The active ClpP protease from *M. tuberculosis* is a complex composed of a heptameric ClpP1 and a ClpP2 ring

Tatos Akopian^{1,3}, Olga Kandror^{1,3},
Ravikiran M Raju², Meera Unnikrishnan²,
Eric J Rubin² and Alfred L Goldberg^{1,*}

¹Department of Cell Biology, Harvard Medical School, Boston, MA, USA and ²Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA, USA

Mycobacterium tuberculosis (Mtb) contains two *clpP* genes, both of which are essential for viability. We expressed and purified Mtb ClpP1 and ClpP2 separately. Although each formed a tetradecameric structure and was processed, they lacked proteolytic activity. We could, however, reconstitute an active, mixed ClpP1P2 complex after identifying N-blocked dipeptides that stimulate dramatically (> 1000-fold) ClpP1P2 activity against certain peptides and proteins. These activators function cooperatively to induce the dissociation of ClpP1 and ClpP2 tetradecamers into heptameric rings, which then re-associate to form the active ClpP1P2 2-ring mixed complex. No analogous small molecule-induced enzyme activation mechanism involving dissociation and re-association of multimeric rings has been described. ClpP1P2 possesses chymotrypsin and caspase-like activities, and ClpP1 and ClpP2 differ in cleavage preferences. The regulatory ATPase ClpC1 was purified and shown to increase hydrolysis of proteins by ClpP1P2, but not peptides. ClpC1 did not activate ClpP1 or ClpP2 homotetradecamers and stimulated ClpP1P2 only when both ATP and a dipeptide activator were present. ClpP1P2 activity, its unusual activation mechanism and ClpC1 ATPase represent attractive drug targets to combat tuberculosis.

The EMBO Journal (2012) 31, 1529–1541. doi:10.1038/emboj.2012.5; Published online 27 January 2012

Subject Categories: proteins; microbiology & pathogens

Keywords: cleavage preferences; dipeptide activator; protease ClpP1P2

Introduction

Tuberculosis is a devastating disease that affects worldwide about 100 million people and causes nearly 2 million deaths annually. It has been estimated that one third of all humans is infected with latent *Mycobacterium tuberculosis* (Mtb). Moreover, Mtb has become increasingly resistant to available antibiotics. Consequently, it is important to identify and

characterize new therapeutic targets in Mtb and to synthesize selective inhibitors. Ideal targets for drug development should be enzymes essential for bacterial viability that differ in physicochemical properties and specificity from those present in humans. ClpP1 has been recently validated to be essential in Mtb (Ollinger *et al*, 2012) and in related studies, we established that the proteases encoded by *clpP1* and by *clpP2* are required both for the growth of Mtb and for its virulence during murine infection (Sassetti *et al*, 2003; Raju *et al*, 2011). Therefore, ClpP1 and ClpP2 are highly attractive drug targets, especially since they are not present in the cytosol of mammalian cells (where protein breakdown occurs primarily by the ubiquitin proteasome pathway or by lysosomes), and these enzymes differ markedly from the mitochondrial Clp complex.

ClpP is a highly conserved, multimeric serine protease originally discovered (Hwang *et al*, 1987; Katayama-Fujimura *et al*, 1987) and extensively characterized in *E. coli* (Maurizi *et al*, 1990b, 1994, 1998; Yu and Houry, 2007). ClpP homologues exist in a wide range of bacteria, as well as in mitochondria and chloroplasts in eukaryotes (Porankiewicz *et al*, 1999). The *clpP* gene in *E. coli* encodes a polypeptide of 207 amino acids, of which 14 residues are autolytically cleaved to yield the mature protein of 21.5 kDa (Maurizi *et al*, 1990a). The active enzyme is a tetradecameric structure composed of two heptameric rings that form a hollow cylinder with 14 proteolytic sites compartmentalized within its central chamber (Flanagan *et al*, 1995; Shin *et al*, 1996; Wang *et al*, 1997). Substrates enter into the chamber through two axial openings with a diameter of about 10 Å, which limits the size of polypeptides degraded. By itself, *E. coli* ClpP is able to rapidly hydrolyse only oligopeptides, but not large globular proteins. The degradation of large proteins requires the presence of an AAA ATPase complex, such as ClpA or ClpX in *E. coli* or ClpC in other species (Kress *et al*, 2009). These hexameric structures associate with both ends of ClpP to form the active 4-ring ATP-dependent protease (Maurizi, 1991; Maurizi *et al*, 1998; Kim *et al*, 2001). These ATPases bind selectively certain protein substrates, unfold them, and translocate the linearized polypeptides into the ClpP proteolytic chamber for degradation (Hoskins *et al*, 1998; Ortega *et al*, 2000; Ishikawa *et al*, 2001; Reid *et al*, 2001). In addition to substrate recognition, the human mitochondrial ClpX complex promotes the assembly of the ClpP complex into an active form (Kang *et al*, 2005).

Most organisms possess a single *clpP* gene, while some microorganisms (e.g., *Streptomyces*, *Actinomycetes*, and *Cyanobacteria*) and plants (e.g., *Arabidopsis thaliana*) have two or more *clpPs* (Porankiewicz *et al*, 1999; Viala *et al*, 2000; Peltier *et al*, 2001, 2004; Schelin *et al*, 2002; Viala and Mazodier, 2002; Butler *et al*, 2006; Sjogren *et al*, 2006; Stanne *et al*, 2007; Andersson *et al*, 2009). The functional significance of these multiple species is unclear. Mtb contains

*Corresponding author. Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA. Tel.: +1 617 432 1854;

Fax: +1 617 232 0173; E-mail: alfred_goldberg@hms.harvard.edu

³These authors contributed equally to this work

two *clpP* genes, *clpP1* and *clpP2*, both of which are essential for viability (Sasseti *et al*, 2003) and infectivity, as shown in Raju *et al* (2011). Although both appear to encode serine proteases, prior attempts (Ingvarsson *et al*, 2007; Benaroudj *et al*, 2011) to express and characterize Mtb ClpP1 and ClpP2 in *E. coli* yielded complexes that lacked proteolytic activity, as did our initial attempts to express ClpP1 and ClpP2 in *E. coli*. We hypothesized that those attempts failed because they were based on the assumption that ClpP1 and ClpP2 are distinct enzymes, while in fact, the active enzyme *in vivo* is a mixed complex.

Here, we demonstrate that ClpP1 and ClpP2, when over-produced independently, form tetradecameric complexes that lack any proteolytic activity. However, when these complexes are mixed together in the presence of certain small activating molecules, which we accidentally discovered, these tetradecamers dissociate into heptameric rings, which then re-associate into a mixed tetradecameric complex that is capable of degrading model peptides as well as some unstructured proteins. These low molecular weight activators clearly represent a novel form of enzyme regulation and stimulate ClpP1P2 activity in a very different manner from the regulatory ATPase complex, ClpC1, which we show enhances specifically the degradation of proteins. Thus, ClpP1P2 differs markedly from other members of the ClpP family and has a number of highly unusual structural, enzymatic, and regulatory properties. These unique qualities of ClpP1P2, taken together with its essential role during infection, make it an attractive target for drug development.

Results

Isolation of processed ClpP1 and ClpP2

Mtb *clpP1* or *clpP2* genes were expressed as C-terminal fusions with 6 × His and/or Myc tags under the control of a tetracycline-inducible promoter. Since previous efforts and our initial attempts to produce active ClpP1 and ClpP2 in *E. coli* were unsuccessful (Ingvarsson *et al*, 2007; Benaroudj *et al*, 2011), we attempted to separately express ClpP1 and ClpP2 under conditions resembling those in Mtb by using the closely related non-pathogenic species *M. smegmatis*. Purification on an Ni-NTA agarose column yielded large amounts of nearly pure proteins, each with an apparent molecular weight of ~22 kDa (Figure 1A). When ClpP1 and ClpP2 were subjected to gel filtration on an S-300 Sephacryl column, both were eluted as single homogenous peaks with a molecular mass of about 300 kDa (Figure 3A, top panel). Thus, both ClpP1 and ClpP2 had the same elution profile as *E. coli* ClpP and appeared to be 14-subunit 2-ring complexes.

The ClpP1 and ClpP2 bands from the SDS-PAGE were digested by trypsin and chymotrypsin and analysed by MS/MS. Eighty-three peptides were identified for ClpP1 (92% coverage by amino acids) and 70 peptides for ClpP2 (94% coverage). Although mass spectrometry thus demonstrated nearly all the expected peptides, N-terminal sequencing indicated that ~70% of both proteins were N-terminally processed with major cleavage sites at Asp⁶-Met⁷ for ClpP1 and Ala¹²-Arg¹³ for ClpP2 (Figure 1B). (In addition, minor cleavages were also detected at Thr⁵-Asp⁶ and Met⁷-Arg⁸ for ClpP1 and Arg¹³-Tyr¹⁴ for ClpP2.) It is noteworthy that the extent of this processing varied in different preparations and correlated with their ability to support enzymatic activity.

