




OPEN

Molecular and structural insights into an asymmetric proteolytic complex (ClpP1P2) from *Mycobacterium smegmatis*

Jyotsna Nagpal, Jason J. Paxman, Jessica E. Zammit, Adam A. Thomas, Kaye N. Truscott , Begoña Heras * & David A. Dougan *

The ClpP protease is found in all kingdoms of life, from bacteria to humans. In general, this protease forms a homo-oligomeric complex composed of 14 identical subunits, which associates with its cognate ATPase in a symmetrical manner. Here we show that, in contrast to this general architecture, the Clp protease from *Mycobacterium smegmatis* (*Msm*) forms an asymmetric hetero-oligomeric complex ClpP1P2, which only associates with its cognate ATPase through the ClpP2 ring. Our structural and functional characterisation of this complex demonstrates that asymmetric docking of the ATPase component is controlled by both the composition of the ClpP1 hydrophobic pocket (Hp) and the presence of a unique C-terminal extension in ClpP1 that guards this Hp. Our structural analysis of *Msm*ClpP1 also revealed openings in the side-walls of the inactive tetradecamer, which may represent sites for product egress.

Bacterial Clp proteases are generally formed by two components, a single peptidase component (ClpP) which associates with one or more members of the AAA+ (ATPases associated with a variety of cellular activities) superfamily (e.g. ClpA, ClpX or ClpC)^{1–6}. In *Escherichia coli*, ClpP is expressed as a proenzyme and the N-terminal propeptide is autocatalytically removed⁷. The active complex is a barrel-shaped oligomer composed of two heptameric rings stacked back-to-back⁸. The catalytic residues (Ser-His-Asp), of this complex, are encapsulated within the barrel-shaped proteolytic chamber and access to the chamber is restricted to a narrow entry portal at either end of the complex. This design hinders entry of correctly folded proteins into the catalytic chamber, and as such prevents the indiscriminate turnover of cytosolic proteins. Substrate recognition and unfolding is mediated by the ATPase component of these machines, and ATPase docking to ClpP also couples substrate delivery to peptidase activation by triggering “gate-opening” of the peptidase entry portal⁹. ATPase docking to ClpP is mediated by two types of contacts – *static* and *dynamic*^{10,11}. The primary contact is mediated by a *static* interaction between a highly conserved loop (commonly referred to as the IGF/L loop) located on the proximal face of the ATPase component, which docks into a hydrophobic pocket (Hp) on ClpP. The Hp is located at the periphery of the interface and is composed of several highly conserved aromatic and hydrophobic residues which are critical for interaction with the ATPase component¹². The second *dynamic* contact, which involves the N-terminal loops of ClpP and the pore-2 loops of the ATPase, is axial in nature and regulated by the nucleotide state of the ATPase component¹³. The Hp is also the site of binding of a novel class of antibiotic that dysregulates ClpP function. These compounds (e.g. acyldepsipeptides (ADEPs)) not only inhibit regulated protein turnover by ClpP by blocking ATPase docking to ClpP, but they also facilitate unregulated access of cytosolic proteins into the catalytic chamber of ClpP thereby triggering their uncontrolled turnover^{14–17}.

In contrast to *E. coli* several bacterial species contain two or more ClpP homologs, which form a diverse array of oligomeric complexes^{18–24}. Interestingly, in *Mycobacterium tuberculosis* (*Mtb*), despite both proteins (i.e. ClpP1 and ClpP2), containing catalytic residues and a propeptide, neither protein alone is processed nor are they proteolytically active²¹. In a landmark study, Goldberg and colleagues identified Benzyloxycarbonyl-Leu-Leucinal (here termed z-LL) as a potent activator of the *Mtb* hetero-oligomeric ClpP1P2 complex, which led the group to propose that z-LL facilitated assembly of an “active” hetero-oligomer²¹. Structural analysis later confirmed that the

Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, 3086, Australia. *email: b.heras@latrobe.edu.au; d.dougan@latrobe.edu.au

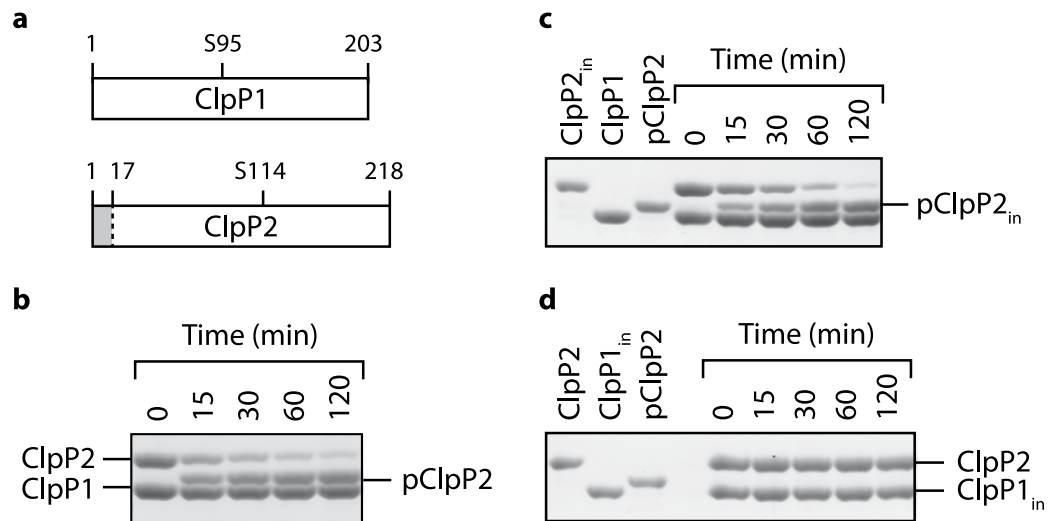


Figure 1. Propeptide processing of *Msm*ClpP2 occurs via the active site residues of *Msm*ClpP1 in the absence of the activator (z-LL). (a) Cartoon representation of *Msm*ClpP1 and *Msm*ClpP2 indicating the position of the propeptide (grey) and the catalytic Ser residues (residue 95 for *Msm*ClpP1 and 114 for *Msm*ClpP2). The *in vitro* processing of (b) wild type *Msm*ClpP2 into processed (p) *Msm*ClpP2 (pClpP2) or (c) inactive *Msm*ClpP2 (ClpP2_{in}) into processed ClpP2_{in} (pClpP2_{in}), was monitored in the absence of activator (z-LL). (d) Processing of *Msm*ClpP2 (in the absence of z-LL), failed to occur in the presence of the proteolytically inactive ClpP1 mutant (ClpP1_{in}).

active *Mtb* Clp protease complex was asymmetric in nature, composed of a single heptameric ring of each protein and that the activator docked into the substrate binding pocket as a substrate agonist²⁴. Finally, the asymmetric nature of the ClpP1P2 complex from *Mtb* was also shown to extend to its peptidase specificity and its interaction with its cognate ATPase components^{25,26}.

In this study we have examined the activation and assembly of the ClpP1P2 complex from *Mycobacterium smegmatis* (*Msm*). Similar to the ClpP1P2 complex from *Mtb*, we show that *Msm*ClpP1P2 forms an asymmetric complex. The asymmetric nature of this complex is not limited to the composition of the tetradecamer but also extends to the peptidase activity of the complex (including propeptide processing), and interaction with its cognate ATPase components. In this case, our biochemical and structural analysis revealed that asymmetric docking of the ATPase components (to *Msm*ClpP1P2) is controlled by two elements within *Msm*ClpP1. Firstly, by a C-terminal extension in *Msm*ClpP1 that obstructs access of the ATPase component to the Hp and secondly by residues that line the Hp. Together with our analysis of the *Ec*ClpP Hp, we have identified Y88 (I104 in *Ec*ClpP) as a key feature of ATPase docking specificity. In addition, our structure of *Msm*ClpP1 reveals small openings in the side wall of the ClpP1 tetradecamer, which may represent exit portals for the egress of cleaved polypeptides.

