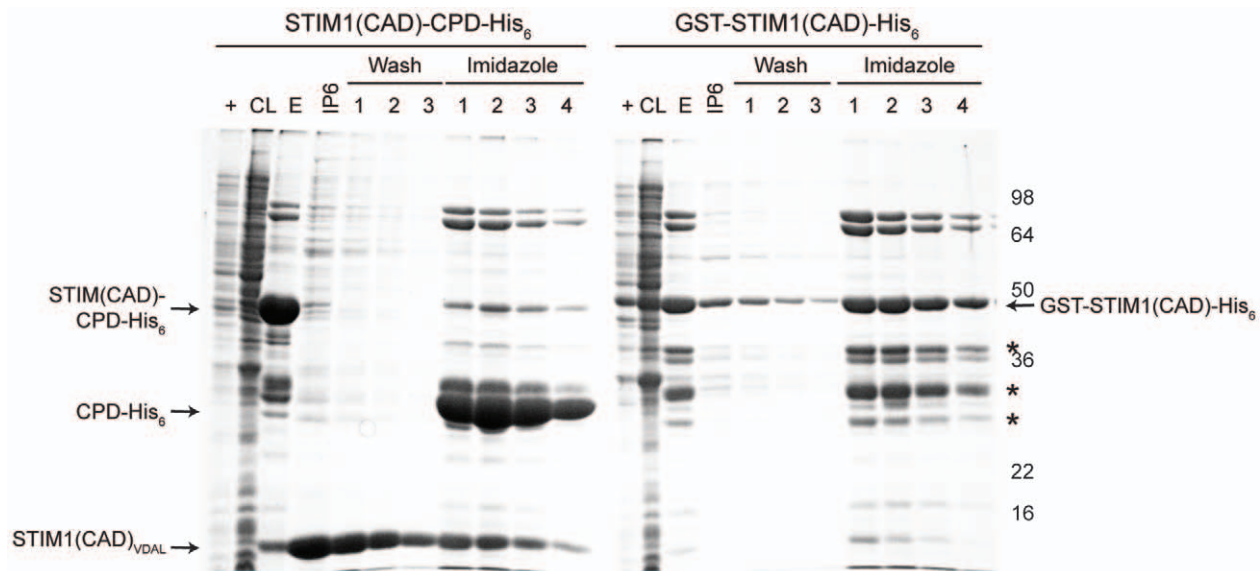


**Figure 6. Comparison of His<sub>6</sub>-tag removal from *Plasmodium falciparum* SENP1 using thrombin relative to the CPD.** (A) Coomassie stain of SDS-PAGE analysis of *P. falciparum* SENP1 (PfSEN1, 25 kDa) purified using either the CPD-His<sub>6</sub> or His<sub>6</sub>-affinity tags. PfSEN1-CPD-His<sub>6</sub> or His<sub>6</sub>-PfSEN1 bound to the Ni<sup>2+</sup>-NTA resin was incubated with 100 μM InsP<sub>6</sub> for 2 hr at room temperature; the resin was washed three times, and wash fractions were collected. Ni<sup>2+</sup>-bound proteins were eluted by adding 200 mM imidazole. +, IPTG induced culture, CL, cleared lysate, E, imidazole elution prior to InsP<sub>6</sub> addition, IP<sub>6</sub>, elution following InsP<sub>6</sub> incubation. (B) UV trace PfSEN1 further purified by gel filtration chromatography following His<sub>6</sub>-tag removal. Inset, Coomassie stain of gel filtration fractions of PfSEN1 purifications. Thrombin refers to PfSEN1 purified following thrombin-mediated removal of the N-terminal His<sub>6</sub>-tag, while InsP<sub>6</sub> refers to InsP<sub>6</sub>-induced, CPD-mediated removal of the C-terminal CPD-His<sub>6</sub>-tag. The residues added to the resulting PfSEN1 protein are shown: GSHM is added to the N-terminus of PfSEN1 following thrombin cleavage, while VDAL is added to the C-terminus of PfSEN1 following InsP<sub>6</sub>-activated CPD cleavage. (C) Coomassie stain of SDS-PAGE analyses of fractions taken during His<sub>6</sub>-PfSEN1 purification prior to thrombin incubation (–), following 12 hr thrombin incubation (+), and following subtractive IMAC to remove uncleaved His<sub>6</sub>-PfSEN1 (Ni<sup>2+</sup>-NTA). The yield of PfSEN1 diminished with each experimental manipulation.