Thus, N-terminal processing of both gene products appears important for the formation of the active enzyme. Moreover, when full-length mutant forms of ClpP1 and ClpP2, which lacked enzymatic activity (see below), were expressed in *M. smegmatis*, a much smaller fraction of N-terminally processed forms could be detected. Therefore, it is likely that the proteolytic processing of mycobacterial ClpPs occurs primarily through an autocatalytic mechanism (possibly involving collaboration with the *M. smegmatis* enzymes). Accordingly, ClpP1 is cleaved after Asp (Figure 1B), which as shown below, is one of the preferred sites for Mtb ClpP (see below, Table I).

In subsequent studies, we therefore expressed the constructs corresponding to the processed versions directly and obtained more homogenous preparations with higher activities. It is noteworthy that these shorter forms, which do not require N-terminal processing, could also be efficiently produced in *E. coli*.

ClpP1 and ClpP2 form a mixed ClpP1P2 protease that requires certain short peptides for activation

Neither ClpP1 nor ClpP2 alone had peptidase activity (Figure 1C), although both formed tetradecameric structures characteristic of the ClpP family. Because both genes are essential (Sasseti *et al*, 2003; Ollinger *et al*, 2012; Raju *et al*, 2011), we hypothesized that ClpP1 and ClpP2 are not two distinct enzymes, but instead associate to form a novel, mixed proteolytic complex. To test this possibility, we first attempted to co-express Mtb ClpP1 and ClpP2 in *M. smegmatis*. The two proteins associated *in vivo* since they could be co-immunoprecipitated from the cell extract (Raju *et al*, 2011). However, due to wide variations in the levels of ClpP1 and ClpP2 expression, the ratios between the co-purified ClpP1 and ClpP2 varied markedly in different preparations, and this heterogeneity prevented rigorous study of the active complex. Therefore, we expressed them separately and attempted to reconstitute such a mixed complex from pure components. In fact, mixing pure ClpP1 and ClpP2 together in high concentrations (up to 0.5 mg/ml) resulted in the appearance of very low peptidase activity against the fluorogenic substrate of *E. coli* ClpP, Suc-Leu-Tyr-AMC.

During attempts to identify transition state-specific inhibitors of this low activity, we accidentally made the surprising, but very valuable discovery that a group of N-blocked peptide aldehydes that were substrate analogues not only did not inhibit, but actually stimulated this activity over 1000-fold. A similar dramatic activation was even found with certain related blocked peptides. For example, as shown in Figure 1C, a mixture of ClpP1 and ClpP2 was inactive in hydrolysing the Z-Gly-Gly-Leu-AMC or the quenched fluorescent substrate Mca-GHQYKMK-Dpa(Dnp)-amide, but in the presence of the activating peptide Z-Leu-Leu, both substrates and the unfolded protein, casein, were efficiently cleaved (Figure 1C and D). The activating peptides and peptide aldehydes only induced peptidase activity if both ClpP1 and ClpP2 were present together. It is noteworthy that at 37°C (under our standard assay conditions) the activation occurred without any noticeable delay after the addition of the peptide activator. Also, the activator had to be continually present for enzymatic activity. When the activator was removed by gel filtration or if its concentration was reduced by dilution, activity was lost, but it could be regained

Figure 1 Purification of processed but inactive ClpP1 and ClpP2 and reconstitution of the active ClpP1P2 complex. (A) Coomassie staining of ClpP1 (2.3 µg) and ClpP2 (2.7 µg) after purification. (B) Sequences of ClpP1 and ClpP2 proteins expressed in *M. smegmatis*. Arrows indicate the sites of proteolytic processing determined by Mass Spectrometry and N-terminal sequencing of ClpP1 and ClpP2 purified as in (A). (C) ClpP1P2 (2.1 µg) possesses peptidase activity but only in the presence of activating dipeptide Z-Leu-Leu. ClpP1 (1.8 µg) or ClpP2 (2.4 µg) alone did not show any activity with or without the activator. Enzymatic activity was measured fluorometrically using Z-Gly-Gly-Leu-AMC as a substrate. (D) Activator also stimulates degradation of longer peptides (Mca-KKPTIQLN-Dpa(Dnp)-amide) and proteins (FITC-casein) by ClpP1P2 (0.36 and 2.9 µg, respectively).

fully upon restoration of activator to its prior concentration (Figure 2C).

The strongest stimulation against Z-Gly-Gly-Leu-AMC, as well as other substrates, was found with Z-Leu-leucinal (Figure 2), but the longer aldehyde Z-Leu-Leu-leucinal was significantly less active. Several other hydrophobic dipeptide aldehydes (e.g., Z-Val-phenylalaninal), acidic peptide aldehydes (e.g., Z-Pro-Nle-aspartal) and alkyl aldehydes did not show any stimulatory capacity. The effective peptide aldehydes presumably should bind to at least some of the enzymes' 14 active sites. However, the related peptide Z-Leu-Leu and its alcohol derivative Z-Leu-leucinol (which presumably should not bind strongly to the active sites) could also activate ClpP1P2, although only at much higher concentrations than the corresponding aldehydes. A much smaller stimulation was observed with blocked peptides Z-Leu, Z-Gly-Leu and Z-Gly-Leu-Leu (Figure 2A).

The concentration dependence for activation, by Z-Leu-leucinal ($K_d = 0.24$ mM) and Z-Leu-Leu ($K_d = 2.2$ mM), revealed a highly cooperative mechanism with a Hill coefficient of 5–7 (Figure 2B). Thus, multiple molecules probably bind to ClpP1P2 to stimulate its activity. Though substrate analogues, these activators are not cleaved, since upon incubation with ClpP1P2, no new amino groups could be detected using the sensitive fluorescamine assay. It is noteworthy that although the aldehyde had a higher affinity, at high concentrations,

Z-Leu-Leu caused a greater activation than Z-Leu-leucinal (Figure 2B). Also because peptides are more stable and much less expensive than the corresponding aldehydes, in subsequent studies, we routinely induce Mtb ClpP1P2 activity using Z-Leu-Leu (subsequently referred to as the 'activator').

Activation involves dissociation of ClpP1 and ClpP2 tetradecamers and formation of 2-ring ClpP1P2 complex

Because the activators stimulate only ClpP1 and ClpP2 together (but not pure ClpP1 or ClpP2; Figure 1C and D), they probably activate by promoting the formation of a new mixed ClpP1P2 complex. We therefore examined how the presence of an activator affects the sizes of these different complexes. Upon size-exclusion chromatography, a mixture of ClpP1 and ClpP2 behaved as tetradecamers exactly like pure Mtb ClpP1 or ClpP2 and *E. coli* ClpP (Figure 3A, upper panel). However, when the activator Z-Leu-Leu was present (Figure 3A, lower panel), both ClpP1 and ClpP2 peaks were eluted as a single lower molecular weight peak, resembling γ -globulin (150 kDa) in size. Thus, the tetradameric (presumably 2-ring) complexes composed of a single subunit type dissociated into heptamers. However, in the presence of the activator, the ClpP1/ClpP2 mixture was eluted as a 300-kDa peak that coincided with the peptidase activity and corresponded in size to ClpP tetradecamers (Figure 3A, lower panel). The ClpP1P2 complexes were isolated from the peak using Ni-NTA (by His-tagged ClpP2) or anti-Myc (by Myc-tagged

Table I Mtb ClpP1P2 preferentially hydrolyses peptides with hydrophobic and acidic residues in P1 position

Peptide substrate	Relative rates of hydrolysis (%)
<i>Hydrophobic P1 residue</i>	
Z-Gly-Gly-Leu-amc	100.0
Suc-Leu-Leu-Val-Tyr-amc	0.8
Suc-Leu-Tyr-amc	11.5
Z-Leu-Leu-Leu-amc	0.15
Z-Leu-Leu-amc	4.7
Z-Ala-Ala-Ala-amc	3.6
Suc-Ala-Leu-Pro-Phe-amc	0.12
Suc-Ala-Ala-Pro-Ala-amc	0.08
Ala-Ala-Phe-amc	87.0
Suc-Ala-Ala-Phe-amc	42.0
<i>Acidic P1 residue</i>	
Ac-Nle-Pro-Nle-Asp-amc	36.0
Z-Leu-Leu-Glu-amc	0.35
<i>Basic P1 residue</i>	
Z-Leu-Leu-Arg-amc	0.25
Z-Gly-Gly-Arg-amc	0.55
Z-Phe-Val-Arg-amc	0.18
<i>Aminopeptidase substrates</i>	
Ala-amc	0.78
Leu-amc	0.36
Phe-amc	0.26
Asp-amc	0.1

Peptidase activity of ClpP1P2 (1.8 µg) was measured with a variety of substrates (0.1 mM) and compared with activity obtained with Z-Gly-Gly-Leu-AMC (100%).

ClpP1) columns, and the presence of both proteins in resin-bound material was confirmed by MS (see Materials and methods).

Thus, the activating peptide causes the dissociation of ClpP1 and ClpP2 tetradecamers into heptamers and favours their subsequent association to form the active tetradameric ClpP1P2 complex. By contrast, no changes in elution pattern were observed when *E. coli* ClpP was incubated with this activator.

