



# HHS Public Access

Author manuscript

*Curr Opin Struct Biol.* Author manuscript; available in PMC 2022 April 01.

Published in final edited form as:

*Curr Opin Struct Biol.* 2021 April ; 67: 120–126. doi:10.1016/j.sbi.2020.09.012.

## Structural Determinants of Regulated Proteolysis in Pathogenic Bacteria by ClpP and the Proteasome

Shoshanna C. Kahne, K. Heran Darwin\*

New York University Robert Grossman School of Medicine, Department of Microbiology, 430 E. 29th Street, Room 312, New York, NY 10016, USA

### Abstract

Bacteria use gated proteolytic machines for routine protein quality control and regulated responses to environmental conditions. This review discusses recent advances in understanding the structure and regulation of ClpP proteases, nanomachines widely distributed across bacteria, and the bacterial proteasome, a protease found in relatively few species. For both machines, activators confer substrate specificity. We highlight new data from organisms encoding two ClpP isoforms and the central role of activators as platforms for integrating regulatory signals. Because proteolytic systems contribute to survival and virulence of many bacterial pathogens, understanding their forms and functions enables new approaches to design targeted therapeutics.

### Keywords

bacteria; pathogens; proteases; regulation

### Introduction

Regulated protein degradation is essential for all organisms. In bacteria, regulated proteolysis is required for cell division, acclimation to environmental stress and transition, and virulence [1]. Because dysregulation of proteolysis can be detrimental, a structural understanding of how these machines are regulated may guide the design of new antimicrobial therapeutics [2].

Most bacteria contain at least two of five well-described compartmentalized proteases, which are barrel shaped and closed on both ends, or gated, to prevent non-specific or constitutive degradation (Table 1, gray columns). Conservation of multiple proteolytic machines, as well as interchangeable ClpP and proteasome activators (Table 1, white columns), within each species suggests they degrade different groups of substrates, under specific conditions, or both. This review summarizes recent insights into structural

\*To whom correspondence should be addressed: heran.darwin@med.nyu.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest statement  
Nothing to declare.

determinants of regulated proteolysis in ClpXP, one of the most ubiquitous proteolytic complexes, and the bacterial proteasome, which is conserved in all domains of life but found only in bacteria of the orders *Actinomycetales* and *Nitrospirales*.

## The Proteases

### ClpP.

Caseinolytic protease, ClpP, is a serine protease found across the bacterial kingdom. Crystal structures from more than a dozen species reveal a conserved architecture: two heptameric rings of ClpP protomers stack co-axially forming a barrel lined with active sites (Figure 1A). In the absence of an activating complex (such as the ClpX hexamer shown in Figure 1A, and discussed below) or a chemical, N-terminal tails of ClpP block entry to both barrel ends, occluding large, folded proteins from the proteolytic chamber.

Pioneering research on ClpP structure and function focused on bacteria with a single *clpP* gene, including *Escherichia coli* and *Staphylococcus aureus*. However, many pathogenic bacteria encode two *clpP* alleles, designated *clpP1* and *clpP2*. The variety of ways these genes are organized in the genome (Figure 2A), evidence of differential expression in some species (Figure 2B), and amino acid variation among isoforms (Figure 2C) has recently inspired investigation into ClpP structure and function in these organisms. Primarily this work has evaluated if ClpP1 and ClpP2 assemble as functional homo- or hetero-tetradecameric complexes.

*Mycobacterium spp.* and *Listeria monocytogenes* proteases are the most thoroughly investigated complexes to date and the only ones with high-resolution structures (Figure 2D). *M. tuberculosis* ClpP1 and ClpP2 are pulled down as hetero-tetradecameric complexes (ClpP1P2) from *Mycobacterium smegmatis* [3], but not from *E. coli* [4], and do not form active homo-tetradecamers [4–6]. Hetero-tetradecameric activity *in vitro* requires purified ClpP1 and ClpP2 lacking pro-peptides [6], an activating complex or chemical (such as hexameric ClpX or acyldepsipeptides, “ADEPs”) bound to hydrophobic clefts between ClpP subunits to induce restructuring of the amino (N)-terminal tails and open the channel into the protease, and a substrate or a substrate mimic (such as N-blocked peptide aldehyde) [6–9].