doi:10.1371/journal.pone.0008119.g006

cleavage is complete within 1–2 hrs at temperatures between 4°C and 37°C, requires only micromolar concentrations of the small molecule InsP<sub>6</sub> (an abundant and inexpensive reagent), and occurs efficiently both in the presence of standard protease inhibitor

cocktails and in the absence of salt. This latter property carries the additional advantage of allowing the user to determine the buffer system in which to elute the target protein, eliminating the need for desalting or buffer exchange steps that can reduce protein



**Figure 7. The CPD can improve the stability of target proteins.** The Crac activation domain (CAD128) of STIM1 (amino acids 342–469, 14 kDa) was expressed in *E. coli* fused to either CPD-His<sub>6</sub> or GST-His<sub>6</sub>. Asterisks indicate GST-STIM1(CAD)-His<sub>6</sub> derived degradation products. His<sub>6</sub>-tagged proteins bound to the Ni<sup>2+</sup>-NTA resin were incubated with 50  $\mu$ M InsP<sub>6</sub> for 1 hr at room temperature, and the resin was washed three times, followed by elution of Ni<sup>2+</sup>-bound proteins by 200 mM imidazole. CL, cleared lysate, +, IPTG induced culture, FT, flowthrough, IP6, elution from InsP<sub>6</sub> incubation.

doi:10.1371/journal.pone.0008119.g007

yields. In addition, we have created a number of vector backbones that can be used to vary the residues that are appended to the target protein following CPD-mediated cleavage, which can range from a single amino acid residue to an HA epitope tag (Figure 2). Thus, the CPD system allows for considerable flexibility in optimizing purification procedures, as is often necessary for uncharacterized target proteins.

This versatility, combined with our observation that it can improve the solubility and integrity of difficult-to-express proteins (Figures 5 to 8), suggests that it will have widespread utility in biological research. The simplicity of this system will also make it amenable for large-scale proteomic, structural genomic, and commercial applications by eliminating the cost and complexity associated with exogenous site-specific proteases, potentially permitting its use in robotic systems for constructing protein arrays for screening purposes.

## Materials and Methods

### Bacterial Growth Condition

Overnight bacterial strains were grown at 37°C in Luria-Bertani (LB) broth. Antibiotics were used at 100  $\mu$ g/mL carbenicillin for pET22b vectors expressed in *E. coli*.

### Strain Construction

Primers used are listed in Table S1; strains constructed are listed in Table S2 in the Supporting Information. For construction of pET-CPD<sub>Sall</sub> vectors, DNA encoding *Vibrio cholerae* MARTX toxin amino acids 3440–3650 from *Vibrio cholerae* N16961 was PCR amplified from genomic DNA using primers #1 and #2. The resulting PCR fragment was cloned into the Sall and XhoI sites of the pET22b and pET28a expression vectors, respectively (Novagen). For construction of the pET-CPD<sub>SacI</sub> vector, DNA encoding *Vibrio cholerae* MARTX toxin amino acids 3442–3650 from *Vibrio cholerae* N16961 was PCR amplified from genomic DNA using primers #3 and #2, and the resulting PCR fragment was cloned