Conformational changes accompanying formation of ClpP1P2 complex

The dissociation and re-association of multimeric ClpP1 and ClpP2 rings must involve activation-induced major changes in subunit conformation. Because ClpP1 (but not ClpP2) contains a Trp residue, we can use it to monitor conformational changes that may accompany the formation of an active ClpP1P2 complex from inactive ClpP1 and ClpP2 ones. Although no spectral changes were observed with dissociation of the ClpP1 tetradecamer upon addition of the activator, the formation of the active ClpP1P2 complex appears to involve changes in ClpP1's conformation, because the fluorescence of Trp174 in ClpP1 shifted its maximal fluorescence from 345 in pure ClpP1 to 338 nm. (Figure 3B). Thus, the interaction between ClpP1 and ClpP2 subunits leading to activation is associated with changes in the subunits' conformation. It is noteworthy that similar changes in Trp174 fluorescence occurred when active-site mutants of ClpP1 and ClpP2 that lack enzymatic activity (see below) were mixed in the presence of the activator. Thus, enzymatic activity of both ClpPs is not necessary for their dissociation—re-association and the major structural changes associated with this activation process.

To confirm that such a mixed ClpP1P2 complex actually exists *in vivo*, we tested whether endogenous ClpP1 and ClpP2 associate in wild-type *M. smegmatis*. As described in the related manuscript (Raju *et al*, 2011), we employed mycobacterial recombineering to add a C-myc tag to the C-terminus of genomic ClpP2. The C-myc-tagged ClpP2 was isolated together with associated proteins using an anti-myc resin, and the material eluted with the Myc peptide was resolved by SDS-PAGE. Bands corresponding by size to ClpP2 and ClpP1 were analysed by Mass Spectrometry, and the presence of both subunits was confirmed, indicating that mixed ClpP1P2 complexes are present in mycobacteria.

Mtb ClpP1P2 is composed of one ClpP1 and one ClpP2 heptameric ring

To determine the subunit composition of this ClpP1P2 complex, we varied the relative concentrations of ClpP1 and ClpP2 in the presence of an activator (Figure 4A). Upon increasing the amount of ClpP1 with a constant amount of ClpP2, peptidase activity gradually increased and reached its maximum when these components were present in close to equimolar amounts. Conversely, when ClpP1 content was held constant and the amount of ClpP2 increased, maximal activity was also obtained with equimolar concentrations. In different experiments using different ClpP1 and ClpP2 preparations, the optimal ClpP1/ClpP2 molar ratio ranged from 0.82 to 1.15. Thus, the active complex contains equal numbers of ClpP1 and ClpP2 subunits.

These findings and the rapidity of activation together strongly suggest that the active enzyme is composed of one ClpP1 and one ClpP2 ring. However, it is also possible that each heptameric ring contains a mixture of ClpP1 and ClpP2 subunits, as has been found for the cyanobacterium *Synechococcus* ClpP complexes (Stanne *et al*, 2007; Andersson *et al*, 2009). To determine the composition of the rings, we crosslinked the neighbouring subunits in the active ClpP1P2 tetradecamer with glutaraldehyde in the presence of the activator (Figure 4B). After a 0.5 h of incubation, seven distinct crosslinked bands were evident on SDS-PAGE corresponding to 1-, 2-, 3- 4-, 5-, 6-, and 7-mers. As expected, the larger crosslinked structures were the least abundant. After an overnight incubation, when crosslinking went to completion, all seven subunits, presumably comprising the rings, were crosslinked together, but still no band was observed with a molecular mass higher than that of a 7-mer. Thus, apparently, no crosslinking occurred between the two rings (which presumably requires a crosslinker with a longer spacer arm than glutaraldehyde). Analysis by mass spectrometry indicated that the crosslinked heptamers were composed only of ClpP1 or ClpP2 subunits, respectively, and no peptides corresponding to ClpP1–ClpP2 crosslinked were found. Thus, each heptameric ring in the Mtb ClpP1P2 protease is homogenous in composition.

Cleavage specificity of Mtb ClpP1P2

To define the substrate preference of the ClpP1P2 active sites, we tested a variety of synthetic fluorescent peptides with hydrophobic, acidic, or basic residues in the P1 position (Table I). The best substrate was Z-Gly-Gly-Leu-AMC, while Suc-Ala-Ala-Phe-AMC and Ala-Ala-Phe-AMC also were readily cleaved. The failure of ClpP1P2 to degrade rapidly the widely used proteasome substrate Suc-Leu-Val-Tyr-AMC

Figure 2 Certain short peptides and peptide aldehydes dramatically activate ClpP1P2 by binding to multiple sites. (A) Effects of various short peptide aldehydes and related peptides or peptide alcohols on activity of ClpP1P2 (1.3 μ g). Peptidase activity was measured with Z-Gly-Gly-Leu-AMC. The specific activity with Z-Leu-norleucinal was 4.25 μ mole/mg/min, which was taken as 100%. (B) Determination of Hill coefficient of Z-Leu-Leu and Z-Leu-leucinal in the hydrolysis of Z-Gly-Gly-Leu-AMC (0.1 mM) and Ac-Nle-Pro-Nle-Asp-amc (0.1 mM) by ClpP1P2 (2.4 μ g). (C) Activation of ClpP1P2 by Z-Leu-Leu is readily reversible. Enzyme (1.8 μ g) was incubated in the presence of activator Z-Leu-Leu (5 mM) and then diluted 200-fold in the buffer with and without the activator. Enzymatic activity was assayed with Z-Gly-Leu-Leu-AMC (0.1 mM). Re-addition of the activator restored the ClpP1P2 activity completely.

indicates major differences from enzymes in the mammalian cytosol. It is noteworthy that Z-Leu-Leu-AMC, the fluorescent peptide corresponding to the peptide activator employed routinely Z-Leu-Leu, was a poor substrate for the enzyme (Table I), and conversely the peptides corresponding to the best substrates, Z-Gly-Gly-Leu or Z-Gly-Leu, were poor as activators (Figure 2A).

In addition to hydrophobic peptides, ClpP1P2 also efficiently hydrolyses a peptide with acidic residues in the P1 position, Ac-Nle-Pro-Nle-Asp-AMC. (We also found that this substrate is degraded by *E. coli* ClpP, which had been reported to cleave after aspartate residues in model polypeptides; Thompson and Maurizi, 1994). However, Mtb ClpP1P2 did not hydrolyse Z-Leu-Leu-Glu-AMC or peptides with basic P1 residue and was also inactive against a variety of unblocked amino acid-AMC substrates used to assay aminopeptidases (Table I). ClpP1P2 also could cleave a variety of longer quenched fluorescent peptides (e.g., Mca-GNTQFKRR-Dpa(Dnp)-amide, Mca-GHQYAMK-Dpa(Dnp)-amide, Mca-GNQYKMK-Dpa(Dnp)-amide and Mca-KKPTIQLN-Dpa(Dnp)-amide), and could degrade slowly the largely unstructured protein FITC-casein, provided an activator was present (Figure 1D).

Though ClpP1 or ClpP2 alone lack enzymatic activity, their catalytic triads are formed

The sequences of both ClpP1 and ClpP2 appear to contain a Ser/His/Asp catalytic triad characteristic of serine proteases

(Figure 1B). Accordingly, Mtb ClpP1P2 was sensitive to most standard inhibitors of serine proteases (Figure 5A), including agents that react with the active-site serine (dichloroisocoumarin, Powers and Kam, 1994 and biotinylated derivative of fluoroethoxyphosphinyl (FP-biotin), Liu *et al*, 1999), and peptide chloromethyl ketones (Szyk and Maurizi, 2006), which modify the catalytic histidine. By contrast, standard inhibitors of metalloproteases and cysteine proteases had no effect. Interestingly, the hydrolysis of both hydrophobic and acidic substrates was inhibited to similar extents by the peptide chloromethyl ketones, Z-LY-CMK or AAF-CMK.

To learn whether both gene products are enzymatically active in the complex, we incubated ClpP1, ClpP2, and the ClpP1/ClpP2 mixture with the biotinylated covalent modifier of active-site serines. As shown in Figure 5B, both ClpP1 and ClpP2 subunits were covalently modified even in the enzymatically inactive ClpP1 or ClpP2 complexes. Thus, the catalytic triad appears functional in these complexes despite their lack of enzymatic activity. To test the possibility of a non-specific binding of biotinylated modifier, the mutant forms of ClpP1 or ClpP2 with active-site Ser substituted for Ala were incubated with FP-biotin. As shown in Figure 5B, no incorporation of the modifier in the mutant proteins occurred, thus confirming its specific reaction with the active sites.