Results and Discussion

Processing of the *Msm*ClpP2 propeptide by the active site residues of *Msm*ClpP1 does not require an activator. Given that correct processing of ClpP propeptides is crucial for ClpP peptidase activity²⁷, we examined the processing of both *Msm*ClpP components. Initially, in order to determine if both *Msm*ClpP subunits were processed, we co-expressed untagged ClpP1 together with His₁₀ tagged ClpP2 (ClpP2_{H10}) in *E. coli* and co-purified the active *Msm*ClpP1P2 complex. Although recovery of the active hetero-oligomeric complex was poor, only one component (*Msm*ClpP2) was processed, and this processing only occurred when the other component (ClpP1) was present (Fig. S1). To determine the location of the processing site we cloned *Msm* *clpP1* and *clpP2*, expressed them in *E. coli* and purified the individual components to analyse propeptide cleavage *in vitro*. Consistent with the co-expression experiment, the processing of *Msm*ClpP2, required *Msm*ClpP1 (Fig. 1). Interestingly, although these data are similar to the processing of the *Mtb*ClpP1P2 complex, in which both ClpP subunits are processed by a hetero-oligomeric complex^{21,28} processing of *Msm*ClpP2 occurred in the absence of any additional components (Fig. 1b). In contrast to the processing of *Msm*ClpP2, processing of both *Mtb*ClpP subunits required either the artificial activator z-LL²¹ (also see Fig. S2) or a cognate ATPase component²⁵. To determine the site of processing, we performed Edman degradation of ClpP1 and processed ClpP2 (pClpP2) from the *in vitro* processing assay. Similar to *Ec*ClpP and *Mtb*ClpP2 (processed between Ala₁₂ and Arg₁₃), *Msm*ClpP2 was processed between residues Ala₁₆ and Arg₁₇. However, in contrast to *Mtb*ClpP1, which was processed between Met₇ and Arg₈^{21,28}, no processing of *Msm*ClpP1 was observed.

Next we asked, how does processing of *Msm*ClpP2 occur? Initially, to address this question, we generated active site mutants of *Msm*ClpP1 (ClpP1_{in}) and *Msm*ClpP2 (ClpP2_{in}), in which the active site Ser (Ser95 and Ser114, respectively) were replaced with Ala. Similar to the processing of *Mtb*ClpP2²⁸, mutation of the active site Ser in *Msm*ClpP2 (*Msm*ClpP2_{in}) did not affect processing (Fig. 1c), while in contrast processing was completely abolished by mutation of the active site Ser in *Msm*ClpP1 (Fig. 1d). Collectively, these data demonstrate that processing of *Msm*ClpP2 is not autocatalytic, rather it appears to occur *in trans* via the catalytic triad of *Msm*ClpP1. An alternative interpretation of these data is that the catalytic triad of *Msm*ClpP2 was not active. Therefore, to ensure that the

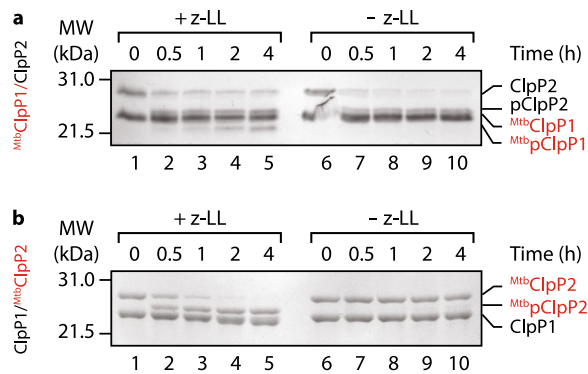


Figure 2. Propeptide processing of mixed *Mtb/Msm*ClpP1P2 complexes in the presence or absence of z-LL. **(a)** *In vitro* processing of *Mtb*ClpP1^{*Msm*}P2 in the presence (lanes 1–5) or absence of z-LL (lane 6–10). **(b)** *In vitro* processing of *Msm*ClpP1^{*Mtb*}P2 in the presence (lanes 1–5) or absence of z-LL (lane 6–10). Following processing proteins were separated by SDS-PAGE and visualised by staining with CBB.

catalytic triad of wild type *Msm*ClpP2 was active, we examined the turnover of different model substrates (from short peptides to a folded protein), by various mixed (wild type and mutant) ClpP1P2 complexes (Fig. S3). These data demonstrated that although the catalytic triad of ClpP2 was active it was not essential for the turnover of all substrates. For example, although the active site of ClpP2 was dispensable for the turnover of short peptide substrates (Fig. S3a,b, compare lanes 2 and 6) it was necessary for efficient *Ec*ClpX-mediated turnover of the model protein substrate, GFP-*Ec*SsrA (Fig. S3c, blue triangles). Interestingly, although mutation of ClpP1 completely inhibited the turnover of peptide substrates it had little effect on the rate of native protein turnover (Fig. S3c, red squares). In contrast mutation of ClpP2 slowed the turnover of GFP-*Ec*SsrA (Fig. S3c, blue triangles). Collectively, these data suggest that both ClpP1 and ClpP2 exhibit unique substrate specificities and that entry of a protein substrate (via *Ec*ClpX) into the proteolytic chamber may be unidirectional.

Next, in order to gain a better understanding of how *in trans* processing might occur we examined propeptide processing of various heterologous component combinations (Fig. 2). Initially we monitored the ability of *Mtb*ClpP1 to facilitate the processing of *Msm*ClpP2 in the absence or presence of z-LL (Fig. 2a). Interestingly, consistent with the processing of *Msm*ClpP2 by *Msm*ClpP1 (Fig. 1b), z-LL was not required for the *Mtb*ClpP1-mediated processing of *Msm*ClpP2, despite an absolute requirement of z-LL for all *Mtb*ClpP1 activities (Fig. S2). Furthermore, the processing of *Mtb*ClpP1 was not required for processing of *Msm*ClpP2 to occur. Hence these data suggest that the initial processing event (of ClpP2) may occur by a mechanism that is distinct from the downstream processing of ClpP1. We then monitored the ability of *Msm*ClpP1 to facilitate the processing of *Mtb*ClpP2 (Fig. 2b). Surprisingly, despite the fact that z-LL was not required for the *Msm*ClpP1-mediated processing of *Msm*ClpP2, it (z-LL) was essential for the processing of *Mtb*ClpP2. This suggests that z-LL is required to trigger a conformational change in *Mtb*ClpP2 that mediates its processing by *Msm*ClpP1.

***Msm*ClpP2 is processed via a transient complex.** Although propeptide processing of homo-oligomeric ClpP complexes such as *Ec*ClpP and *Homo sapiens* ClpP (*Hs*ClpP) has long been described as autocatalytic⁷, the mechanism by which this step occurs is poorly defined and currently it remains unclear if processing occurs before (or as a result of) assembly of the tetradecamer, or indeed if processing occurs via the *cis* or *trans* ring^{7,27,29}. Therefore, we examined the composition of the active “processing” complex. For homo-oligomeric ClpP complexes, it is plausible that processing is mediated by the *cis* ring and the processing site is determined via a molecular “ruler” mechanism. For hetero-oligomeric complexes such as ClpP1P2²⁴, the distance between the processing site (within the propeptide) and the active site within the *trans* ring (~47 Å) is estimated to be much greater than the distance between the propeptide and the active site in the *cis* ring (~25 Å). Hence, to determine if processing of ClpP2 is mediated by formation of the ClpP1P2 complex, we generated single point mutations in both *Msm*ClpP proteins to disrupt the ring-ring interface. This mutation was based on a critical Arg residue that stabilises the ring-ring interface of *Staphylococcus aureus* ClpP (*Sa*ClpP)³⁰. Specifically, we replaced the Arg-finger residue in *Msm*ClpP1 and *Msm*ClpP2 (Arg168 and Arg189, respectively) with Ala. Initially, to ensure the overall structure of each mutant protein was not compromised, we compared the oligomeric state of each protein in the absence of z-LL using analytical ultracentrifugation (AUC). Importantly, both wild type and mutant *Msm*ClpP1 and *Msm*ClpP2 each formed heptamers (Fig. S4, red circles). Next, to monitor the effect of these mutations on *Msm*ClpP1P2 complex formation we performed a series of pull-down experiments, in which wild type or mutant *Msm*ClpP1_{H10} was immobilised to Ni-NTA-agarose beads. Importantly, mutation of a single component was sufficient to significantly reduce its interaction with the other component (Fig. 3), while mutation of both components almost completely abolished the interaction of the two components (Fig. 3b, lane 6). These data demonstrate that, similar to *Sa*ClpP, the Arg finger plays a crucial role in stabilising the *Msm*ClpP1P2 complex. Consistent with this loss of the ClpP1P2 tetradecamer, the turnover of a model peptide (Fig. 3c) or protein (Fig. 3d) substrate, by each of the different mutant protein complexes, was completely abolished. Surprisingly, and in contrast to the peptidase activity of each mutant protein complex, all three complexes retained the ability to process ClpP2, with only a modest change to the rate of processing (Fig. 3e). One explanation for these data is that the substrate (i.e. the propeptide) is, in this case, tethered to the peptidase and hence any residual interaction between the two rings