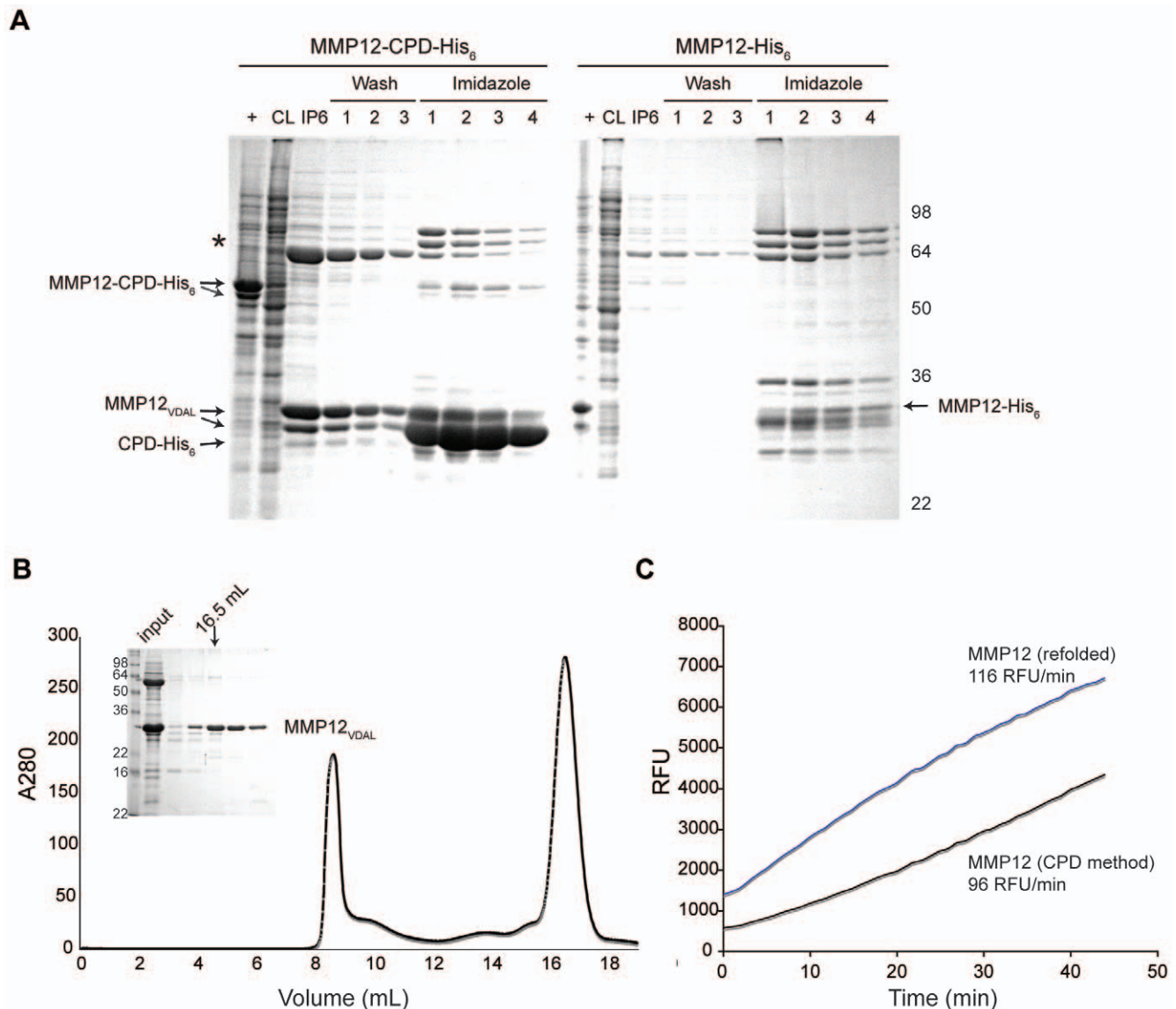
into the SacI and XhoI sites of pET22b. To construct the pET-HA-CPD<sub>Sall</sub> vectors, DNA encoding the HA epitope tag was added to the 5' end of primer #4, and PCR amplification using primers #4 and #2 was used to fuse the HA tag directly to amino acid 3440 of *V. cholerae* MARTX CPD. The resulting PCR fragment was cloned into the Sall and XhoI sites of the pET22b and pET28a expression vectors, respectively. For construction of the pET-CPD<sub>BamHI-Leu</sub> vector, DNA encoding *Vibrio cholerae* MARTX toxin amino acids 3444–3650 from *Vibrio cholerae* N16961 was PCR amplified from genomic DNA using primers #5 and #2, and the resulting PCR fragment was cloned into the BamHI and XhoI sites of pET22b. For construction of the pET-CPD<sub>BamHI</sub> vector, DNA encoding *Vibrio cholerae* MARTX toxin amino acids 3440–3650 from *Vibrio cholerae* N16961 was PCR amplified from genomic DNA using primers #6 and #2, and the resulting PCR fragment was cloned into the BamHI and XhoI sites of pET22b.