We anticipated that the activator would enhance the modification of the active-site serines by promoting active-site formation. Surprisingly, the presence of the activator did

Figure 3 With the activator present, ClpP1 and ClpP2 tetradecamers dissociate into heptamers but ClpP1/ClpP2 mixture forms an enzymatically active tetradecamer. (A) Size-exclusion chromatography of ClpP1 (0.28 mg), ClpP2 (0.34 mg), and ClpP1/ClpP2 mixture (0.31 mg) was carried out using Sephacryl S300 column in the absence (upper panel) and presence (lower panel) of activator Z-Leu-Leu, which was also present in the running buffer. Samples (0.2 ml) were applied to the column and 0.5 ml fractions were collected and assayed for protein content and activity with Z-Gly-Gly-Leu-AMC. The column was calibrated with γ -thyroglobulin (670 K), γ -globulin (158 K), ovalbumin (44 K), and *E. coli* ClpP (300 K). (B) The change in fluorescence emission spectrum of ClpP1/ClpP2 mixture (12.8 μ g in 100 μ l) upon addition of activator indicates a conformational change due to complex formation between ClpP1 and ClpP2. The addition of activator to ClpP1/ClpP2 mixture shifted the peak of ClpP1 Trp174 emission from 345 to 338 nm (lower panel), while no change was observed with ClpP1 alone (upper panel).

not stimulate the modification of either homogeneous ClpP1 or ClpP2 or the mixed complex. In fact, the activator even reduced slightly (but reproducibly) this reaction, as also occurred in the presence of a substrate. Similar inhibition was observed in the presence of a substrate. Thus some activator molecules, which are structurally related to peptide substrates, appear to bind to the active sites. In any case, because active-site labelling occurred with ClpP1 and ClpP2 alone, these results prove that activation is not through formation of the catalytic triad, and instead that formation of the mixed tetradecamer probably enables the substrate to access and bind to the previously latent active sites.

ClpP1 and ClpP2 have distinct cleavage specificities

To estimate how ClpP1 and ClpP2 subunits contribute to enzymatic activity of the complex, pure ClpP1 or ClpP2 was inactivated by pretreatment with dichloroisocoumarin at concentrations that completely inhibit ClpP1P2. When the covalently inactivated ClpP1 or ClpP2 was incubated with its normal counterpart plus activator, the hydrolysis of hydrophobic and acidic peptide substrates, as well as casein, was significantly less than with untreated subunits (Figure 5C). Thus, both types of subunits appear to function enzymatically and contribute to the activity of the complex. However,

inactivation of ClpP1 caused a much greater loss of these activities than did inactivation of ClpP2, especially with the hydrophobic peptide substrate (Figure 5C).

To confirm these different roles of each subunit, we reconstituted enzymatic complexes using wtClpP1 and an active-site mutant ClpP2 (Ser¹¹⁰ to Ala) or with wtClpP2 and active-site mutant ClpP1 (Ser⁹⁸ to Ala). The complexes containing only one type of active subunits showed lower peptidase activity than the wild-type enzyme against both hydrophobic and acidic peptide substrates, and casein (Figure 5C). As was found upon derivatization with isocoumarin, the lack of functional ClpP1 caused a greater loss of activity against these various substrates than did the loss of ClpP2, particularly with the hydrophobic substrate. These data indicate that ClpP1 and ClpP2 active sites have different substrate preferences and suggest that ClpP1 sites are more important than ClpP2's in cleaving the most abundant bonds in proteins.

The ClpC1 ATPase complex stimulates protein degradation by ClpP1P2 but only in the presence of both an activator and ATP

The mechanism for the dramatic ClpP1P2 activation by small peptides uncovered here was surprising and unprecedented.

Figure 4 ClpP1P2 has maximal activity when equimolar amounts of ClpP1 and ClpP2 are present (**A**) and is composed of heptameric rings containing only ClpP1 or ClpP2 subunits (**B**). (**A**) Activity of ClpP1P2 at different ClpP1:ClpP2 ratios. Constant amounts of ClpP2 (0.85 μ g) were mixed with increasing amounts of ClpP1, and Z-Gly-Gly-Leu-AMC hydrolysis was measured in the presence of activator. (**B**) Crosslinking of ClpP1P2 subunits by glutaraldehyde. After 0.5 and 20 h incubation at room temperature of ClpP1P2 (12 μ g) with 0.125% glutaraldehyde, the reaction mixture was analysed by SDS-PAGE, followed by Mass Spectrometry. High molecular weight bands corresponding to seven crosslinked subunits were found to contain exclusively ClpP1 or ClpP2 peptides, indicating that each ring contains seven identical ClpP1 or ClpP2 subunits. In the insert, one tenth of the crosslinked material used in Figure 4B was resolved by SDS-PAGE and stained by Coomassie Blue. The gel was scanned, and the image enlarged using Photoshop.

Figure 5 Both ClpP1 and ClpP2 contain functional active sites, but ClpP1's active sites are more important than ClpP2's in ClpP1P2 activity. (**A**) ClpP1P2 protease has the inhibitor sensitivity characteristic of serine proteases. Peptidase activity against Z-Gly-Gly-Leu-AMC (0.1 mM) was measured after 30 min preincubation of the enzyme (1.2 μ g) with or without inhibitors at room temperature. The error ranges of the inhibition did not exceed 5%. (**B**) To learn whether the active-site charge relay system is functional in the absence of activator, 1.2 μ g of wild-type or active-site Ser-Ala mutants of ClpP1, ClpP2, or a ClpP1/ClpP2 mixtures were incubated for 30 min with a biotinylated derivative of an active-site inhibitor fluoroethoxyphosphinyl. The binding of the inhibitor was determined by SDS-PAGE followed by western blot analysis with an anti-biotin antibody. (**C**) To determine how ClpP1 and ClpP2 contribute to the enzymatic activity of ClpP1P2, either ClpP1 or ClpP2 was inactivated separately by pretreatment with dichloroisocoumarin (0.1 mM) or by an active-site mutation (active site Ser to Ala). The inactivated ClpP1 or ClpP2 was then mixed with its normal (untreated WT) counterpart in the presence of activator Z-Leu-Leu, and hydrolysis of hydrophobic and acidic peptide substrates and casein was measured. FITC-casein was not used to measure proteinase activity of DCL-treated enzymes due to the overlapping of the fluorescence spectra of FITC and DCL. Figure source data can be found in Supplementary data.

One attractive possible mechanism would be that the conformational changes and complex formation induced by the activator resemble those changes caused by the binding of the Mtb regulatory ATPase complex, ClpC1 or ClpX, since in some ATP-dependent proteases (e.g., human mitochondrial ClpP), no peptidase activity was demonstrated in the absence of nucleotides and the regulatory ATPase. To address the possibility that the peptide activator mimicked the regulatory ATPases, we cloned the ClpC1 ATPase from Mtb, expressed it in *M. smegmatis*, and tested whether it can stimulate ClpP1P2 activity. As expected, the resulting complex was of high molecular weight and had ATPase activity. (The isolation and characterization of Mtb ClpC1 will be reported in detail in a separate manuscript.) Unlike the dipeptide activator, addition of ClpC1 did not increase the hydrolysis of any of various fluorescent peptides assayed, nor did it enhance the effect of the activator on their degradation (Figure 6A). However, when the ATPase was added in the presence of the activator, it markedly increased the degradation of the model protein, FITC-casein (Figure 6A). This stimulation of protein degradation by ClpP1P2 was only observed when both the activator and ATP were present (Figure 6C). Thus, ClpC1 and the activator must increase proteolysis by quite different, but clearly additive mechanisms.

It is noteworthy that ClpC1 did not stimulate proteolysis by ClpP1 or ClpP2 alone (Figure 6B). Therefore, it is very likely that ClpC1 *in vivo* also functions only with the mixed ClpP1P2 complex and requires an additional factor resembling the activator for protein degradation. The peptide activator did not influence ATP hydrolysis by ClpC1. By contrast, the protein substrate casein stimulated ClpC1's ATPase activity two-fold (Figure 6D) in a similar fashion to the activation by substrates of the homologous *E. coli* AAA ATPases, Lon (Waxman and Goldberg, 1982), ClpA (Hwang

et al, 1988; Thompson and Maurizi, 1994), and HslU (Seol *et al*, 1997). Thus, the peptide activator is necessary only for ClpP1P2 assembly, while ClpC1 binds casein directly and facilitates its degradation by ClpP1P2.

Discussion

Mtb ClpP1P2 is a novel enzyme complex in multiple respects

Several observations led us to hypothesize that ClpP1 and ClpP2 function together in a single complex. (1) Our initial attempts and those of others (Ingvarsson *et al*, 2007; Benaroudj *et al*, 2011) to isolate active ClpP1 or ClpP2 from Mtb were unsuccessful, even though they formed tetradecameric complexes similar in size to other ClpP family members, as shown in Figure 3 and by others (Ingvarsson *et al*, 2007). (2) In our related paper (Raju *et al*, 2011), we showed that both *clpP1* and *clpP2* genes are essential for growth and infectivity of Mtb, and thus, they cannot compensate for the loss of the other. (3) When ClpP1 and ClpP2 were co-expressed in *M. smegmatis*, they could be co-immunoprecipitated (Raju *et al*, 2011). We then showed that neither purified ClpP1 nor ClpP2 alone is enzymatically active, but if they were both present, together with an activating peptide or peptide derivative, then mixed complexes were formed that showed robust proteolytic activity. Furthermore, only when the mixed ClpP1P2 complex was formed in the presence of the activator, casein degradation was stimulated in an ATP-dependent manner by ClpC1 ATPase. Since this latter process resembles the conditions for protein degradation *in vivo* and since ClpC1 is also essential for viability, it seems very likely that the ClpP1P2 is the functional protease *in vivo*.

One initial approach that enabled us to obtain active ClpP1P2 was the use of an unusual expression system.