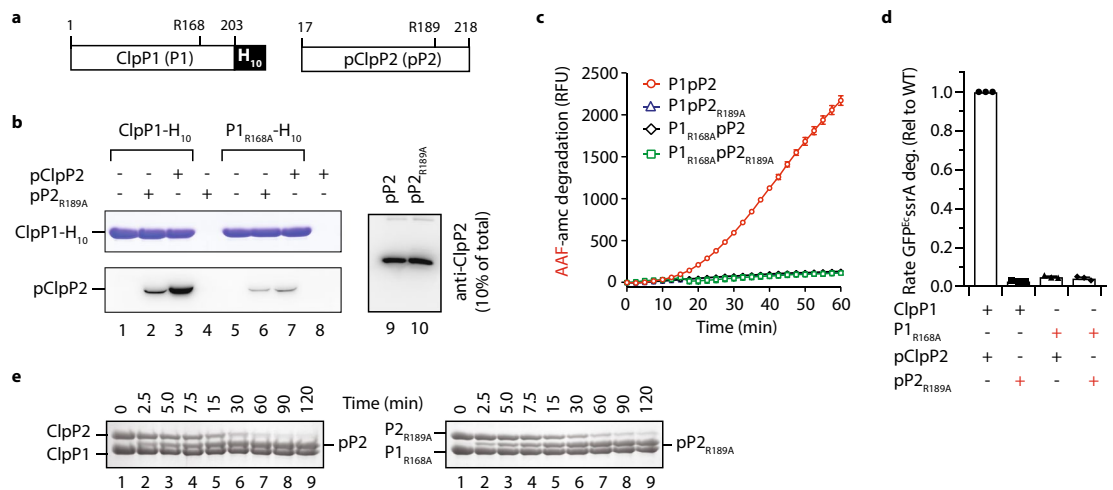


Figure 3. In contrast to peptide and protein degradation, propeptide processing of *Msm*ClpP2 does not require stable interaction between *Msm*ClpP1 and *Msm*ClpP2. (a) Cartoon representation of *Msm*ClpP1-H₁₀ and *Msm*ClpP2 indicating the position of the conserved Arg finger in ClpP1 (R168) and ClpP2 (R189). (b) Co-immunoprecipitation (co-IP) of wild type (lanes 1–3) or mutant ClpP1-H₁₀ (lanes 5–7) in the presence of either wild type (lanes 3 and 7) or mutant pClpP2 (lane 2 and 6). The total amount of wild type or mutant pClpP2 (lane 9 and 10, respectively) added to the co-IP is indicated. (c–d) The turnover of AAF-amc peptide (c) or rate of GFP^{Ec}ssrA degradation was monitored by fluorescence using wild type or mutant ClpP1P2 complexes (e) comparison of propeptide processing by the wild type ClpP1P2 complex (left panel) and the ClpP1_{R168A}P2_{R189A} complex (right panel).

may be sufficient to facilitate rapid cleavage of the propeptide. An alternative explanation for these data is that processing of ClpP2 is not mediated by the classic ClpP1P2 tetradecamer, but rather by an interaction that does not require inter-ring contacts mediated by the Arg fingers. Consistent with this idea, propeptide processing of *Msm*ClpP2 (by *Msm*ClpP1) is independent of z-LL activity (Fig. 1), while in contrast the activator is essential for all other proteolytic activities of the *Msm*ClpP1P2 complex.

Mutation of the hydrophobic pocket modulates the peptidase activity of the *Msm*ClpP1pP2 complexes. Next, in order to study the interaction of the *Msm*ClpP1P2 complex with its cognate ATPase components (ClpX and ClpC1), we generated specific point mutations within the hydrophobic pockets of *Msm*ClpP1 and *Msm*ClpP2. Initially, we targeted two residues in the Hp (the first was a highly conserved tyrosine residue found in all ClpP sequences, Y60 in ClpP1 and Y79 in ClpP2, the second hydrophobic residue was less conserved across ClpP sequences, Y110 in ClpP1 and L129 in ClpP2). Each of the hydrophobic residues (described above) was replaced with alanine, to generate a series of single point mutants (P1_{Y60A}, P1_{Y110A}, P2_{Y79A} or P2_{L129A}) and one double mutant of *Msm*ClpP2 in which both hydrophobic residues (Y79 and L129) were replaced with alanine, here referred to as P2_{dbl}. To ensure that mutation of the Hp of *Msm*ClpP1 and *Msm*ClpP2, did not alter the overall structure of each protein, we examined the oligomeric state of each Hp mutant using AUC and compared them to the wild type proteins (Fig. S4). Importantly, with the exception of P1_{Y60A}, the oligomeric state (as determined by AUC) of each mutant was largely unaffected (Fig. S4). A similar trend for ClpP1 mutants was also observed using Native-PAGE (Fig. S4), although in this case, again with the exception of P1_{Y60A} the gel appeared to stabilize the 14-mer. Stabilisation of the 14-mer by Native-PAGE has also been observed for human ClpP²⁹. Surprisingly, neither the 7-mer nor the 14-mer complexes of ClpP2 were observed in Native-PAGE.

Next, we compared the peptidase activity of the wild type *Msm*ClpP1P2 complex with the various *Msm*ClpP1 and *Msm*ClpP2 mutant proteins (Fig. S5). Interestingly, despite the changes to the oligomeric state of P1_{Y60A} only a small change in the rate of AAF-amc turnover (by P1_{Y60A}P2) was observed (Fig. S5a), while in contrast, the rate of LY-amc turnover by P1_{Y60A}P2 was unexpectedly increased by ~4-fold (Fig. S5b). Collectively these data appear to suggest that the replacement of aromatic residues (e.g. Y60A) within the Hp of ClpP1 may affect the conformation of the substrate binding pocket (S1) directly, which in the presence of the peptide agonist z-LL, could modulate substrate affinity, specificity and/or cleavage. Consistently, ClpP1 is directly responsible for the turnover of both peptide substrates, as replacement of the active site Ser of *Msm*ClpP1 with Ala (P1_{in}) abolished the turnover of both substrates (Fig. S3). In contrast to P1_{Y60A} the relative peptidase activity of P1_{Y110A} (by the P1_{Y110A}P2 complex) was essentially unchanged for both substrates (Fig. S5a,b). Importantly, given that P1_{Y110A} did not exhibit any substrate-dependent defects in peptidase activity, we limited our subsequent analysis to the ClpP1_{Y110A}P2 complex. Next, we examined the effect of the different Hp mutations in *Msm*ClpP2. Surprisingly, the peptidase rate for both substrates was significantly reduced (by up to 80%) for both single point mutants (Y79A and L129A) of *Msm*ClpP2 (P2_{Y79A} and P2_{L129A}) (Fig. S5c,d). Intriguingly, the rate of peptide degradation by the double mutant (P2_{dbl}) was unchanged for both substrates (Fig. S5c,d, compare columns 1 and 4). Currently it remains unclear why the peptidase activity of the two single mutants in complex with wild type ClpP1 is reduced. However, given

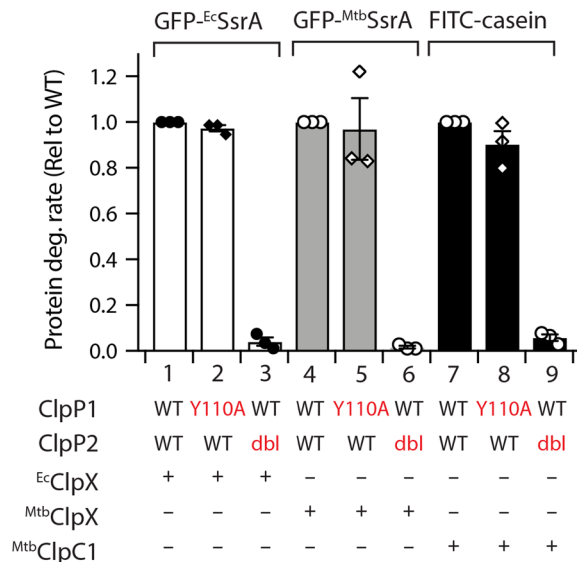


Figure 4. *Msm*ClpP1P2 form an obligate single-headed complex with its cognate ATPase components. (a) The *Ec*ΔNClpX-mediated degradation of GFP-*Ec*SsrA is mediated by docking to *Msm*ClpP2. Although mutation of the Hp residue (Y110) to Ala did not affect the ATPase-mediated delivery of GFP-*Ec*SsrA (red squares), mutation of the Hp residues (Y79 and Y129) abolished turnover (blue triangles). (b) The rate of either GFP-*Ec*SsrA degradation by *Ec*ΔNClpX (white bars), GFP-*Mtb*SsrA degradation by *Mtb*ClpX (grey bars) or FITC-casein degradation by *Mtb*ClpC1 (black bars) was determined from three independent experiments (n = 3) using either wild type ClpP1P2 (lanes 1, 4 and 7), ClpP1_{Y110A}P2 (lanes 2, 5 and 8) or ClpP1P2_{dbl} (lanes 3, 6 and 9). Error bars represent SEM.

that the double mutant exhibited a similar peptidase activity to the wild type complex, all further analysis was limited to the ClpP1P2_{dbl} complex.