The pET22b-GFP-CPD construct was cloned by PCR amplifying GFP from pEGFPN3 (Clontech) using primers #7 and #8. To construct the pET22b-gp130(ICD)-CPD vector, amino acids 642–918 of gp130 corresponding to the intracellular domain were PCR amplified using primers #9 and #10 and pET21a-gp130(ICD) as a template. The pET22b-BirA-CPD vector was constructed by PCR amplifying the *birA* gene from a pGEX4T1-BirA template using primers #9 and #10. The pET22b-STIM1(CAD)-CPD plasmid was constructed by PCR amplifying DNA encoding amino acids 342–469 of STIM1 using pGEX6-CAD128 as a template and primers #13 and #14. The pET22b-mMMP12-CPD construct was constructed by PCR amplifying the catalytic domain of mouse MMP12 (amino acids 29–267) using pET41a-mMMP12 as a template using primers #15 and #16. In all cases, the resulting PCR products were cloned into the NdeI and Sall sites of pET22b-CPD<sub>Sall</sub>.

### Protein Expression and Purification

For purification of His<sub>6</sub>-tagged CPD fusion proteins, overnight cultures of the appropriate strain were diluted 1:500 into 1 L 2YT





**Figure 8. The CPD improves the solubility of mouse macrophage metalloelastin (MMP12) relative to a His<sub>6</sub>-affinity tag.** (A) SDS-PAGE analysis of His<sub>6</sub>-tagged MMP12 using Coomassie stain. MMP12 was fused to His<sub>6</sub>-tagged CPD, and the expression of the fusion protein relative to His<sub>6</sub>-tagged MMP12 was compared. Asterisk indicates a putative chaperone protein that co-purifies with MMP12<sub>VDAL</sub>. The diagonal arrows indicate a His<sub>6</sub>-tagged truncated MMP12 product that is also observed during MMP12 purification from inclusion bodies [17]. His<sub>6</sub>-tagged proteins bound to the Ni<sup>2+</sup>-NTA resin were incubated with 50  $\mu$ M InsP<sub>6</sub> for 1 hr at room temperature, and the resin was washed three times, followed by elution of Ni<sup>2+</sup>-bound proteins by 200 mM imidazole. CL, cleared lysate, +, IPTG induced culture, FT, flowthrough, IP6, elution from InsP<sub>6</sub> incubation, E, imidazole elution prior to InsP<sub>6</sub> addition. (B) Additional purification of MMP12<sub>VDAL</sub> by gel filtration chromatography. Inset, Coomassie stain of SDS-PAGE analysis of gel filtration fractions of MMP12<sub>VDAL</sub>. (C) MMP12 activity assay. The activity of MMP12 purified under denaturing conditions and refolded (MMP12 (Refolded)) and MMP12<sub>VDAL</sub> purified using the CPD system (CPD method) against a standard fluorogenic substrate were compared. Comparable rates of fluorogenic substrate cleavage are observed for MMP12 purified by the CPD method relative to the refolding method.  
doi:10.1371/journal.pone.0008119.g008