Figure 6 ClpC1 activates degradation of casein, but not peptides, by ClpP1P2 only in the presence of activator and ATP. (A) ClpC1 activates casein but not peptide degradation by ClpP1P2. ClpP1P2 (2.5 μg) and ClpC1 (32 μg) were mixed in 100 μl of reaction buffer containing 2 mM ATP and 8 mM MgCl₂ (see Materials and methods) in the presence or absence of the peptide activator. Enzymatic activity was measured fluorometrically using Z-Gly-Gly-Leu-AMC or FITC-casein as substrates. The rate of degradation of ClpP1P2 in the presence of the activator was taken as 100%. (B) ClpC1 activates casein degradation by ClpP1P2 but not by ClpP1 or ClpP2 alone. Degradation of FITC-casein by ClpP1 (2.5 μg), ClpP2 (2.9 μg), or ClpP1P2 (2.7 μg) was measured as in (A). (C) The stimulation of FITC-casein degradation by ClpP1P2 requires ATP and the peptide activator. Degradation was measured as in (A) and (B) with different concentrations of ClpC1. (D) Casein stimulates ATPase activity of ClpC1. In all, 2 μg of ClpC1 was incubated with 5 μg FITC-casein, or with ClpP1P2 +/− activator, or with the activator alone, in 20 μl of reaction buffer for 30 min and the ATPase activity was measured by Malachite Green method (see Materials and methods).

Expressing the cloned Mtb ClpP1 and ClpP2 proteins in *E. coli* did not yield the mature, processed subunits, although some limited processing of Mtb ClpP1 and ClpP2 (by ClpP1) has been recently reported in *E. coli* (Benaroudj *et al*, 2011). In the present studies, we used the closely related mycobacterium, *M. smegmatis*, in which ClpP1 and ClpP2 underwent efficient N-terminal processing, which appears necessary for enzymatic activity. In fact, in the crystal structure of the inactive ClpP1 tetradecamer composed of unprocessed subunits, the distance between the active-site Ser and His residue was too large to support the formation of the active catalytic triad (Ingvarsson *et al*, 2007). Once we had identified the N-termini of the fully active ClpP1 and ClpP2, we directly expressed these shorter sequences, which yielded homogenous tetradecamers in *M. smegmatis* as well as in *E. coli*. Also, by using an *E. coli* mutant lacking endogenous ClpP, we could obtain WT or mutant ClpP1 and ClpP2 without contamination by endogenous ClpP subunits. It is noteworthy that these ClpP1 tetradecamers composed only of 'processed' subunits do contain a functional catalytic triad, although it is still unable to catalyze peptide hydrolysis. Such a lack of catalytic activity despite the presence of a functional catalytic triad has been previously shown for other pro-enzymes, such as trypsinogen, which though enzymatically inactive can react with various active-site titrants (Smith *et al*, 1992). Its lack of enzymatic activity is attributed to an inability of substrates to access the active sites, and presumably a similar explanation accounts for the inactivity of pure ClpP1 and ClpP2 (see below).

Recently, Mtb ClpP2 and ClpC1 were reported to catalyze the degradation of an endogenous Mtb protein Rse A (Barik *et al*, 2010). While degradation of that substrate perhaps may not require formation of a ClpP1P2 complex (unlike casein or the peptides studied here), several features of that study are difficult to reconcile with our findings. For example, they expressed Mtb ClpP proteins in *E. coli*, which we and others (Ingvarsson *et al*, 2007; Benaroudj *et al*, 2011) found to yield mostly non-processed inactive tetradecamers. Possibly, their use of WT *E. coli* for expression may have resulted in contamination by the highly active *E. coli* ClpP, as was reported by Benaroudj *et al* (2011). To avoid these problems, we expressed the mature Mtb proteins only in *E. coli* strain lacking ClpP.

Very little is known about ClpP proteases from organisms that contain multiple *clpP* genes. In *Streptomyces lividans*, five *clpP* genes are organized in two operons (*clpP1* and *clpP2*; *clpP3* and *clpP4*), and one is monocistronic (Viala and Mazodier, 2002). Both ClpP1 and ClpP2 are required for degradation of the transcriptional activator PopR, which suggests that they also form a single mixed complex (Viala and Mazodier, 2002). ClpP3 and ClpP4 can also function together in PopR degradation, and their coordinate regulation suggests that they also comprise a mixed complex (Viala and Mazodier, 2002). In plant organelles, the organization of the ClpP proteases is much more complex (Peltier *et al*, 2001, 2004); for example, tetradecameric ClpP complexes have been isolated from *Arabidopsis thaliana* that contain five different ClpP proteins and six different non-proteolytic ClpP homologues (ClpR) (Peltier *et al*, 2004). Although their composition and activities have not been studied, the different ClpP and ClpR proteins may be present in the same tetradecameric complex (Peltier *et al*, 2004). In fact, a novel

form of ClpP has recently been characterized from the cyanobacterium *Synechococcus* that contains two identical heptameric rings, composed of three active ClpP3 and four inactive ClpR subunits (Andersson *et al*, 2009), which though inactive, are essential for the ClpC-dependent proteolytic activity.

Our crosslinking experiments demonstrate that Mtb ClpP1P2 heptameric rings are composed of seven identical subunits. Therefore, the active Mtb enzyme must be composed of one ClpP1 ring and one ClpP2 ring. Accordingly, optimal activity was obtained when ClpP1 and ClpP2 were present in a 1:1 molar ratio (Figure 4A). This association of the ClpP1 and ClpP2 rings with each other causes conformational changes that allow both complementary rings to become enzymatically active. By monitoring the changes in tryptophan fluorescence in ClpP1 that accompany activation, we confirmed that the active ClpP1 conformation is achieved only in the presence of ClpP2 and a dipeptide activator. Interestingly, the ClpP1 and ClpP2 rings can activate each other, even if either or both were inactivated by mutation or derivatization of their active-site serines.

Although only a limited number of peptide substrates were screened, Mtb ClpP1P2 clearly has a rather broad substrate specificity. In its preference for large hydrophobic residues in the P1 position, ClpP1P2 resembles chymotrypsin; however, it also cleaves a peptide with an acidic residue in the P1 position that is a typical substrate for caspases and the caspase-like site on the proteasome (Kisselev *et al*, 2003, 2006). The active sites on ClpP1 and ClpP2 clearly differ in their cleavage specificity. ClpP1 clearly is predominant in the hydrolysis of casein and after hydrophobic residues (which are highly abundant in cell proteins). Its loss also reduces the rate of hydrolysis of acidic peptides but ClpP2 rings also contribute significantly to this activity. Further analysis of the hydrolysis of other substrates should be of value in the development of selective inhibitors.

Activation of ClpP1P2 by small molecules

The most unexpected and novel aspect of these findings is the discovery of small molecules that dramatically activate ClpP1P2 and enable it to degrade even unstructured polypeptides. *In vitro*, these agents were essential for both the appearance and the maintenance of enzyme activity. The most potent among these activators are short N-blocked peptide aldehydes, but the corresponding peptide alcohols and peptides also stimulate, though at higher concentrations. (Z-Leu-Leu was used routinely here because it is inexpensive and yields the largest maximal activation.) All these compounds markedly stimulate hydrolysis of all peptide substrates tested, as well as casein. In fact, without this surprising finding, the remaining observations on ClpP1P2 would not have been possible because its inherent activity is too low for most studies.

Even though ClpP1 and ClpP2 were enzymatically inactive by themselves, their active sites, even in the absence of the activator, could react with agents that covalently modify active-site serines or histidines, and did so as strongly as in the active enzyme (Figure 5B). Thus, in the absence of the activator, the catalytic triads appear to be functional in ClpP1 and ClpP2, unlike in the unprocessed ClpP1 (Ingvarsson *et al*, 2007). Presumably, these sites are unable to hydrolyze peptides in the absence of the complementary ring, because of a

failure of the substrate to enter the pure ClpP1 and ClpP2 complexes, as occurs with the latent form of the 20S proteasome, which is activated by a gating mechanism allowing substrate entry (Smith *et al*, 2007). Alternatively, formation of the mixed complex may involve structural rearrangements that enable catalysis.

The exact site where the activator binds to ClpP1P2 remains uncertain, and several possibilities exist. The structures of the activators closely resemble those of some hydrophobic substrates (Figure 2A; Table I), which suggest that the activators bind to the active sites. Accordingly, peptide aldehydes, which should bind tightly to active-site serine, were the strongest activators, while related peptides (which resemble products of substrate cleavage) are about 10 times less potent. Another observation supporting binding to the active sites was that the addition of activators instead of increasing enzyme interaction with the active-site titrant, actually decreased the extent of this modification. Because the structural changes that accompany ClpP1P2 activation (dissociation of tetradecamers into heptamers, formation of mixed complexes and changes in Trp fluorescence) were also induced by the activator in the inactive ClpP1 and ClpP2 active-site mutants, these activating dipeptides do not require interaction with the catalytic serines to induce dissociation—re-association.