***Msm*ClpP1P2 forms asymmetric complexes with its cognate ATPase.** To determine the mode of ATPase docking to the *Msm*ClpP1P2 complex, we examined the degradation of several ATPase-dependent substrates in the presence of wild type or mutant complexes of *Msm*ClpP1P2. Initially, we examined the ATPase-dependent turnover of GFP-*Ec*SsrA by Hp mutants of *Msm*ClpP1 (in the presence of wild type ClpP2) using the non-cognate ATPase, *Ec*ΔNClpX. Consistent with recent findings³¹, *Ec*ΔNClpX was able to mediate the turnover of GFP-*Ec*SsrA by *Msm*ClpP1P2, either in the absence or presence of the peptide activator, z-LL. Significantly, despite differences in peptidase activity of the various ClpP1 Hp mutants (Fig. S5) the *Ec*ΔNClpX-mediated turnover of GFP-*Ec*SsrA by each mutant protein complex was equivalent to the wild type complex (Fig. S6). Next, we examined the *Ec*ΔNClpX-mediated turnover of the same substrate by ClpP1P2 complexes bearing Hp mutations in ClpP2 (Fig. S6). Importantly, mutation of either or both Hp residues on ClpP2 abolished the ATPase-mediated turnover of GFP-*Ec*SsrA (Fig. S6). Collectively, these data suggest that the ATPase component docks exclusively to ClpP2. To confirm the asymmetric nature of the *Msm*ClpP1P2 complex we monitored the turnover of two additional model substrates that are mediated by Mycobacterial ATPase components (i.e. GFP-*Mtb*SsrA for ClpX and fluorescently-labelled model unfolded protein, FITC-casein for ClpC1). In this case, given that the “ClpP docking loop” of each ATPase component (ClpX and ClpC1) is conserved across *Msm* and *Mtb* (Fig. S7), the turnover of each substrate by the various Hp mutant complexes was examined in the presence of either *Mtb*ClpX or *Mtb*ClpC1, respectively. Consistent with the asymmetric binding observed for *Ec*ΔNClpX, the *Mtb*ClpX-dependent turnover of GFP-*Mtb*SsrA was unaffected by mutation of the Hp in *Msm*ClpP1 (Fig. 4 lane 5). In contrast, the equivalent Hp mutation in *Msm*ClpP2 completely abolished substrate turnover (Fig. 4, lane 6). Similarly, the *Mtb*ClpC1-dependent turnover of FITC-casein was unaffected by mutation of the Hp in *Msm*ClpP1 (Fig. 4, lane 8), while the equivalent Hp mutation in *Msm*ClpP2 effectively abolished the turnover of FITC-casein (Fig. 4, lane 9). Collectively these data indicate that, mutation of the Hp in *Msm*ClpP2 is sufficient to completely abolish the turnover of substrates by all ATPases tested. Indeed, both the cognate ATPases (*Mtb*ClpX and *Mtb*ClpC1) and the heterologous ATPase (*Ec*ClpX) bind specifically to *Msm*ClpP2 and not to *Msm*ClpP1 forming single-headed, asymmetric complexes (ClpP1P2X and ClpP1P2C1). These data are consistent with the findings by Weber-Ban and colleagues who showed that the *Mtb*ClpP1P2 complex only binds to its partner unfoldase components through ClpP2²⁵.

Next, to better understand the molecular basis of this asymmetric specificity we compared the Hp residues of several ClpP homologs. From this analysis we noticed that in contrast to most ClpP sequences, *Msm*ClpP1 contained an additional aromatic residue (Y88) within the Hp and speculated that this aromatic residue inhibited ATPase docking (Fig. S8). To test this hypothesis, we first replaced the equivalent residue in *Ec*ClpP (I104) with tyrosine to generate *Ec*ClpP_{I104Y} and tested the ability of this mutant protein to interact with its cognate ATPase components (ClpA and ClpX). Consistent with the idea that Y88 is a crucial inhibitory element within the Hp, *Ec*ClpP_{I104Y} prevented the ClpA-mediated degradation of GFP-*Ec*SsrA (Fig. 5a). However, the same mutation had

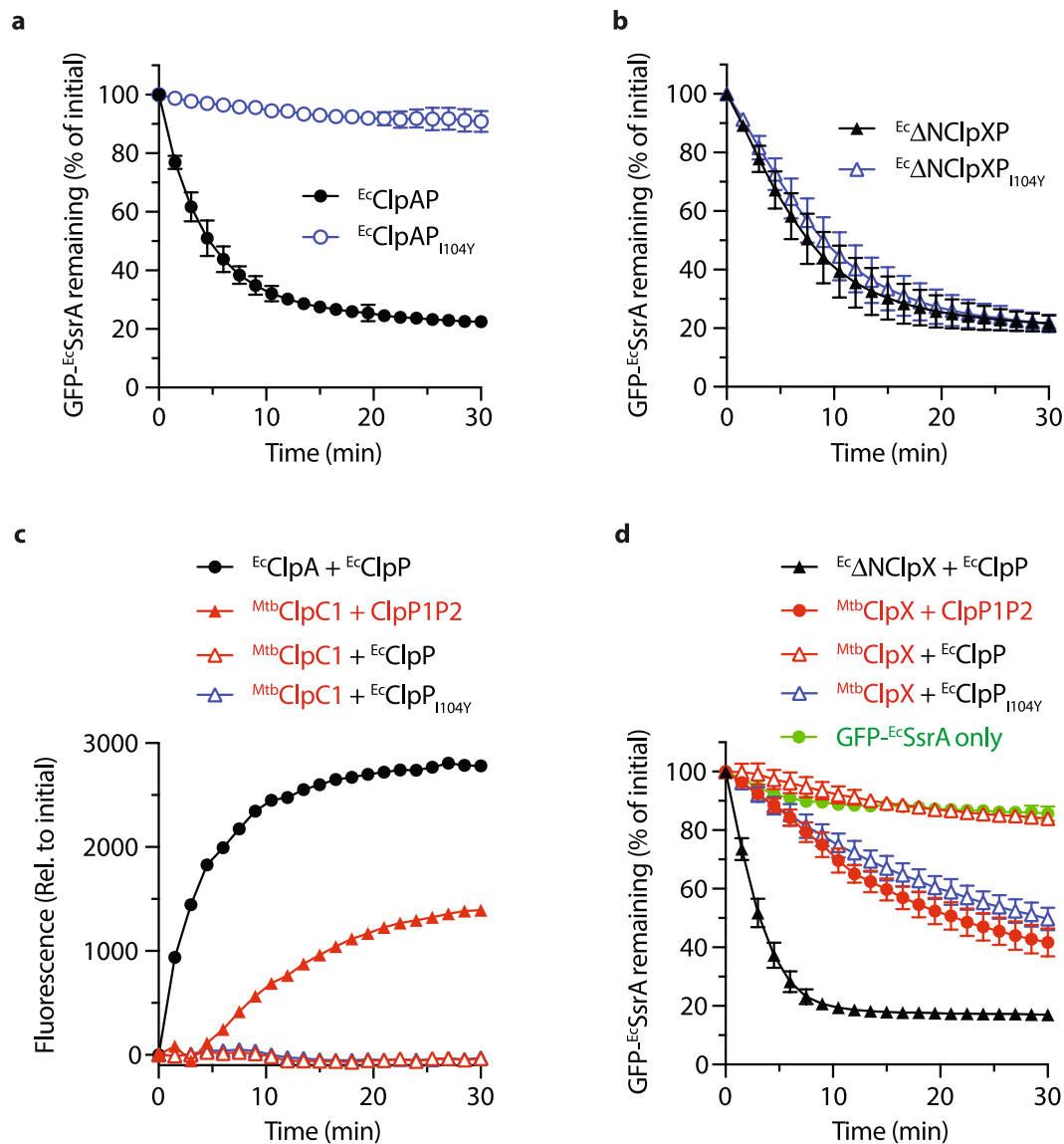


Figure 5. Single point mutations in the Hp of *EcClpP* inhibit ClpA-mediated substrate-turnover but facilitate *MtbClpX*-mediated substrate-turnover. The degradation of GFP-*EcSsrA* mediated by (a) *EcClpA* and (b) *EcΔNClpX* was monitored by fluorescence in the presence of wild type *EcClpP* (black symbols) or *EcClpP*_{I104Y} (blue symbols). (c) The degradation of FITC-casein was monitored by fluorescence in the presence of either *EcClpAP* (filled circles) or *MtbClpC1* (triangles) together with *MsmClpP1P2* (filled triangles), *EcClpP* (open red triangles) or *EcClpP*_{I104Y} (open blue triangles). (d) The turnover of GFP-*EcSsrA* was monitored by fluorescence in the absence of any addition (green circles) or in the presence of either *EcΔNClpXP* (filled black triangles) or *MtbClpX* together with either *MsmClpP1P2* (filled red circles), *EcClpP* (open red triangles) or *EcClpP*_{I104Y} (open blue triangles). Degradation rates were determined from three independent experiments (n = 3). Error bars represent SEM.

no effect on the ClpX-mediated turnover of SsrA-tagged GFP (Fig. 5b). Interestingly, although the introduction of tyrosine at residue 104 (in *EcClpP*) was sufficient to inhibit ClpA docking, *EcClpP* (bearing Ile at residue 104) was unable to functionally interact with either *MtbClpC1* (Fig. 5c, open red triangles) or *MtbClpX* (Fig. 5d, open red triangles). In contrast to ClpA docking, replacement of Ile104 with Tyr in *EcClpP* (*EcClpP*_{I104Y}), did not affect *EcClpX*-docking. However, this mutation was sufficient to recover wild type-like activity with *MtbClpX*. Collectively these data indicate that tyrosine (at residue 104/88) although inhibitory to ClpA and ClpC1 docking is permissive to ClpX docking. This in turn suggests that residue 104/88 may act as a “sensor” of ATPase docking and not simply a key inhibitory element within *MsmClpP1*, that obstructs ATPase docking.