media and grown shaking at 37°C. When an OD<sub>600</sub> of 0.6 was reached, IPTG was added to 250  $\mu$ M, and cultures were grown for 3–4 hrs at 30°C. Cultures were pelleted, resuspended in 25 mL lysis buffer [500 mM NaCl, 50 mM Tris-HCl, pH 7.5, 15 mM imidazole, 10% glycerol] and flash frozen in liquid nitrogen. Lysates were thawed, then lysed by sonication and cleared by centrifugation at 15,000 $\times$ g for 30 minutes. His<sub>6</sub>-tagged CPD fusion proteins were affinity purified by incubating the lysates in batch with 0.5–1.0 mL Ni-NTA Agarose beads (Qiagen) with shaking for 2–4 hrs at 4°C. The binding reaction was pelleted at 1,500 $\times$ g, the supernatant was set aside, and the pelleted Ni<sup>2+</sup>-

NTA agarose beads were washed three times with lysis buffer. In some cases, 10% of the Ni<sup>2+</sup>-NTA beads containing immobilized CPD-His<sub>6</sub> fusion proteins were removed, pelleted and then His<sub>6</sub>-tagged fusion protein eluted using high imidazole buffer [500 mM NaCl, 50 mM Tris-HCl, pH 7.5, 175 mM imidazole, 10% glycerol].

To liberate untagged target proteins into the supernatant fractions, 300–500  $\mu$ L lysis buffer was added to the Ni<sup>2+</sup>-NTA beads containing CPD-His<sub>6</sub> fusion proteins and the indicated amount of inositol hexakisphosphate (InsP<sub>6</sub>, Calbiochem) was added. In general, on-bead cleavage was allowed to proceed by



**Table 1.** Target proteins expressed and purified by CPD purification method.

Target protein	Yield <sup>a</sup> (mg/L culture)	Yield (nmol/L culture)	Activity
GFP <sub>VDAL</sub> (CPD method)	3.3	105	Fluorescence at 511 nm
gp130(ICD) <sub>VDAL</sub> (CPD method)	5.9	188	n/a
gp130(ICD)-His <sub>6</sub>	3.7	115	n/a
BirA <sub>VDAL</sub> (CPD method)	10.9	202	Biotinylates LHHILDAQKMVWNHR BirA biotinylation site
GST-BirA-His <sub>6</sub>	12.0	90	Biotinylates LHHILDAQKMVWNHR BirA biotinylation site
PfSEN1 <sub>VDAL</sub> (CPD method)	2.0	67	Cleaves PfSUMO
GSHM-PfSEN1	1.4	46	Cleaves PfSUMO
STIM1(CAD) <sub>VDAL</sub>	2.1	148	Binds Orai1
GST-STIM1(CAD)-His <sub>6</sub>	2.5 <sup>b</sup>	62	n/a
MMP12 <sub>VDAL</sub> (CPD method)	1.4	47	Cleaves fluorogenic peptide substrate Mca-PLGLDL(Dpa)AR
MMP12 (refolded)	23	767	Cleaves fluorogenic peptide substrate Mca-PLGLDL(Dpa)AR

<sup>a</sup>Protein yield per litre of culture.

<sup>b</sup>Yield difficult to assess since GST-fusion protein degrades and falls out of solution over time.

doi:10.1371/journal.pone.0008119.t001

nutating the beads in the presence of 50–100  $\mu$ M InsP<sub>6</sub> for 1–2 hr at either room temperature or 4°C. The beads were pelleted at 1,500 $\times$ g, and the supernatant fraction was removed. The beads were then washed 3–4 times with 300–500  $\mu$ L lysis buffer, and supernatant fractions retained. His<sub>6</sub>-tagged proteins remaining on the beads (i.e. cleaved CPD-His<sub>6</sub>) were eluted using high imidazole buffer [500 mM NaCl, 50 mM Tris-HCl, pH 7.5, 175 mM imidazole, 10% glycerol] in 300–500  $\mu$ L volumes. The elution was repeated 3–4 times, and eluate fractions were collected. Purification of His<sub>6</sub>-tagged proteins lacking the CPD was performed in parallel.