Although agents that bind to the active sites should be competitive inhibitors if they bind to all the active sites, in the tetradameric HslV protease complex from *E. coli*, Chung and coworkers have elegantly shown that inactivation of about half the proteolytic sites can occur without a decrease in maximal proteolytic rate (Lee *et al*, 2009). Thus, in Mtb ClpP1P2 partial occupancy of active sites by activators probably could occur without reducing activity, while also perhaps inducing conformational changes in remaining subunits that favour formation of the active state. To induce these structural changes, the activators exhibit very strong cooperativity with a Hill coefficient between 5 and 7, which suggests that multiple molecules bind to either a fraction of the active sites (or to a distinct allosteric site) to induce the active conformation. Unfortunately, the activators identified thus far have low affinities; consequently, identifying their specific binding sites is difficult by standard biochemical approaches.

Although these activators resemble peptide substrates, there does not appear to be a simple correspondence between sequences that are preferentially hydrolysed and ones that support activation. Peptides corresponding to the peptide activators were poor substrates for the enzyme (Table I), and conversely the peptide corresponding to the best substrate was poor as an activator. Thus, it is possible that these peptides activate by also binding to an additional regulatory site.

The active sites of the cylindrical proteases, such as ClpPs, HslUV, or proteasomes, are sequestered within the proteolytic chamber and by themselves cannot degrade protein substrates (Baumeister *et al*, 1998; Yu and Houry, 2007; Striebel *et al*, 2009). Activation of these compartmentalized proteases can occur if the binding of the ATPase alters the conformation of the active site as in HslUV system (Yoo *et al*, 1996; Huang and Goldberg, 1997; Sousa *et al*, 2002; Yu and Houry, 2007) or opens an entry channel to allow substrate access, as occurs in gating of proteasome (Groll *et al*, 2000;

Whitby *et al*, 2000; Smith *et al*, 2007; Rabl *et al*, 2008) and proteases ClpXP or ClpAP (Grimaud *et al*, 1998; Maurizi *et al*, 1998; Kirstein *et al*, 2009; Lee *et al*, 2010b). The latter mechanism is important in action of acyldepsipeptide antibiotics, which are cytotoxic in *B. subtilis* and *E. coli* (Brotz-Oesterhelt *et al*, 2005; Kirstein *et al*, 2009) by causing activation of ClpP and excessive degradation of cellular proteins (Kirstein *et al*, 2009; Lee *et al*, 2010a). These molecules bind to the two ends of the ClpP tetradecamer in the cavities between adjacent ClpP monomers, which are the same sites to which loops of the regulatory ATPases bind (Kim *et al*, 2001; Bewley *et al*, 2006; Lee *et al*, 2009, 2010a). Thus, the acyldepsipeptides stimulate proteolysis by facilitating substrate access to the degradative chamber (Kirstein *et al*, 2009; Lee *et al*, 2010a), and prevent association of the protease with regulatory ATPases (Kirstein *et al*, 2009; Lee *et al*, 2010a). Recently, additional activators have been identified that function by a similar mechanism (Leung *et al*, 2011).

One initially attractive possible mechanism of our dipeptide activators was that they function in a similar fashion as acyldepsipeptide antibiotics. However, the dipeptide activator and ClpC1 ATPase have very different effects on peptide degradation, and stimulation of casein degradation requires the presence of both an activator and ClpC1. Furthermore, in related studies using acyldepsipeptide ADEP2, we recently found that these agents also can activate Mtb ClpP1P2 only in the presence of an activating peptide. It is therefore highly unlikely that the activator binds to the same regulatory cavities as ClpC1 or the ADEPs.

The physiological relevance of this activation mechanism for ClpP1P2 is an important issue for future study. The finding that ClpC1 functions only in the presence of a small molecule activator suggests that factors with similar activity exist in Mtb and allow ClpP1P2 association and function with ClpC1. In this respect, these activating dipeptides resemble a protein or 'chemical' chaperone that prevents formation of inactive conformations and favours formation of the active enzyme. However, unlike a chaperone, the dipeptide activators have to be continuously present to maintain the active ClpP1P2 complex. These findings imply that Mtb contains endogenous activators, either small molecules or perhaps protein(s), that promote the formation and the maintenance of ClpP1P2 mixed tetradecamers *in vivo*, and the identification of these activators and the isolation of the active complexes from Mtb will be important goals for future work.

In human mitochondrial ClpP (unlike *E. coli* ClpP), the ClpX ATPase complex is necessary not only for substrate recognition and translocation, but also for the formation of the ClpP tetradecamers from inactive heptamers (Kang *et al*, 2005). This action resembles the second stage in the activation mechanism demonstrated here (Figure 7). However, ClpC1 was unable by itself to induce activation of the Mtb enzyme. In Mtb, the expression of ClpC1 is regulated coordinately with ClpP1 and ClpP2 by ClgR factor (Sherid *et al*, 2010); therefore, it is very likely that ClpP1P2 functions *in vivo* together with ClpC1. Consequently, our finding that ClpC1 ATPase promotes the degradation of casein only in the presence of an activator argues strongly that the activator serves a unique, essential function, and that in Mtb some endogenous factor serves a similar role in promoting assembly of the mixed complex.

Figure 7 Mechanism of formation of active ClpP1P2. The proposed mechanism is based on the results in (1) Figure 4A showing that the optimal activity of ClpP1P2 is reached at the equal molar amounts of ClpP1 and ClpP2; (2) Figure 3A showing the dissociation of ClpP1 and ClpP2 tetradecamers into heptamers, and their re-association into mixed ClpP1P2 complex in the presence of the activator; (3) Figure 3B showing physical interaction between the rings; and (4) Figure 4B demonstrating crosslinking of only one type of subunits within the rings.

ClpP1P2 is an attractive drug target

The present findings and our related *in-vivo* observations (Raju *et al*, 2011) provide strong evidence that inhibition of ClpP1P2 is a promising new approach to combat tuberculosis. Not only is ClpP1P2 essential in Mtb, but no similar proteolytic complex exists in the mammalian cytosol, and its cleavage specificity, as defined with model peptides, clearly differs from those of the major mammalian cytosolic proteases (proteasomes). Furthermore, Mtb ClpP1P2 in structure and substrate specificity differs considerably from known ClpP family members in the mitochondria. Also, despite its unusual activation mechanism, ClpP1P2 is sensitive to typical inhibitors of serine proteases. It also seems likely that agents that activate Mtb ClpP1P2 may have therapeutic applications, since acyldepsipeptide antibiotics are toxic in many bacteria by activating ClpP and causing excessive proteolysis (Kirstein *et al*, 2009). Finally, the identification of inhibitors of ClpP1P2 activity or blockers of its activation should now be a straightforward undertaking, since screening can be carried out without the need for ATPases that regulate its activity *in vivo*. Nevertheless, characterization of these essential regulatory ATPases will be an important step to understand the physiological functions of ClpP1P2, and they may also represent therapeutic targets. In fact, Mtb ClpC1 has been recently identified as a drug target (Schmitt *et al*, 2011).

Materials and methods

Materials

Fluorogenic peptide substrates with C-terminal aminomethylcoumarin (amc) groups and protease inhibitor were from Bachem (Switzerland), Sigma (USA), or Biomol International (USA). Peptide substrates for the FRET assay, FITC-casein, and glutaraldehyde were from Sigma. Ni-NTA agarose was from Qiagen and Sephacryl S-300 from Pharmacia. Black 96-well microplates used in the enzyme assays were from Greiner (Germany). FP-Biotin, 10-(fluoroethoxyphosphoryl)-N-(biotinamidopentyl)decanamide, was a gift of Dr Francesco Parlanti (Onyx Inc). ADEP2 was kindly provided by Dr Heike Brötz-Oesterhelt (Heinrich-Heine-University Düsseldorf).

Bacterial strains, plasmids, expression and growth of cells

M. smegmatis mc²155 were grown at 37°C in Middlebrook 7H9 broth with 0.05% Tween-80 and ADC (0.5% BSA, 0.2% dextrose, 0.085% NaCl, 0.003 g catalase/1 L media) supplemented with hygromycin (50 µg/ml) and in case of inducible expression of ClpP1, ClpP2, or ClpC1, with anhydrotetracycline (100 ng/ml). Full-length C-terminally 6 × His-, Myc- or 6 × His and Myc-tagged ClpP1 and ClpP2 proteins were expressed in *M. smegmatis* on pMV plasmid under the regulation of a constitutive GroEL promoter. For expression of shorter forms of polypeptides corresponding to processed ClpP1 (lacking 6 N-terminal amino acids) and ClpP2 (lacking 11 N-terminal amino acids), pTetOR plasmid, which has an inducible tetracycline promoter was used. After overnight induction with ATc (100 ng/ml), cells were collected and kept at −70°C.