Structural basis for docking asymmetry of *MsmClpP1P2*. To better understand the molecular features that define ATPase docking we crystallised *MsmClpP1* (Fig. 6). The structure of *MsmClpP1* was determined to 2.0 Å resolution and refined to an R_{free} value of 21.5% (Table 1). Similar to other ClpP structures, *MsmClpP1* was composed of two heptameric rings stacked back-to-back (the asymmetric unit consisted of one heptamer, the

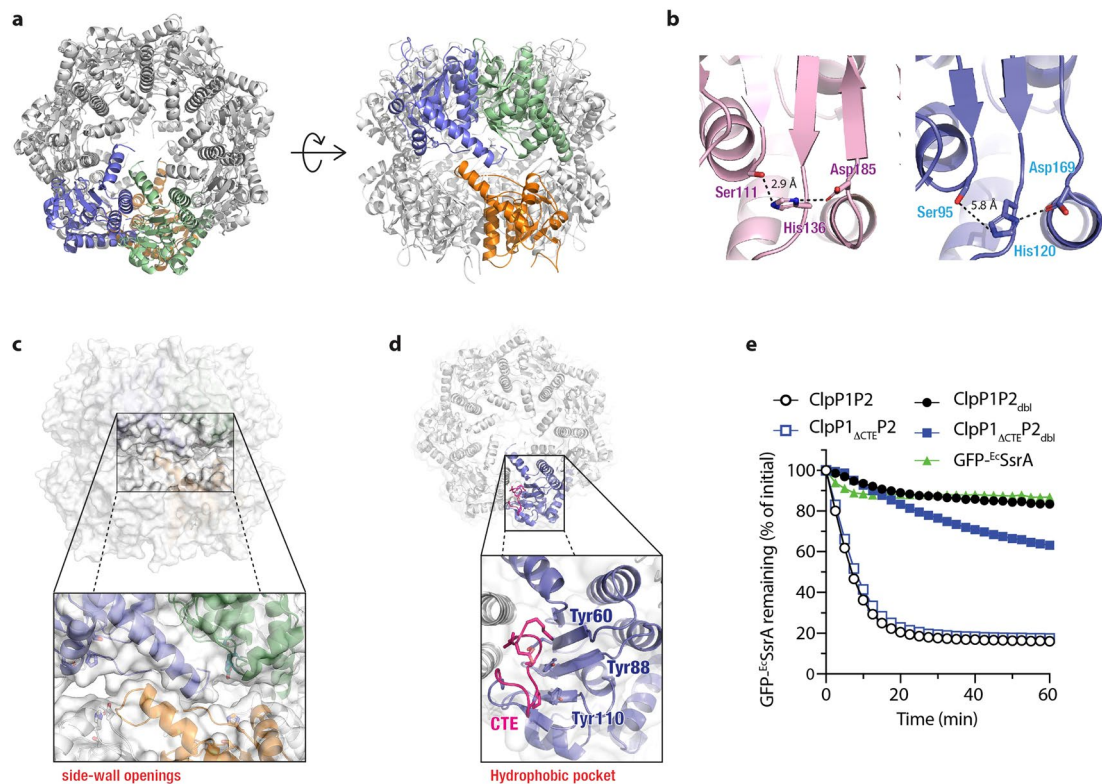


Figure 6. The C-terminal extension (CTE) of *Msm*ClpP1 obstructs ATPase docking and substrate delivery. (a) Ribbon representation of the oligomeric tetradecamer of *Msm*ClpP1 in top view (left panel) and side-view (right panel). Individual subunits are indicated in green, blue and orange. (b) Close up view of the catalytic triad of *Msm*ClpP1 (slate blue) in comparison to *Ec*ClpP (light pink, PDB code 3MT6) demonstrating that the His120 is dislocated ~ 5.8 Å from the catalytic Ser residue (Ser 95) (c) Surface representation of *Msm*ClpP1 tetradecamer highlighting side-wall openings located between the rings, adjacent to the catalytic triad. (d) Three tyrosine residues (Tyr60, Tyr88 and Tyr110) line the Hp and interact with the CTE (pink). (e) The *Ec* Δ ClpX-mediated degradation of GFP-*Ec*SsrA was monitored by fluorescence in the presence of various ClpP1P2 complexes. Although the turnover of GFP-*Ec*SsrA by wild type ClpP1P2 (open black circles) was not affected by deletion of the CTE (open blue squares) it was blocked by mutation of the Hp in ClpP2 (filled black circles). Importantly, degradation via the ClpP1P2_{dbl} was recovered when the CTE was removed from ClpP1 (filled blue squares). Degradation assays were determined from three independent experiments (n = 3). Error bars represent SEM.

biological tetradecamer is formed from crystallographically related heptamers). The tetradecamer forms a compact barrel-shaped oligomer with approximate dimensions 86×102 Å (Fig. 6a). The overall fold of the *Msm*ClpP1 protomer is similar to most ClpP structures; structural superposition of *Msm*ClpP1 with *Mtb*ClpP1 (2CBY)³², *Mtb*ClpP1P2 (4U0G)²⁴ and *Ec*ClpP (3MT6)¹⁷ resulted in overall r.m.s.d. values of 0.8 Å (1183 C α aligned), 2.3 Å (1024 C α aligned) and 2.6 Å (1071 C α aligned), respectively. However, there are a number of notable differences between these structures. Firstly, the structure of *Msm*ClpP1, similar to *Mtb*ClpP1³², is in an inactive conformation, as His120 is dislocated from the catalytic triad (i.e. the distance between Ser95 and His120 is 5.8 Å (Fig. 6b, right panel) (6.55 Å for the equivalent residues *Mtb*ClpP1). In contrast, in the ClpP1 component of the *Mtb*ClpP1P2 complex (4U0G) and in *Ec*ClpP, these residues are only 2.9 Å apart (Fig. 6b, left panel). Another key difference between these structures relates to the formation of the tetradecamer, via the handle domains. In *Ec*ClpP and *Mtb*ClpP1P2 complex, the two heptameric rings associate via the formation of an antiparallel β -sheet in the handle domains, whereby each β -sheet is composed of strands from opposing monomers across the ring-ring interface. Compared to these active structures, both *Msm*ClpP1 and *Mtb*ClpP1, have shorter handle domains, which lack a structured β -strand region. As a result, association of the heptameric rings in *Msm*ClpP1 (and *Mtb*ClpP1) is mediated largely by a *trans* association of a short section of the α -helix in the handle domain creating several openings in the side-walls of the ClpP1 tetradecamer (Figs. 6c and S9). Although handle flexibility has been observed in a selection of inactive ClpP structures, the size and precise location of pores resulting from this flexibility has been difficult to assess due to the large number of unmodeled residues in these structures. Significantly, the breaches in the side wall of *Msm*ClpP1 are located adjacent to the catalytic residues of the protease and as such provide a direct path between the active site of the protease and the external solution. Hence as has been proposed by Kay and Houry³³, these openings could represent an exit portal for the egress of peptides. Therefore, we postulate that this structure of *Msm*ClpP1 represents a post “substrate-cleavage” snapshot of the protease. Following peptide cleavage, this protease acquires an inactive conformation where the catalytic His residue adopts a distorted configuration,