This general procedure was followed with the following exceptions: for purification of MMP12 constructs, the cultures were grown at 16°C for 8 hr after IPTG induction, and 1 mM tris(2-carboxyethyl)phosphine (TCEP) was added to the lysis buffer to prevent misfolding of the protein. PfSEN1 and BirA protein purifications were performed exclusively at room temperature, since at 4°C, protein aggregation was observed. For removal of the His<sub>6</sub>-tag from His<sub>6</sub>-PfSEN1, thrombin beads (Calbiochem) that had been washed in PBS were added to the eluted His<sub>6</sub>-PfSEN1, which had been buffer exchanged into PBS according to the manufacturer's instructions. Thrombin cleavage was allowed to proceed with shaking overnight for 12 hr at room temperature.

Aliquots were taken before and after thrombin addition to monitor cleavage efficiency. Thrombin cleaved, untagged PfSEN1 was enriched by performing a subtractive Ni<sup>2+</sup>-NTA pull-down. Untagged PfSEN1 from both methods was then buffer-exchanged into gel filtration buffer (50 mM NaCl, 20 mM Tris pH 8.0). Protein purifications were analyzed by SDS-PAGE and Coomassie staining using GelCode Blue (Pierce). Purified protein concentrations of purified were determined by Bradford assay (Pierce).

#### Purification of MMP12-His<sub>6</sub>

MMP12-His<sub>6</sub> was purified as previously described [17] with the following modifications. The cell pellet was resuspended in 100 mM NaCl, 100 mM Tris pH 8.0, 5.0 mM EDTA, 0.5 mM DTT, 100  $\mu$ g/mL lysozyme and stirred for 2 hr. The cells were sonicated then centrifuged at 10,000 rpm for 10 min. The resulting inclusion bodies were washed two times and then resuspended in 50 mL 6M guanidine hydrochloride, 10 mM Tris pH 8.0 by stirring at 4°C overnight. The mixture was centrifuged at 15,000 rpm for 30 min, and 2 mL aliquots of supernatant were prepared. The supernatant was diluted 1:100 into denaturing buffer [6M Urea, 50 mM Tris pH 8.0, 10 mM CaCl<sub>2</sub>, 30 mM NaCl, 5 mM DTT] to a final concentration of 0.1–0.2 mg/mL.

**Table 2.** Comparison of CPD-mediated and thrombin-mediated purification of PfSEN1.

	PfSEN1-CPD-His <sub>6</sub>	His <sub>6</sub> -PfSEN1
Step 1	Prepare soluble lysate (1 hr)	Prepare soluble lysate (1 hr)
Step 2	IMAC purification (2 hr)	IMAC purification (1 hr)
Step 3	On-bead cleavage; collect supernatant (2 hr)	Imidazole elution
Step 4	Concentrate protein (0.5 hr)	Buffer exchange and concentrate protein (0.5 hr)
Step 5	Gel filtration chromatography (1 hr)	Thrombin cleavage overnight (>12 hr)
Step 6	Concentrate protein (0.5 hr)	Remove His <sub>6</sub> -tag and uncleaved fusion with IMAC (1 hr)
Step 7		Concentrate protein and buffer exchange (0.5 hr)
Step 8		Gel filtration chromatography (1hr)
		Concentrate protein (0.5 hr)
<b>Total time</b>	5 hr	>17.5 hr

doi:10.1371/journal.pone.0008119.t002

**Table 3.** Comparison of CPD method to published method for purifying matrix metalloelastin.

	MMP12-CPD-His <sub>6</sub>	MMP12-His <sub>6</sub> [17]
Step 1	Prepare soluble lysate (1 hr)	Prepare and dissolve inclusion bodies (16 hr)
Step 2	IMAC purification (2 hr)	Protein refolding by dialysis in 1/20 volume (48 hr)
		a. 24 hr–6 M Urea, 4 L
		b. 12 hr–3 M Urea, 4 L
		c. 12 hr–1 M Urea, 4 L
Step 3	On-bead cleavage; collect supernatant (2 hr)	Load partially refolded protein on anion and cation exchange tandem columns (2.5 hr)
Step 4	Concentrate protein (0.5 hr)	Wash columns with three buffers to fully refold protein on column (2 hr)
Step 5	Gel filtration chromatography (1 hr)	Elute protein from cation exchange column (0.5 hr)
Step 6	Concentrate protein (0.5 hr)	Concentrate protein (0.5 hr)
<b>Total time</b>	<b>7 hr</b>	<b>3–4 days</b>