In contrast, *L. monocytogenes* ClpP1P2 can be purified as an active hetero-tetradecamer from *E. coli* [10, 11]. However, homo-tetradecameric ClpP2 from *L. monocytogenes* is also active. Homo-tetradecameric ClpP2 has greater peptidase activity than the hetero-tetradecamer, but in complex with ClpX it has almost ten times lower proteolytic activity than the hetero-tetradecamer, even compared to ClpP1P2 complexes containing either inactive ClpP1 or ClpP2 [11]. The greater activity of ClpXP1P2 over ClpXP2 is due in part to a greater affinity of ClpX for the heterotetradecamer and faster unfolding of substrates by ClpX bound to ClpP1P2 [12], although the structural determinants of these differences have not yet been identified. Because ClpP1 and ClpP2 have different substrate preferences in *L. monocytogenes* [12], determining if there are biological consequences that depend on the relative abundance of the two ClpP machines will be a very exciting area for further work.

Hetero-tetradecameric complexes from *M. tuberculosis* and *L. monocytogenes* are similar: both include a homo-heptamer of each ClpP isoform, coaxially stacked face-to-face [3, 6, 8, 10, 11]. In *M. tuberculosis*, *clpP1* and *clpP2* are arranged in a bicistronic operon [13] that is essential for viability [3, 14], a unique feature among ClpP-containing bacteria. However, operonic organization is not required for assembly of hetero-tetradecameric complexes, as *clpP1* and *clpP2* are not co-transcribed in *L. monocytogenes* (Figure 2A).

Investigation of ClpP in other bacteria with two ClpP isoforms is more limited (Figure 2D). *clpP1* and *clpP2* are differentially expressed in *Pseudomonas aeruginosa* and *Clostridium difficile*, and some homo-tetradecamers show *in vitro* activity (Figure 2B, 2D) [15, 16]••. However, without pull-downs or other *in vivo* data from their endogenous organisms, the physiologic relevance of these structures is unknown.

Contradictory data from *Chlamydia trachomatis* provide a cautionary example of the limits of extrapolation of *in vitro* data. Initial experiments with *C. trachomatis* ClpP2 purified from *E. coli* revealed an apparently proteolytically active homo-tetradecamer [17]. However, ClpP2 purified from *E. coli* lacking endogenous ClpP has no *in vitro* activity and robustly forms hetero-tetradecamers when mixed with ClpP1 purified from the same strain [18] suggesting *E. coli* ClpP contaminated proteins purified in the first experiments (precedent established by [4, 16]••). Investigation of these structures from their native organisms could help avoid this issue as well as help identify other factors needed for protease processing or assembly (as has been found for *Mycobacteria spp.*).

Fascinatingly, active hetero-tetradecameric ClpP structures appear to be common among species containing multiple *clpP* genes (Figure 2D). Whether a second *clpP* isoform arose in these species by gene duplication or horizontal gene transfer, studying diverse ClpP structures may reveal novel functions distinguishing hetero-tetradecamers from homo-tetradecamers that help explain their prevalence in organisms containing multiple *clpP* genes.

### Proteasomes.

Proteasomes are not as widely conserved across the bacterial kingdom as ClpP, but are present in several human pathogens, including *M. tuberculosis* and *M. leprae* (Table 1). For comprehensive details, we refer the reader to a recent review of bacterial proteasome structure and function [19]. Bacterial proteasomes, like their eukaryotic homologs, are composed of 28 subunits: two homo-heptameric rings of  $\beta$ -subunits containing the active sites are sandwiched between two homo-heptameric rings of  $\alpha$ -subunits that block entry into the cylindrical proteolytic chamber in the absence of an activator. Most bacterial proteasomes studied to date consist of single  $\alpha$  and  $\beta$  isoforms. The *Rhodococcus erythropolis* proteasome is a notable exception, as proteasomes purified from *R. erythropolis* contain two  $\alpha$  and two  $\beta$  isoforms [20, 21]. Although all pairwise combinations of  $\alpha$  and  $\beta$  gene products form functional proteasomes [20], their physiological relevance is unknown. The arrangement of the  $\alpha$  and  $\beta$  isoforms within *R. erythropolis* remains unknown, and as such, so do any physiological consequences of having multiple proteasome subunit alleles.

## Activators: platforms for regulating substrate specificity

While many bacteria encode multiple ClpP activators, proteasome-containing bacteria studied to date encode only two well-established activators (Table 1). Activators of both proteolytic machines are associated with the degradation of distinct substrate sets and thus represent a modular regulatory mechanism determining which, when, and, in at least one case, where substrates are degraded.