Data collection	
PDB ID: 6BPU	
Wavelength (Å)	0.9537
Oscillation, total degrees (°)	(1°/180°)
Resolution range (Å)*	50.0–2.00 (2.03–2.00)
Space group	$P4_22_12$
Unit-cell parameters (Å, °)	$\alpha = \beta = 169.62$; $\gamma = 114.22$ $\alpha = \beta = \gamma = 90$
No. of molecules in asymmetric unit	7
Observed reflections	1386, 431
Unique reflections	111, 562
R _{pim}	0.035 (0.364)
Completeness (%) [*]	99.7 (98.3)
$\langle I/\sigma(I) \rangle$ [*]	9.29 (2.9)
Multiplicity [*]	12.4 (11.0)
Refinement	
Resolution (Å) [*]	37.15–2.00 (2.07–2.00)
R _{factor} ^{a*}	0.178 (0.224)
R _{free} ^{b*}	0.214 (0.260)
No. of non-hydrogen atoms	
- Protein	10172
- Waters	1087
- Malonate (molecules)	63
B factors	
- Wilson (Å ²)	26.68
- Average B factor (Å ²): All atoms	31.32
- Average B factor (Å ²): Macromolecules	30.29
- Average B factor (Å ²): Ligands	51.60
- Average B factor (Å ²): Solvent	39.80
R.m.s.d. from ideal geometry	
- Bonds (Å)	0.007
- Angles (°)	0.88
MolProbity analysis	
- Ramachandran favored/outliers (%)	96.50/0.3
- Clashscore [percentile]	5.21 (97th)

Table 1. Crystallographic data and refinement statistics. *Values in parentheses refer to the highest resolution shell. ^aR_{fac} = $\sum_h |F_o - F_c| / \sum_h |F_o|$, where F_o and F_c are the observed and calculated structure-factor amplitudes for each reflection “h”. ^bR_{free} was calculated with 5% of the diffraction data selected randomly and excluded from refinement.

which initiates a conformational change in the adjacent handle domain whereby the β -strand that typically forms an antiparallel β -sheet with the equivalent β -strand in the opposite monomer, becomes disordered. As a result of this disorder in the β -strand, several interactions which stabilise the tetradecamer are lost. This results in a shorter interface between opposing subunits in the tetradecamer and opening of “large pores” along the surface of the protein which could allow release of the cleaved peptides. Smaller solvent exposed channels have also been observed in the side walls of the *Mtb*ClpP1P2 complex (Fig. S9), however these openings were only located between ClpP1 subunits, above the equatorial interface, ~15 Å from the catalytic triad²⁴.

Another distinctive feature of *Msm*ClpP1, which is absent in many ClpP homologues including *Mtb*ClpP1^{24,32}, is a short C-terminal extension (CTE) ~12 residues long, which docks into its own Hp (Fig. 6d). Interestingly, the CTE interacts with residues within the Hp of ClpP in a similar manner to ADEP^{16,17,24}. Specifically, the CTE interacts with three tyrosine residues (Y60, Y88 and Y110) within the Hp (Fig. 6d). Hence, we speculated that, similar to ADEPs, the CTE of *Msm*ClpP1 may restrict ATPase docking. To test this idea, we deleted the CTE of *Msm*ClpP1 and monitored the ability of various wild type and mutant *Msm*ClpP1P2 complexes to interact with different ATPase components (Figs. 6e and S10). First, we examined the *Ec*ClpX-mediated turnover of GFP-*Ec*SsrA by *Msm*ClpP1 Δ CTE in the presence of wild type *Msm*ClpP2. Importantly, this complex retained wild type-like activity against each ATPase/substrate combination tested, demonstrating that removal of the CTE does not alter ClpP1P2 activity (Fig. 6e, open symbols). Next, in order to switch ATPase-docking specificity (from *Msm*ClpP2 to *Msm*ClpP1), we repeated the above experiments in the presence of *Msm*ClpP2_{dbl} (which is unable to dock to any of the ATPase components tested). Remarkably, deletion of the CTE was sufficient to facilitate the *Ec*ClpX-mediated turnover of GFP-*Ec*SsrA (Fig. 6e, filled squares). These data demonstrate that the CTE of *Msm*ClpP1 plays a crucial inhibitory role in ATPase docking. Next, we monitored the docking of the cognate ATPase components (*Mtb*ClpX

and *Mtb*ClpC1) to the ClpP1 $_{\Delta$ CTE $_{P2}$ complex. Surprisingly, neither *Mtb*ClpX nor *Mtb*ClpC1 were able to mediate the turnover of either GFP-*Mtb*SsrA (Fig. S10b) or FITC-casein (Fig. S10c) respectively. Taken together, these data suggest that although the CTE obstructs ATPase docking (of *Ec*ClpX), removal of this feature alone is not sufficient to permit docking to the physiological relevant ATPase components (*Mtb*ClpX/*Mtb*ClpC1). Hence, both the presence of the CTE (that occludes the Hp of ClpP1) and the altered specificity of the Hp may have important implications for the development of novel ADEP-like antibiotics that dysregulate Clp proteases.

Interestingly, several ClpP homologs contain a C-terminal extension (including human ClpP which contains a 28-residue extension). To date however, the role of these CTEs has remained unclear. In contrast to our structure of *Msm*ClpP1, the CTE of *Hs*ClpP extends away from the heptameric ring²⁷. Despite this “snapshot” placing the CTE of human ClpP away from the ATPase-interface we speculate that this region (in various ClpP homologues) could play an important role in regulating ClpP function *in vivo*. One possibility is that the CTE (or in the case of *Msm*ClpP1 - the atypical Hp) may have co-evolved with an alternate Clp-protease activator to further diversify the function of the peptidase. Consistent with this idea, novel (ATP-independent) Clp protease activators have recently been identified in plants³⁴. Likewise, the Mycobacterial proteasome has also been shown to function with a variety of activators (both ATP-dependent and ATP-independent)^{3,35–38}. Therefore, we speculate that novel ClpP activators may exist in a variety of species which contain ClpP homologs with either an extended C-terminus or an atypical Hp.

In conclusion, our findings clearly demonstrate that, similar to *Mtb*ClpP1P2, the hetero-oligomeric Clp-protease complex in *Msm* is highly regulated. Although processing of *Msm*ClpP2 (and hence “activation” of the *Msm*ClpP1P2 complex) can proceed in the absence of an activator, the *in vitro* peptidase activity of *Msm*ClpP1P2 (for peptide and protein turnover) requires either a cognate ATPase component or a chemical activator (i.e. z-LL). Significantly, both cognate ATPase components dock to only one face of the peptidase. This asymmetry provides direct competition for ATPase-docking to the peptidase, and as a result (dependent on the abundance of each component), likely controls the delivery of specific substrates to this peptidase and hence their turnover in a tightly regulated fashion. The specific docking of both ATPase components to a single face of this machine may also provide an opportunity for further diversification of the peptidase, through docking of additional specific activators to the vacant platform on the ClpP1P2 complex.

Material and Methods

Cloning. *Msm clpP1* (MSMEG_4673) and *clpP2* (MSMEG_4672) were amplified from *M. smegmatis* mc²155 genomic DNA (kindly provided by Prof. R. Manganelli) with specific primers (see Table S1), using Phusion DNA polymerase (New England Biolabs). *Mtb clpP1* (Rv2460c) and *clpP2* (Rv2461c) were amplified from *Mtb* genomic DNA (kindly provided by Ms. M. Globan, VIDRL). The amplified DNA was digested with the appropriate restriction enzymes and ligated into similarly digested plasmids. Fragments coding for unprocessed ClpP1 and ClpP2 were cloned into pET10C (and pET10N)³⁹ to generate either a C- or N-terminal His₁₀-tagged fusion protein, while fragments coding for processed forms of ClpP1 and ClpP2 (i.e. lacking propeptide) were cloned into pHUE⁴⁰ to generate N-terminal His₆-Ub fusion proteins. *Mtb clpX* cloned into pET15(b) was a kind gift from Dr. P. Genevax (Université de Toulouse, France), and *Mtb clpC1* cloned into pET30(a) was a kind gift from Dr. D. Vasudevan (Institute of Life Sciences, Bhubaneswar, India)⁴¹. The SsrA-tag from *Mtb/Msm* (AADSNQRDYALAA) was cloned into pDD173⁴² by annealing specific primers (see Table S1) to generate a C-terminal GFP fusion (GFP-*Mtb*SsrA). All clones (Table S2) were verified by nucleotide sequencing.

Protein expression and purification. His₆-tagged *Ec*ClpX and *Ec*ClpP were expressed in *E. coli* and purified as described previously⁴³. *Mtb*ClpC1 and *Mtb*ClpX were both expressed as C-terminal His₆-tagged fusion proteins. *Mtb*ClpC1-His₆ was expressed as described in⁴¹, while *Mtb*ClpX-His₆ was expressed at 16 °C in *E. coli* BL21 (DE3) codon+ RIL cells, following the addition of 0.5 mM IPTG. ClpP1 and ClpP2 (from either *Msm* or *Mtb*) were expressed, with either a C-terminal His₁₀ tag, essentially as described in³⁹ or alternatively with a His₆-Ub tag (which was subsequently cleaved) as described⁴⁴. All His₆-tagged fusion proteins were purified using Ni-NTA-agarose beads as described previously⁴³, while His₁₀-tagged fusion proteins were purified essentially as described⁴⁵. Untagged ClpP1 and ClpP2 (wild type and specific point mutants) were also generated using the Ub-fusion system⁴⁰ and purified essentially as described, using a combination of IMAC and preparative grade size exclusion chromatography (SEC).