doi:10.1371/journal.pone.0008119.t003

The protein was then dialyzed for 24 hr in 2 L refolding buffer 1 [3 M Urea, 50 mM Tris pH 8.0, 10 mM CaCl<sub>2</sub>, 30 mM NaCl, 5 mM DTT]. The partially refolded protein was then dialyzed in 4 L of refolding buffer 2 [1 M Urea, 50 mM HEPES pH 7.4, 10 mM CaCl<sub>2</sub>, 5 mM DTT]. The buffer exchanged protein was then purified using tandem 5 mL MonoQ and SP Sepharose (GE Healthcare) at 4°C. After loading the protein on the column, the column was washed with 50 mL of refolding buffer 2 without DTT at 1 M, 0.5, and 0 M urea, respectively. The protein was eluted from the SP column in 500 mM NaCl, 50 mM HEPES pH 7.4, 10 mM CaCl<sub>2</sub>.

### Gel Filtration Chromatography

Untagged PfSEN1 obtained from either thrombin or InsP<sub>6</sub>-mediated cleavage was concentrated using a 10 kDa Centricon concentrator (Millipore) and buffer exchanged into 50 mM NaCl, 20 mM Tris pH 8.0 and purified on a Superdex 200 10/30 column (GE Healthcare) equilibrated in the same buffer. For MMP12, the gel filtration buffer contained 150 mM NaCl, 50 mM Tris pH 7.4, 10 mM TCEP. Gel filtrations were performed at 4°C.

### Activity Assays

Fluorescence of purified GFP at 511 nm was verified using a Molecular Devices fmax plate reader in black 96-well plates and 488 nm excitation. The activity of MMP12 was determined using the fluorogenic substrate Mca-PLGLDL(Dpa)AR (Mca, (7-methoxycoumarin-4-yl)acetyl, Dpa, *N*-3-(2,4-dinitrophenyl)-L-2,3-diaminopropionyl, Anaspec). Reactions were performed in the assay buffer (50 mM Tris pH 7.4, 150 mM NaCl, 10 mM CaCl<sub>2</sub>, 0.02% NaN<sub>3</sub>, 5 mM TCEP) at 37°C. The substrate was used at

10 μM and the protein at 0.2 μM. The substrate hydrolysis was monitored continuously in a fluorescent plate reader (Molecular Devices) using an excitation wavelength of 325 nm and an emission wavelength of 395 nm.

### Supporting Information

**Table S1** Primers used in Study. <sup>a</sup> Restriction enzyme sequences are underlined, and the HA tag is shown in italics. <sup>b</sup> RE - Restriction site

Found at: doi:10.1371/journal.pone.0008119.s001 (0.06 MB DOC)

**Table S2** Strains used in study. 1. Skinotis G and Lupardus, PJ, *et al.* (2008) Mol Cell 31: 737–748. 2. Ponder EL, *et al.* (2009) under review Nat Chem Biol. 3. Park CY, *et al.* (2009) Cell 136: 876–890.

Found at: doi:10.1371/journal.pone.0008119.s002 (0.07 MB DOC)

### Acknowledgments

We thank Chris Overall for kindly providing the pET41a-mMMP12 expression construct, and Chan Young Park for providing the GST-CAD-His<sub>6</sub> construct used to clone the CAD domain of Stim1.

### Author Contributions

Conceived and designed the experiments: AS PJL MM ELP AMS KCG MB. Performed the experiments: AS PJL MM ELP AMS. Analyzed the data: AS PJL MM ELP AMS. Contributed reagents/materials/analysis tools: AS PJL MM ELP AMS KCG. Wrote the paper: AS MB.