### ClpP activators and regulatory models.

The degradation of folded proteins by ClpP complexes requires binding of a hexameric AAA+ (ATPases Associated with diverse cellular Activities) activator. Activators help recruit substrates and couple ATP hydrolysis with substrate unfolding and threading into a proteolytic chamber. Perhaps the best characterized activator is ClpX, in part because of its role in degrading *ssrA*-tagged polypeptides (for history and biotechnological implications, see [22]). Furthermore, ClpX is found in all of the organisms covered in this review (Table 1). The N-terminus of ClpX has a zinc binding domain (ZBD) that can facilitate degradation specificity by binding to regulators known as adaptors, which modulate degradation efficiency of substrates (for an excellent review on this subject, see [1]). One of the first adaptors discovered was SspB (stringent starvation protein B), which can positively or negatively regulate substrates of ClpXP [23]. SspB can bind to *ssrA*-tagged proteins [24–27] as well as untagged proteins like the N-terminal fragment of anti-sigma-E factor RseA [23, 28]; SspB-adaptor complexes then bind to ClpX's ZBD. This positioning may be critical to point substrates into the activator pore towards the AAA+ domains that are required for facilitating degradation.

Another adaptor is RssB (regulator of sigma S protein B), which stimulates degradation of the sigma factor RpoS, required for the expression of numerous genes during stationary phase, starvation and other stressful conditions. Under unstressed growth conditions, RssB interacts with both RpoS and the ClpX ZBD to degrade RpoS by ClpP [29, 30]. During stress, anti-adaptors are produced that bind to RssB, preventing RpoS degradation. How RpoS is transferred from RssB to ClpXP for degradation, and how anti-adaptors negatively regulate this process, has been difficult to dissect in the absence of a full-length crystal structure of RssB. The recently solution of the first crystal structure of full-length RssB bound to an anti-adaptor, IraD (Inhibitor of RssB activity after DNA damage)[31] • and identification of the RssB sequence important for ClpX interaction[29] support a hand off model of delivery of RpoS to ClpXP by RssB. In the IraD-bound RssB structure, RssB adopts a closed conformation preventing RpoS binding [31]•. Because IraD and ClpX interact with an overlapping region of RssB, it is reasonable to hypothesize that RssB-RpoS binding to ClpX induces closing of the RssB structure resulting in the hand off of RpoS to ClpXP [29]•.

The two newest models of ClpXP regulation come from *Caulobacter crescentus*. Although this organism is non-pathogenic, its distinct life cycle that includes asymmetric differentiation during cell division provides useful phenotypes for investigating the regulation of bacterial development by proteolysis. In *C. crescentus*, a group of adaptors bind to ClpXP in a hierarchical manner to stimulate the degradation of substrates in an

ordered sequence [32]. Additionally, during cell division, ClpXP must be localized to one pole to limit degradation[33]•, providing the first example of sub-cellular organization as a regulatory model for ClpXP regulation.

### Proteasome regulators.

All proteasome-containing bacteria identified to date have an ATP-dependent and an ATP-independent proteasome activator (Table 1). Mpa/ARC (mycobacterial proteasome ATPase/ATPase forming ring-shaped complexes in non-mycobacteria) is an AAA ATPase that forms hexameric rings, and acts as a substrate receptor, unfoldase, and protease gate opener [34, 35]. Mpa/ARC interacts with substrates post-translationally modified with Pup (prokaryotic ubiquitin-like protein) [36, 37]. Although numerous proteins are pupylated in mycobacteria [38, 39], few structural or functional pupylation determinants are known, with the exception that the sole Pup ligase, PafA (proteasome accessory factor A), only pupylates surface exposed lysines [39–41].

Depupylation activity by the enzyme Dop (deamidase of Pup) also determines pupylation status [42, 43]. Deletion of a conserved disordered loop in *M. smegmatis* Dop increases depupylation without affecting substrate binding *in vitro* [44]. Like pupylation, determinants of depupylation have not been identified in *M. tuberculosis*. However, a recent discovery of the first condition in which pupylome abundance is altered in *M. tuberculosis* (growth in nitrate broth), without a change in abundance of known proteasome components, may enable the discovery of regulatory mechanisms of proteasomal degradation [45]•.