Protein analysis by electrophoresis. For the analysis of protein purity and protein turnover, samples were separated by 16.5% Tricine SDS-PAGE⁴⁶. To analyse the oligomerisation/native structure of wild type and mutant ClpP1 complexes, 5 µg of purified protein was separated using 4–16% Native-PAGE Novex Bis-Tris gels (Invitrogen) essentially as described²⁹ and visualised by staining with Coomassie Brilliant Blue (CBB) R250.

Processing and peptidase assays. For processing assays, either full length *Msm* or *Mtb* ClpP1 was incubated together with full length ClpP2 (from either *Msm* or *Mtb*) at 30 °C in Buffer XP (25 mM Tris-HCl pH 7.5, 100 mM KCl, 100 mM NaCl, 20 mM MgCl₂, 10% (v/v) glycerol, 0.025% (v/v) Triton X-100, 1 mM DTT) in the absence or presence of the activator, z-Leu-Leu-H (z-LL). Processing of ClpP (2.8 µM) was initiated, either by addition of the activator (0.5 mM), or equimolar amounts of the partner ClpP protein. The reaction was stopped (at the indicated time points), by the addition of sample buffer followed by incubation at 95 °C for 5 min. Processing of *Msm*ClpP was analysed by 16.5% Tris-tricine SDS-PAGE while processing of *Mtb*ClpP was analysed by 15% glycine SDS-PAGE. To monitor the peptidase activity of wild type and mutant ClpPs, peptide degradation assays were performed using fluorescently labelled peptides essentially as described²⁹. Similarly, the turnover of GFP-*Ec*SsrA, GFP-*Mtb*SsrA or FITC-casein (by 1 µM *Ec*ΔNClpX, *Mtb*ClpX or *Mtb*ClpC1, respectively) was and monitored by fluorescence as described²⁹. All peptide and protein degradation assays were performed in the presence of 0.5 mM z-LL (unless otherwise stated).

Crystallisation of *Msm*ClpP1 and diffraction data collection. Purified *Msm*ClpP1 (in 50 mM Tris-HCl pH 8.0, 200 mM KCl) was concentrated using a Centricon-30 centrifugal concentrator (Amicon) up to 30 mg/mL. *Msm*ClpP1 crystallization experiments were performed using the vapor diffusion method. Initial high-throughput crystallization experiments were performed in house or at the CSIRO Collaborative Crystallization Centre (www.csiro.au/C3; Melbourne, Australia). For crystal optimization experiments, drops were set in 24-well plates by mixing 1 μ L of protein with 1 μ L of well condition and drops were equilibrated at 20 °C against a reservoir volume of 500 μ L. Small hexagonal crystals were obtained in 1.8M sodium malonate pH 6.4. After two rounds of optimisation, which included changing pH, malonate concentration and protein concentration a significant improvement in the quality of the crystals was obtained. Larger hexagonal crystals (with approximate dimensions 0.3 mm \times 0.2 mm \times 0.1 mm) were obtained from solutions consisting of 2.1–3.4M sodium malonate pH 6.6–6.8 and a protein concentration of 5.5 mg/mL.

The crystals were cryoprotected using 3M sodium malonate pH 6.5–6.9 and flash cooled in liquid nitrogen. Diffraction data was collected at 100 K at the protein crystallography beamline MX2 at the Australian synchrotron using an ADSC Quantum 315r detector. 1° oscillation images were collected for a total of 180° using a crystal-to-detector distance of 260 mm. Diffraction data were integrated and scaled with HKL2000⁴⁷.

Structure determination and refinement. The crystal structure of *Msm*ClpP1 was solved by molecular replacement using BALBES⁴⁸ using the structure of *M. tuberculosis* ClpP as a search model (PDB: 2CBY, sequence identity 96%). The model was built using Coot⁴⁹ and refined using phenix.refine⁵⁰ and TLS (translation/libration/screw) refinement⁵¹. Most of the structure could be unambiguously assigned in the electron density map except residues 1–10 at the N-terminus of each chain, and the loop region between residues 125–129, which in chains A, B, C and E were difficult to model because of poor density. The final model was validated using Molprobity⁵². Table 1 provides the statistics for the X-ray data collection and final refined model. All structural figures were generated with PyMOL. Superposition of molecules was carried out using the Secondary Structure Matching (SSM) option from the program Coot⁴⁹.

Received: 16 August 2019; Accepted: 25 October 2019;

Published online: 02 December 2019

References

- Gur, E., Ottofueling, R. & Dougan, D. A. Machines of destruction - AAA+ proteases and the adaptors that control them. *Subcell Biochem* **66**, 3–33, https://doi.org/10.1007/978-94-007-5940-4_1 (2013).
- Olivares, A. O., Baker, T. A. & Sauer, R. T. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat Rev Microbiol* **14**, 33–44, <https://doi.org/10.1038/nrmicro.2015.4> (2016).
- Alhuwaidar, A. A. H. & Dougan, D. A. AAA+ Machines of Protein Destruction in Mycobacteria. *Front Mol Biosci* **4**, 49, <https://doi.org/10.3389/fmolb.2017.00049> (2017).
- Elsholz, A. K. W., Birk, M. S., Charpentier, E. & Turgay, K. Functional Diversity of AAA+ Protease Complexes in *Bacillus subtilis*. *Front Mol Biosci* **4**, 44, <https://doi.org/10.3389/fmolb.2017.00044> (2017).
- Mahmoud, S. A. & Chien, P. Regulated Proteolysis in Bacteria. *Annu Rev Biochem* **87**, 677–696, <https://doi.org/10.1146/annurev-biochem-062917-012848> (2018).
- Striebel, F., Kress, W. & Weber-Ban, E. Controlled destruction: AAA+ ATPases in protein degradation from bacteria to eukaryotes. *Curr Opin Struct Biol* **19**, 209–217, <https://doi.org/10.1016/j.sbi.2009.02.006> (2009).
- Maurizi, M. R., Clark, W. P., Kim, S. H. & Gottesman, S. Clp P represents a unique family of serine proteases. *The Journal of biological chemistry* **265**, 12546–12552 (1990).
- Wang, J., Hartling, J. A. & Flanagan, J. M. The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. *Cell* **91**, 447–456 (1997).
- Liu, K., Ologbenla, A. & Houry, W. A. Dynamics of the ClpP serine protease: a model for self-compartmentalized proteases. *Crit Rev Biochem Mol Biol* **49**, 400–412, <https://doi.org/10.3109/10409238.2014.925421> (2014).
- Kim, Y. I. *et al.* Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nat Struct Biol* **8**, 230–233, <https://doi.org/10.1038/84967> (2001).
- Martin, A., Baker, T. A. & Sauer, R. T. Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease. *Mol Cell* **27**, 41–52, <https://doi.org/10.1016/j.molcel.2007.05.024> (2007).
- Leung, E. *et al.* Activators of cylindrical proteases as antimicrobials: identification and development of small molecule activators of ClpP protease. *Chemistry & biology* **18**, 1167–1178, <https://doi.org/10.1016/j.chembiol.2011.07.023> (2011).
- Martin, A., Baker, T. A. & Sauer, R. T. Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nat Struct Mol Biol* **15**, 1147–1151, <https://doi.org/10.1038/nsmb.1503> (2008).
- Brotz-Oesterhelt, H. *et al.* Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat Med* **11**, 1082–1087, <https://doi.org/10.1038/nm1306> (2005).
- Kirstein, J. *et al.* The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO Mol Med* **1**, 37–49, <https://doi.org/10.1002/emmm.200900002> (2009).
- Lee, B. G. *et al.* Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nat Struct Mol Biol* **17**, 471–478, <https://doi.org/10.1038/nsmb.1787> (2010).
- Li, D. H. *et al.* Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: A model for the ClpX/ClpA-bound state of ClpP. *Chemistry & biology* **17**, 959–969, <https://doi.org/10.1016/j.chembiol.2010.07.008> (2010).
- Balogh, D. *et al.* Insights into ClpXP proteolysis: heterooligomerization and partial deactivation enhance chaperone affinity and substrate turnover in *Listeria monocytogenes*. *Chem Sci* **8**, 1592–1600, <https://doi.org/10.1039/c6sc03438a> (2017).
- Cohen, S. E., McKnight, B. M. & Golden, S. S. Roles for ClpXP in regulating the circadian clock in *Synechococcus elongatus*. *Proc Natl Acad Sci USA* **115**, E7805–E7813, <https://doi.org/10.1073/pnas.1800828115> (2018).
- Dahmen, M., Vielberg, M. T., Groll, M. & Sieber, S. A. Structure and mechanism of the caseinolytic protease ClpP1/2 heterocomplex from *Listeria monocytogenes*. *Angewandte Chemie* **54**, 3598–3602, <https://doi.org/10.1002/anie.201409325> (2015).
- Akopian, T. *et al.* The active ClpP protease from *M. tuberculosis* is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO J* **31**, 1529–1541, <https://doi.org/10.1038/emboj.2012.5> (2012).
- Lavey, N. P., Shadid, T., Ballard, J. D. & Duerfeldt, A. S. Clostridium difficile ClpP Homologues are Capable of Uncoupled Activity and Exhibit Different Levels of Susceptibility to Acyldepsipeptide Modulation. *ACS Infect Dis* **5**, 79–89, <https://doi.org/10.1021/acscinfed.8b00199> (2019).