### References

- Graslund S, Nordlund P, Weigelt J, Hallberg BM, Bray, et al. (2008) Protein production and purification. Nat Methods 5: 135–146.
- Arnau J, Lauritzen C, Petersen GE, Pedersen J (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expr Purif 48: 1–13.
- Sheahan KL, Cordero CL, Satchell KJ (2007) Autoprocessing of the *Vibrio cholerae* RTX toxin by the cysteine protease domain. EMBO J 26: 2552–2561.
- Shen A, Lupardus PJ, Albrow VE, Guzzetta A, Powers JC, et al. (2009) Mechanistic and structural insights into the proteolytic activation of *Vibrio cholerae* MARTX toxin. Nat Chem Biol 5: 469–478.
- Lupardus PJ, Shen A, Bogoy M, Garcia KC (2008) Small Molecule-Induced Allosteric Activation of the *Vibrio cholerae* RTX Cysteine Protease Domain. Science 322: 265–268.
- Prochazkova K, Satchell KJ (2008) Structure-function analysis of inositol hexakisphosphate-induced autoprocessing of the *Vibrio cholerae* multifunctional autoprocessing RTX toxin. J Biol Chem 283: 23656–23664.
- Michell RH (2008) Inositol derivatives: evolution and functions. Nat Rev Mol Cell Biol 9: 151–161.
- Raboy V (2003) myo-Inositol-1,2,3,4,5,6-hexakisphosphate. Phytochemistry 64: 1033–1043.
- Mao H (2004) A self-cleavable sortase fusion for one-step purification of free recombinant proteins. Protein Expr Purif 37: 253–63.
- Zhang A, Gonzalez SM, Cantor EJ, Chong S (2001) Construction of a mini-intein fusion system to allow both direct monitoring of soluble protein expression and rapid purification of target proteins. Gene 275: 241–52.

11. Skiniotis G, Lupardus PJ, Martick M, Walz T, Garcia KC (2008) Structural organization of a full-length gp130/LIF-R cytokine receptor transmembrane complex. *Mol Cell* 31: 737–748.
12. Hibi M, Murakami M, Saito M, Hirano T, Taga T, et al. (1990) Molecular cloning and expression of an IL-6 signal transducer, gp130. *Cell* 63: 1149–1157.
13. O'Callaghan CA, Byford MF, Wyer JR, Willcox BE, Jakobsen, et al. (1999) BirA enzyme: production and application in the study of membrane receptor-ligand interactions by site-specific biotinylation. *Anal Biochem* 266: 9–15.
14. Ponder EL, Leader BA, Albrow VE, Bekes M, Mikolajczyk J, et al. (2009) Functional characterization of PiSENPI, a SUMO deconjugating protease of *Plasmodium falciparum* using small molecule inhibitors. under review *Nat Chem Biol*.
15. Mehlin C, Boni E, Buckner FS, Engel L, Feist T, et al. (2006) Heterologous expression of proteins from *Plasmodium falciparum*: results from 1000 genes. *Mol Biochem Parasitol* 148: 144–160.
16. Park CY, Hoover PJ, Mullins FM, Bachhawat P, Covington ED, et al. (2009) STIM1 clusters and activates CRAC channels via direct binding of a cytosolic domain to Orai1. *Cell* 136: 876–890.
17. Parker AA, Stow MD, Smith K, Panicker AK, Guilloteau JP, et al. (2000) Large-scale expression, refolding, and purification of the catalytic domain of human macrophage metalloelastase (MMP-12) in *Escherichia coli*. *Protein Expr Purif* 20: 152–161.
18. Butt TR, Edavettal SC, Hal JP, Mattern MR (2005) SUMO fusion technology for difficult-to-express proteins. *Protein Expr Purif* 43: 1–9.
19. Banki MR, Feng L, Wood DW (2005) Simple bioseparations using self-cleaving elastin-like polypeptide tags. *Nat Methods* 2: 659–661.