The proteasome can also be activated by an unusual, single-ring dodecameric ATP-independent activator, PafE/Bpa (proteasome accessory factor E/bacterial proteasome activator) (Figure 1B), which does not require a recognition tag to degrade substrates [46, 47]. In the absence of ATP-powered unfolding or recognition of post-translationally modified targets, PafE-mediated degradation may be somewhat stochastic, although specific substrates have been identified [47, 48]•. In contrast to Mpa, PafE is sufficient for the activation of gated proteasomes *in vitro* [47]. X-ray crystallography showed PafE interacts with proteasome  $\alpha$ -subunit rings using C-terminal glycine-glutamine-tyrosine-leucine, or “GQYL”, motifs that induce opening of proteasome core particles [48]•. Although the C-terminal tails of Mpa also contain GQYL motifs essential for function *in vivo* [47], they are “tucked” away in a crystal structure, and cannot facilitate degradation with wild-type proteasomes *in vitro* [35]. Collectively these data suggest Mpa-dependent degradation requires additional factors.

Recently a potential third proteasome activator, Cpa (Cdc48-like protein of actinobacteria), was identified based on its homology to eukaryotic activators and cooccurrence in genomes of proteasome-containing bacteria [49]. *R. erythropolis* Cpa forms hexameric rings in a nucleotide-dependent manner and competes with *M. tuberculosis* Mpa for complexing with “open gate” (OG) proteasomes *in vitro* [49]. OG proteasomes lack N-terminal residues that occlude access into the proteolytic chamber as well as allow for weak interactions with Mpa. However, Cpa does not contain the C-terminal GQYL motif and has not been pulled down in association with proteasomes. It remains to be determined how Cpa associates with, and affects activity of, the proteasome *in vivo*.

## Asymmetry between proteases and activators.

Activators with six- or 12-fold symmetry must somehow interact with the seven-fold symmetry of proteolytic barrel surfaces to stimulate degradation. Until recently, examination of this interface was not possible. Crystal structures of ClpP activated by ADEPs, antibiotics that bind to ClpP similarly to ClpX, enabled the first high-resolution structure-function studies of activated ClpP [50]. Although disordered in many structures, N-terminal loops of ClpP in some ADEP bound ClpP crystals were characterized in “up” or “down” conformations [50]. Until direct high-resolution of the ClpXP interface was possible, much theorizing regarding the asymmetry mismatch relied on the possibility that pseudo six-fold symmetry might be achieved by the conformations of these loops.

Although high-resolution structures of activator-bound proteasomes have not yet been obtained, three high-resolution structures of ClpXP have recently been determined: *L. monocytogenes* ClpXP1P2 [51]••, substrate-bound *E. coli* ClpXP [52]••, and substrate-bound *Neisseria meningitidis* ClpXP [53]••. In all structures, ClpX interacts with ClpP by inserting loops containing the motif isoleucine/leucine/valineglycine-phenylalanine/leucine (IGF) loops, into hydrophobic clefts between ClpP subunits (Figure 1A). IGF loop flexibility enables distortions necessary to accommodate the asymmetry, and consequently ClpX is offset and tilted relative to ClpP (Figure 1A), generating a bend in the central channel. In all three structures, all of the N-terminal loops of ClpP adopt the up conformation, clarifying the importance of this restructuring for activation of the protease, though not through formation of a pseudo six-fold symmetric interface.

Notwithstanding technical differences between the materials and methods, there are notable differences between these high-resolution ClpXP structures. In *E. coli* and *N. meningitidis*, ClpX subunits arrange in a spiral staircase resulting in a distorted “seam” subunit bridging the top and bottom of the ring. The position of the seam is related to nucleotide binding in both complexes. In contrast, *L. monocytogenes* ClpX subunits are regularly arranged with no seam. Additionally, ClpX binding to ClpP1P2 in *L. monocytogenes* did not induce widening of the ClpP entry pore, as observed in the *E. coli* and *N. meningitidis* structures.

The *L. monocytogenes* ClpXP1P2 is the sole hetero-tetradecameric complex solved at high-resolution to date. In all ClpP1P2 structures identified so far, ClpX interacts with only one heptameric ring (Figure 2D), in contrast to frequently observed doubly-capped homo-tetradecamers (used to model both the *E. coli* and *N. meningitidis* ClpXP structures). Structural characterization of diverse ClpXP complexes will ultimately establish if differences between these structures are due to *bona fide* species-specific differences or differences in complex isolation and analysis.