Purification of Mtb ClpP1 and ClpP2

All procedures for enzyme purification were carried out at 4°C using the following buffers:

Buffer A: 50 mM potassium phosphate pH 7.6 100 mM KCl, 5 mM MgCl₂, β-mercaptoethanol, and 10% glycerol; **buffer B:** 50 mM potassium phosphate pH 7.6 100 mM KCl, 5 mM MgCl₂, and 5% glycerol. In a typical purification, frozen cells (5–10 g) expressing ClpP1 or ClpP2 were suspended in two volumes of buffer A and broken by French press at 1500 p.s.i. The extract was centrifuged at 100 000 g and mixed with 5 ml Ni-NTA agarose previously equilibrated in buffer A. After incubating at 4°C for 4 h, Ni-NTA agarose resin was transferred to empty column and proteins were eluted using step gradient (0, 25, 50, 100, and 200 mM of imidazole in buffer B). The fractions containing near homogeneous ClpP1 or ClpP2 proteins were combined, concentrated on Millipore MWCO 10 000 cut filter and purified further by gel filtration on Sephacryl S-300 column (1.5 × 12 cm) equilibrated with buffer B. The specific activity of the purified enzyme was in the range 4–5 µmole/mg/min. High molecular weight protein peaks were combined, concentrated to ~2–5 mg/ml, and kept at −80°C.

Determination of enzymatic activities

All assays of a proteolytic activity were performed at 37°C in 96-well plate using Plate Reader SpectraMax M5 (Molecular Devices, USA). Wells contained 0.1 mM fluorescent peptide, 0.3–3 µg ClpP1P2, 0.5 mM Z-Leu-Leu-aldehyde or 5 mM Z-Leu-Leu in 80 µl of 50 mM phosphate buffer pH 7.6 with 5% glycerol and 100 mM KCl. After shaking for 20 s, peptidase activities were assayed at 37°C by continuously monitoring the rate of production of fluorescent 7-amino-4-methylcoumarin (AMC) from fluorogenic peptide substrates at 460 nm (Ex at 380 nm). Cleavage of longer octa- and nano-peptides was measured using FRET assay using the same conditions, except that 0.1–0.5 µg enzyme and 2–5 mM quenched substrates were used. Increase in fluorescence was monitored continuously at 405 nm (Ex at 340 nm). FITC-casein was purified using PD-10 column, and its hydrolysis (2–5 µg) was continuously monitored at 518 nm (Ex at 492 nm). All assays were performed in triplicate and average results presented. Deviations in the measurements for AMC and quench peptide substrates were <5%, while in the case of FITC-casein, it was <10%. Potential cleavage of enzyme activator Z-Leu-Leu was tested by fluorescence method as described previously (Akopian *et al*, 1997).

ATPase activity of ClpC1 was measured in the buffer containing 50 mM Tris-HCl pH 7.8, 50 mM KCl, 10% glycerol, 1 mM DTT, 2 mM ATP, and 8 mM MgCl₂. The amount of generated orthophosphate was measured colorimetrically by Malachite Green method (Baykov *et al*, 1988). The deviation between the measurements was <5%.

Fluorescent emission spectra

Emission spectra of ClpP1 complexes were determined in buffer B in microplate format using a SpectroMax M5 (USA) plate reader with a step of 1 nm. Preliminary spectra registration indicated a max of excitation of Trp174 in ClpP1 environment is 279 nm in buffer B.

MS and N-terminal analysis

Protein bands from SDS-PAGE were digested by sequence grade trypsin (Promega) or chymotrypsin (Roche). Obtained peptides were analysed by Thermo Electron LTQ-Orbitrap MassSpec after their separation by Agilent 1100 HPLC system. Automatic N-terminal sequencing of purified proteins (Edman degradation)

was done using ABI 494 Protein Sequencers after transfer of proteins onto PVDF membranes.

Crosslinking of neighbouring subunits

Crosslinking of ClpP1P2 was carried out with 0.125% glutaraldehyde in the buffer B containing activator Z-Leu-Leu. After 0.5 and 20 h incubation at room temperature, the reaction mixture was resolved by SDS-PAGE, and proteins analysed by MS.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

References

- Akopian TN, Kisselev AF, Goldberg AL (1997) Processive degradation of proteins and other catalytic properties of the proteasome from *Thermoplasma acidophilum*. *J Biol Chem* **272**: 1791–1798
- Andersson FI, Tryggvesson A, Sharon M, Diemand AV, Classen M, Best C, Schmidt R, Schelin J, Stanne TM, Bukau B, Robinson CV, Witt S, Mogk A, Clarke AK (2009) Structure and function of a novel type of ATP-dependent Clp protease. *J Biol Chem* **284**: 13519–13532
- Barik S, Sureka K, Mukherjee P, Basu J, Kundu M (2010) RseA, the SigE specific anti-sigma factor of *Mycobacterium tuberculosis*, is inactivated by phosphorylation-dependent ClpC1P2 proteolysis. *Mol Microbiol* **75**: 592–606
- Baumeister W, Walz J, Zühl F, Seemüller E (1998) The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92**: 367–380
- Baykov AA, Evtushenko OA, Avaeva SM (1988) A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay. *Anal Biochem* **171**: 266–270
- Benaroudj N, Raynal B, Miot M, Ortiz-Lombardia M (2011) Assembly and proteolytic processing of mycobacterial ClpP1 and ClpP2. *BMC Biochem* **12**: 61
- Bewley MC, Graziano V, Griffin K, Flanagan JM (2006) The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *J Struct Biol* **153**: 113–128
- Brotz-Oesterhelt H, Beyer D, Kroll HP, Endermann R, Ladel C, Schroeder W, Hinzen B, Raddatz S, Paulsen H, Henninger K, Bandow JE, Sahl HG, Labischinski H (2005) Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat Med* **11**: 1082–1087
- Butler SM, Festa RA, Pearce MJ, Darwin KH (2006) Self-compartmentalized bacterial proteases and pathogenesis. *Mol Microbiol* **60**: 553–562
- Flanagan JM, Wall JS, Capel MS, Schneider DK, Shanklin J (1995) Scanning transmission electron microscopy and small-angle scattering provide evidence that native *Escherichia coli* ClpP is a tetradecamer with an axial pore. *Biochemistry* **34**: 10910–10917
- Grimaud R, Kessel M, Beuron F, Steven AC, Maurizi MR (1998) Enzymatic and structural similarities between the *Escherichia coli* ATP-dependent proteases, ClpXP and ClpAP. *J Biol Chem* **273**: 12476–12481
- Groll M, Bajorek M, Kohler A, Moroder L, Rubin DM, Huber R, Glickman MH, Finley D (2000) A gated channel into the proteasome core particle. *Nat Struct Biol* **7**: 1062–1067
- Hoskins JR, Pak M, Maurizi MR, Wickner S (1998) The role of the ClpA chaperone in proteolysis by ClpAP. *Proc Natl Acad Sci USA* **95**: 12135–12140
- Huang H, Goldberg AL (1997) Proteolytic activity of the ATP-dependent protease HslVU can be uncoupled from ATP hydrolysis. *J Biol Chem* **272**: 21364–21372
- Hwang BJ, Park WJ, Chung CH, Goldberg AL (1987) *Escherichia coli* contains a soluble ATP-dependent protease (Ti) distinct from protease La. *Proc Natl Acad Sci USA* **84**: 5550–5554
- Hwang BJ, Woo KM, Goldberg AL, Chung CH (1988) Protease Ti, a new ATP-dependent protease in *Escherichia coli*, contains protein-activated ATPase and proteolytic functions in distinct subunits. *J Biol Chem* **263**: 8727–8734

Acknowledgements

These studies were supported by NIGMS 5R01 GM51923-16 and Harvard Catalyst Pilot Grants to ALG.

Author contributions: TA, OK, and ALG conceived and designed experiments; TA, OK, RR, and MU performed experiments; TA, OK, ER, and ALG analysed the data; TA, OK, and ALG wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

- Ingvarsson H, Mate MJ, Hogbom M, Portnoi D, Benaroudj N, Alzari PM, Ortiz-Lombardia M, Unge T (2007) 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
- Ishikawa T, Beuron F, Kessel M, Wickner S, Maurizi MR, Steven AC (2001) Translocation pathway of protein substrates in ClpAP protease. *Proc Natl Acad Sci USA* **98**: 4328–4333
- Kang SG, Dimitrova MN, Ortega J, Ginsburg A, Maurizi MR (2005) Human mitochondrial ClpP is a stable heptamer that assembles into a tetradecamer in the presence of ClpX. *J Biol Chem* **280**: 35424–35432
- Katayama-Fujimura Y, Gottesman S, Maurizi MR (1987) A multiple-component, ATP-dependent protease from *Escherichia coli*. *J Biol Chem* **262**: 4477–4485
- Kim YI, Levchenko I, Fraczowska K, Woodruff RV, Sauer RT, Baker TA (2001) Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nat Struct Biol* **8**: 230–233
- Kirstein J, Hoffmann A, Lilie H, Schmidt R, Rubsamen-Waigmann H, Brotz-Oesterhelt H, Mogk A, Turgay K (2009) The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. *EMBO Mol Med* **1**: 37–49
- Kisselev AF, Callard A, Goldberg AL (2006) Importance of the different proteolytic sites of the proteasome and the efficacy of inhibitors varies with the protein substrate. *J Biol Chem* **281**: 8582–8590
- Kisselev AF, Garcia-Calvo M, Overkleeft HS, Peterson E, Pennington MW, Ploegh HL, Thornberry NA, Goldberg AL (2003) The QJ; caspase-like sites of proteasomes, their substrate specificity, new inhibitors and substrates, and allosteric interactions with the trypsin-like sites. *J Biol Chem* **278**: 35869–35877
- Kress W, Maglica Z, Weber-Ban E (2009) Clp chaperone-proteases: structure and function. *Res Microbiol* **160**: 618–628
- Lee BG, Park EY, Lee KE, Jeon H, Sung KH, Paulsen H, Rubsamen-Schaeff H, Brotz-Oesterhelt H, Song HK (2010a) Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nat Struct Mol Biol* **17**: 471–478
- Lee JW, Park E, Jeong MS, Jeon YJ, Eom SH, Seol JH, Chung CH (2009) HslVU ATP-dependent protease utilizes maximally six among twelve threonine active sites during proteolysis. *J Biol Chem* **284**: 33475–33484
- Lee ME, Baker TA, Sauer RT (2010b) Control of substrate gating and translocation into ClpP by channel residues and ClpX binding. *J Mol Biol* **399**: 707–718
- Leung E, Datti A, Cossette M, Goodreid J, McCaw SE, Mah M, Nakhamchik A, Ogata K, El Bakkouri M, Cheng YQ, Wodak SJ, Eger BT, Pai EF, Liu J, Gray-Owen S, Batey RA, Houry WA (2011) Activators of cylindrical proteases as antimicrobials: identification and development of small molecule activators of ClpP protease. *Chem Biol* **18**: 1167–1178
- Liu Y, Patricelli MP, Cravatt BF (1999) Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci USA* **96**: 14694–14699
- Maurizi MR (1991) ATP-promoted interaction between Clp A and Clp P in activation of Clp protease from *Escherichia coli*. *Biochem Soc Trans* **19**: 719–723
- Maurizi MR, Clark WP, Katayama Y, Rudikoff S, Pumphrey J, Bowers B, Gottesman S (1990a) Sequence and structure of Clp