23. Wood, N. A. *et al.* Initial Characterization of the Two ClpP Paralogs of *Chlamydia trachomatis* Suggests Unique Functionality for Each. *Journal of bacteriology* **201**, <https://doi.org/10.1128/JB.00635-18> (2019).
24. Schmitz, K. R., Carney, D. W., Sello, J. K. & Sauer, R. T. Crystal structure of *Mycobacterium tuberculosis* ClpP1P2 suggests a model for peptidase activation by AAA+ partner binding and substrate delivery. *Proc Natl Acad Sci USA* **111**, E4587–4595, <https://doi.org/10.1073/pnas.1417120111> (2014).
25. Leodolter, J., Warweg, J. & Weber-Ban, E. The *Mycobacterium tuberculosis* ClpP1P2 Protease Interacts Asymmetrically with Its ATPase Partners ClpX and ClpC1. *PLoS one* **10**, e0125345, <https://doi.org/10.1371/journal.pone.0125345> (2015).
26. Personne, Y., Brown, A. C., Schuessler, D. L. & Parish, T. *Mycobacterium tuberculosis* ClpP proteases are co-transcribed but exhibit different substrate specificities. *PLoS one* **8**, e60228, <https://doi.org/10.1371/journal.pone.0060228> (2013).
27. Kang, S. G., Maurizi, M. R., Thompson, M., Mueser, T. & Ahvazi, B. Crystallography and mutagenesis point to an essential role for the N-terminus of human mitochondrial ClpP. *J Struct Biol* **148**, 338–352, <https://doi.org/10.1016/j.jsb.2004.07.004> (2004).
28. Benaroudj, N., Raynal, B., Miot, M. & Ortiz-Lombardia, M. Assembly and proteolytic processing of mycobacterial ClpP1 and ClpP2. *BMC Biochem* **12**, 61, <https://doi.org/10.1186/1471-2091-12-61> (2011).
29. Brodie, E. J., Zhan, H., Saiyed, T., Truscott, K. N. & Dougan, D. A. Perrault syndrome type 3 caused by diverse molecular defects in CLPP. *Sci Rep* **8**, 12862, <https://doi.org/10.1038/s41598-018-30311-1> (2018).
30. Gersch, M., List, A., Groll, M. & Sieber, S. A. Insights into structural network responsible for oligomerization and activity of bacterial virulence regulator caseinolytic protease P (ClpP) protein. *The Journal of biological chemistry* **287**, 9484–9494, <https://doi.org/10.1074/jbc.M111.336222> (2012).
31. Schmitz, K. R. & Sauer, R. T. Substrate delivery by the AAA+ ClpX and ClpC1 unfoldases activates the mycobacterial ClpP1P2 peptidase. *Molecular microbiology* **93**, 617–628, <https://doi.org/10.1111/mmi.12694> (2014).
32. Ingvarsson, H. *et al.* Insights into the inter-ring plasticity of caseinolytic proteases from the X-ray structure of *Mycobacterium tuberculosis* ClpP1. *Acta Crystallogr D Biol Crystallogr* **63**, 249–259, <https://doi.org/10.1107/S0907444906050530> (2007).
33. Sprangers, R., Gribun, A., Hwang, P. M., Houry, W. A. & Kay, L. E. Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. *Proc Natl Acad Sci USA* **102**, 16678–16683, <https://doi.org/10.1073/pnas.0507370102> (2005).
34. Kim, J. *et al.* Structures, Functions, and Interactions of ClpT1 and ClpT2 in the Clp Protease System of *Arabidopsis Chloroplasts*. *Plant Cell* **27**, 1477–1496, <https://doi.org/10.1105/tpc.15.00106> (2015).
35. Bai, L. *et al.* Structural analysis of the dodecameric proteasome activator PafE in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **113**, E1983–1992, <https://doi.org/10.1073/pnas.1512094113> (2016).
36. Hu, K. *et al.* Proteasome substrate capture and gate opening by the accessory factor PafE from *Mycobacterium tuberculosis*. *The Journal of biological chemistry* **293**, 4713–4723, <https://doi.org/10.1074/jbc.RA117.001471> (2018).
37. Delley, C. L. *et al.* Bacterial proteasome activator bpa (rv3780) is a novel ring-shaped interactor of the mycobacterial proteasome. *PLoS one* **9**, e114348, <https://doi.org/10.1371/journal.pone.0114348> (2014).
38. Muller, A. U. & Weber-Ban, E. The Bacterial Proteasome at the Core of Diverse Degradation Pathways. *Front Mol Biosci* **6**, 23, <https://doi.org/10.3389/fmolb.2019.00023> (2019).
39. Truscott, K. N. *et al.* A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. *Nat Struct Biol* **8**, 1074–1082, <https://doi.org/10.1038/nsb726> (2001).
40. Catanzariti, A. M., Soboleva, T. A., Jans, D. A., Board, P. G. & Baker, R. T. An efficient system for high-level expression and easy purification of authentic recombinant proteins. *Protein Sci* **13**, 1331–1339, <https://doi.org/10.1110/ps.04618904> (2004).
41. Vasudevan, D., Rao, S. P. & Noble, C. G. Structural basis of mycobacterial inhibition by cyclomarin A. *The Journal of biological chemistry* **288**, 30883–30891, <https://doi.org/10.1074/jbc.M113.493767> (2013).
42. Dougan, D. A., Reid, B. G., Horwich, A. L. & Bukau, B. ClpS, a substrate modulator of the ClpAP machine. *Mol Cell* **9**, 673–683 (2002).
43. Dougan, D. A., Weber-Ban, E. & Bukau, B. Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX. *Mol Cell* **12**, 373–380 (2003).
44. Ninnis, R. L., Spall, S. K., Talbo, G. H., Truscott, K. N. & Dougan, D. A. Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli*. *EMBO J* **28**, 1732–1744, <https://doi.org/10.1038/emboj.2009.134> (2009).
45. Lowth, B. R. *et al.* Substrate recognition and processing by a Walker B mutant of the human mitochondrial AAA+ protein CLPX. *J Struct Biol* **179**, 193–201, <https://doi.org/10.1016/j.jsb.2012.06.001> (2012).
46. Schagger, H. & von Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**, 368–379 (1987).
47. Otwinowski, Z. & Minor, W. In *Methods in Enzymology* Vol. Volume 276, 307–326 (Academic Press, 1997).
48. Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64**, 125–132, <https://doi.org/10.1107/S0907444907050172> (2008).
49. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132, <https://doi.org/10.1107/S0907444904019158> (2004).
50. Afonine, P. V. *et al.* Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **66**, 1153–1163, <https://doi.org/10.1107/S0907444910026582> (2010).
51. Painter, J. & Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439–450, <https://doi.org/10.1107/S0907444906005270> (2006).
52. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12–21, <https://doi.org/10.1107/S0907444909042073> (2010).

Acknowledgements

D.A.D. was supported by an Australian Research Council (ARC) Australian Research Fellowship (DP110103936) and B.H. was supported by an ARC Future Fellowship (FT130100580). The work was supported by a Victoria India Doctoral Scholarship (to J.N.), La Trobe University Postgraduate Research Scholarships (to J.E.Z. and A.A.T.) and La Trobe University Full Fee Research Scholarships (to J.N. and A.A.T.). We thank Dr. Pierre Genevaux (Université de Toulouse, France) for providing pET15(b)^{Mtb}clpX, Dr. Dileep Vasudevan (Institute of Life Sciences, Bhubaneswar, India) for providing pET30(a)^{Mtb}clpC1, Nilushika Silva for technical assistance in cloning of MsmClpP_{in} and MsmClpP_{in}. Crystals were analysed at the Australian Synchrotron. The atomic coordinates and structure factors for 6PBU have been deposited in the Protein Data Bank (<http://wwpdb.org>). Preparation of a draft manuscript was supported by a “write-up” award from the Understanding Disease Research Focus Area at La Trobe University to J.N.

Author contributions

The experiments were conceived and designed by D.A.D., B.H. and K.N.T., performed by J.N., J.P., J.Z., A.A.T. and B.H. Data analysis was carried out by J.N., B.H., J.P., K.N.T. and D.A.D. The paper was written by D.A.D., J.N., B.H., J.P. and K.N.T. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-53736-8>.

Correspondence and requests for materials should be addressed to B.H. or D.A.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019