## Future perspectives

Given the critical roles of protein degradation in bacteria, the design of inhibitors or stimulants based on structural information may yield potent and specific antimicrobials. Understanding how these proteases are assembled and regulated requires more detailed structural studies, as well as continued genetic and biochemical analyses to identify accessory factors controlling proteolytic activity.



## Acknowledgements

Proteasome work in the Darwin lab is supported by NIH grants AI144851 and AI088075 awarded to K.H.D.

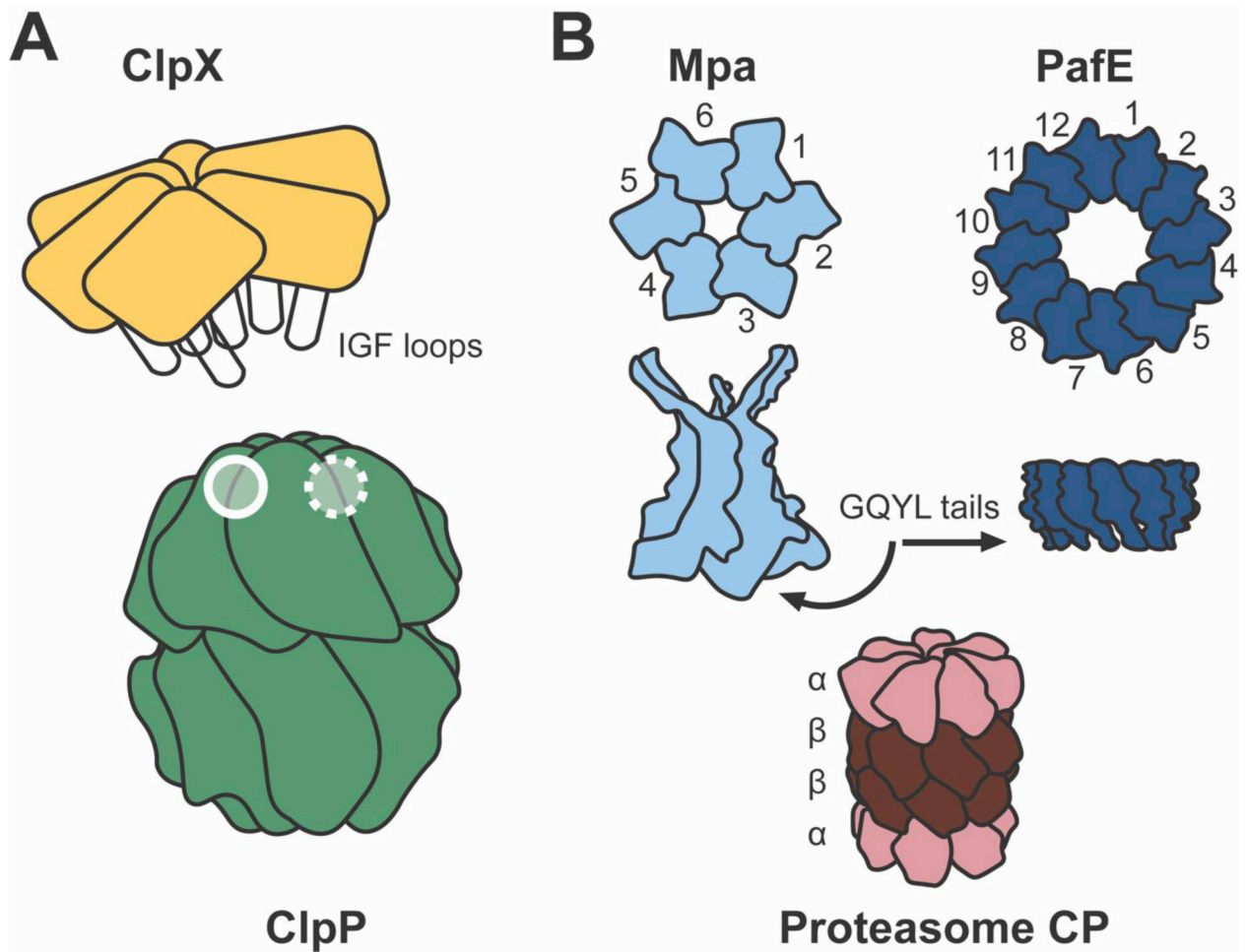
## References and recommended reading

1. Mahmoud SA and Chien P, Regulated Proteolysis in Bacteria. *Annu Rev Biochem*, 2018. 87: p. 677–696. [PubMed: 29648875]
2. Culp E and Wright GD, Bacterial proteases, untapped antimicrobial drug targets. *J Antibiot (Tokyo)*, 2017. 70(4): p. 366–377. [PubMed: 27899793]
3. Raju RM, et al., Mycobacterium tuberculosis ClpP1 and ClpP2 function together in protein degradation and are required for viability in vitro and during infection. *PLoS Pathog*, 2012. 8(2): p. e1002511. [PubMed: 22359499]
4. Benaroudj N, et al., Assembly and proteolytic processing of mycobacterial ClpP1 and ClpP2. *BMC Biochem*, 2011. 12: p. 61. [PubMed: 22132756]
5. 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*, 2007. 63(Pt 2): p. 249–59. [PubMed: 17242518]
6. 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*, 2012. 31(6): p. 1529–41. [PubMed: 22286948]
7. Schmitz KR and Sauer RT, Substrate delivery by the AAA+ ClpX and ClpC1 unfoldases activates the mycobacterial ClpP1P2 peptidase. *Mol Microbiol*, 2014. 93(4): p. 617–28. [PubMed: 24976069]
8. Schmitz KR, et al., Crystal structure of Mycobacterium tuberculosis ClpP1P2 suggests a model for peptidase activation by AAA+ partner binding and substrate delivery. *Proc Natl Acad Sci U S A*, 2014. 111(43): p. E4587–95. [PubMed: 25267638]
9. Vahidi S, et al., An allosteric switch regulates Mycobacterium tuberculosis ClpP1P2 protease function as established by cryo-EM and methyl-TROSY NMR. *Proc Natl Acad Sci U S A*, 2020. 117(11): p. 5895–5906. [PubMed: 32123115]