- P, the proteolytic component of the ATP-dependent Clp protease of *Escherichia coli*. *J Biol Chem* **265**: 12536–12545
- Maurizi MR, Clark WP, Kim SH, Gottesman S (1990b) Clp P represents a unique family of serine proteases. *J Biol Chem* **265**: 12546–12552
- Maurizi MR, Singh SK, Thompson MW, Kessel M, Ginsburg A (1998) Molecular properties of ClpAP protease of *Escherichia coli*: ATP-dependent association of ClpA and clpP. *Biochemistry* **37**: 7778–7786
- Maurizi MR, Thompson MW, Singh SK, Kim SH (1994) Endopeptidase Clp: ATP-dependent Clp protease from *Escherichia coli*. *Methods Enzymol* **244**: 314–331
- Ollinger J, O'Malley T, Kesicki EA, Odingo J, Parish T (2012) Validation of the essential ClpP protease in *Mycobacterium tuberculosis* as a novel drug target. *J Bacteriol* **194**: 663–668
- Ortega J, Singh SK, Ishikawa T, Maurizi MR, Steven AC (2000) Visualization of substrate binding and translocation by the ATP-dependent protease, ClpXP. *Mol Cell* **6**: 1515–1521
- Peltier JB, Ripoll DR, Friso G, Rudella A, Cai Y, Ytterberg J, Giacomelli L, Pillardy J, van Wijk KJ (2004) Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted three-dimensional structures, and functional implications. *J Biol Chem* **279**: 4768–4781
- Peltier JB, Ytterberg J, Liberles DA, Roepstorff P, van Wijk KJ (2001) Identification of a 350-kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of *Arabidopsis thaliana*. *J Biol Chem* **276**: 16318–16327
- Porankiewicz J, Wang J, Clarke AK (1999) New insights into the ATP-dependent Clp protease: *Escherichia coli* and beyond. *Mol Microbiol* **32**: 449–458
- Powers JC, Kam CM (1994) Isocoumarin inhibitors of serine peptidases. *Methods Enzymol* **244**: 442–457
- Rabl J, Smith DM, Yu Y, Chang SC, Goldberg AL, Cheng Y (2008) Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases. *Mol Cell* **30**: 360–368
- Raju R, Unnikrishnan M, Rubin D, Krishnamoorthy V, Kandror O, Akopian T, Goldberg AL, Rubin EJ (2011) *Mycobacterium tuberculosis* encodes a heteromeric ClpP protease required for normal growth and virulence. *PLoS Pathog* (in press)
- Reid BG, Fenton WA, Horwich AL, Weber-Ban EU (2001) ClpA mediates directional translocation of substrate proteins into the ClpP protease. *Proc Natl Acad Sci USA* **98**: 3768–3772
- Sasseti CM, Boyd DH, Rubin EJ (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**: 77–84
- Schelin J, Lindmark F, Clarke AK (2002) The clpP multigene family for the ATP-dependent Clp protease in the cyanobacterium *Synechococcus*. *Microbiology* **148**: 2255–2265
- Schmitt EK, Riwanto M, Sambandamurthy V, Roggo S, Miault C, Zwingelstein C, Krastel P, Noble C, Beer D, Rao SP, Au M, Niyomrattanakit P, Lim V, Zheng J, Jeffery D, Pethe K, Camacho LR (2011) The natural product cyclomarin kills *Mycobacterium tuberculosis* by targeting the ClpC1 subunit of the caseinolytic protease. *Angew Chem Int Ed Engl* **50**: 5889–5891
- Seol JH, Yoo SJ, Shin DH, Shim YK, Kang MS, Goldberg AL, Chung CH (1997) The heat-shock protein HslVU from *Escherichia coli* is a protein-activated ATPase as well as an ATP-dependent proteinase. *Eur J Biochem* **247**: 1143–1150
- Sherid AM, Rustad TR, Cangelosi GA, Sherman DR (2010) Characterization of a Clp protease gene regulator and the re-aeration response in *Mycobacterium tuberculosis*. *PLoS One* **5**: e11622
- Shin DH, Lee CS, Chung CH, Suh SW (1996) Molecular symmetry of the ClpP component of the ATP-dependent Clp protease, an *Escherichia coli* homolog of 20 S proteasome. *J Mol Biol* **262**: 71–76
- Sjogren LL, Stanne TM, Zheng B, Sutinen S, Clarke AK (2006) Structural and functional insights into the chloroplast ATP-dependent Clp protease in *Arabidopsis*. *Plant Cell* **18**: 2635–2649
- Smith DM, Chang SC, Park S, Finley D, Cheng Y, Goldberg AL (2007) Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry. *Mol Cell* **27**: 731–744
- Smith N, Naude RJ, Oelofsen W, Lazure C, Patthy A (1992) The isolation and partial characterization of trypsinogen, pancreatic secretory trypsin inhibitor and multiple forms of chymotrypsinogen and trypsin from the pancreas of the ostrich (*Struthio camelus*). *Int J Biochem* **24**: 877–885
- Sousa MC, Kessler BM, Overkleeft HS, McKay DB (2002) Crystal structure of HslUV complexed with a vinyl sulfone inhibitor: corroboration of a proposed mechanism of allosteric activation of HslV by HslU. *J Mol Biol* **318**: 779–785
- Stanne TM, Pojidaeva E, Andersson FI, Clarke AK (2007) Distinctive types of ATP-dependent Clp proteases in cyanobacteria. *J Biol Chem* **282**: 14394–14402
- Striebel F, Kress W, Weber-Ban E (2009) Controlled destruction: AAA+ ATPases in protein degradation from bacteria to eukaryotes. *Curr Opin Struct Biol* **19**: 209–217
- Szyk A, Maurizi MR (2006) Crystal structure at 1.9 Å of *E. coli* ClpP with a peptide covalently bound at the active site. *J Struct Biol* **156**: 165–174
- Thompson MW, Maurizi MR (1994) Activity and specificity of *Escherichia coli* ClpAP protease in cleaving model peptide substrates. *J Biol Chem* **269**: 18201–18208
- Viala J, Mazodier P (2002) ClpP-dependent degradation of PopR allows tightly regulated expression of the clpP3 clpP4 operon in *Streptomyces lividans*. *Mol Microbiol* **44**: 633–643
- Viala J, Rapoport G, Mazodier P (2000) The clpP multigenic family in *Streptomyces lividans*: conditional expression of the clpP3 clpP4 operon is controlled by PopR, a novel transcriptional activator. *Mol Microbiol* **38**: 602–612
- Wang J, Hartling JA, Flanagan JM (1997) The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. *Cell* **91**: 447–456
- Waxman L, Goldberg AL (1982) Protease La from *Escherichia coli* hydrolyzes ATP and proteins in a linked fashion. *Proc Natl Acad Sci USA* **79**: 4883–4887
- Whitby FG, Masters EI, Kramer L, Knowlton JR, Yao Y, Wang CC, Hill CP (2000) Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature* **408**: 115–120
- Yoo SJ, Seol JH, Shin DH, Rohrwild M, Kang MS, Tanaka K, Goldberg AL, Chung CH (1996) Purification and characterization of the heat shock proteins HslV and HslU that form a new ATP-dependent protease in *Escherichia coli*. *J Biol Chem* **271**: 14035–14040
- Yu AY, Houry WA (2007) ClpP: a distinctive family of cylindrical energy-dependent serine proteases. *FEBS Lett* **581**: 3749–3757