10. Zeiler E, et al., Vibrilactone as a tool to study the activity and structure of the ClpP1P2 complex from *Listeria monocytogenes*. *Angew Chem Int Ed Engl*, 2011. 50(46): p. 11001–4. [PubMed: 21954175]
11. Dahmen M, et al., Structure and mechanism of the caseinolytic protease ClpP1/2 heterocomplex from *Listeria monocytogenes*. *Angew Chem Int Ed Engl*, 2015. 54(12): p. 3598–602. [PubMed: 25630955]
12. Balogh D, et al., Insights into ClpXP proteolysis: heterooligomerization and partial deactivation enhance chaperone affinity and substrate turnover in *Listeria monocytogenes*. *Chem Sci*, 2017. 8(2): p. 1592–1600. [PubMed: 28451288]
13. Personne Y, et al., Mycobacterium tuberculosis ClpP proteases are co-transcribed but exhibit different substrate specificities. *PLoS One*, 2013. 8(4): p. e60228. [PubMed: 23560081]
14. Sassetti CM, Boyd DH, and Rubin EJ, Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol*, 2003. 48(1): p. 77–84. [PubMed: 12657046]
15. Hall BM, et al., Two Isoforms of Clp Peptidase in *Pseudomonas aeruginosa* Control Distinct Aspects of Cellular Physiology. *J Bacteriol*, 2017. 199(3).
- 16••. Lavey NP, et al., Clostridium difficile ClpP Homologues are Capable of Uncoupled Activity and Exhibit Different Levels of Susceptibility to Acyldepsipeptide Modulation. *ACS Infect Dis*, 2019. 5(1): p. 79–89. [PubMed: 30411608] Investigation of ClpP isoforms in *C. difficile* revealed two active homo-tetradecameric structures, an exception among bacteria.
17. Wood NA, et al., Initial Characterization of the Two ClpP Paralogs of *Chlamydia trachomatis* Suggests Unique Functionality for Each. *J Bacteriol*, 2019. 201(2).
18. Pan S, et al., The functional ClpXP protease of *Chlamydia trachomatis* requires distinct clpP genes from separate genetic loci. *Sci Rep*, 2019. 9(1): p. 14129. [PubMed: 31575885]
19. Harris JR and Marles-Wright J, *Macromolecular Protein Complexes II: Structure and Function*. 2019: Springer.

20. Zuhl F, et al., Subunit topology of the Rhodococcus proteasome. *FEBS Lett*, 1997. 400(1): p. 83–90. [PubMed: 9000518]
21. Tamura T, et al., The first characterization of a eubacterial proteasome: the 20S complex of Rhodococcus. *Curr Biol*, 1995. 5(7): p. 766–74. [PubMed: 7583123]
22. Fritze J, et al., An overview of the bacterial SsrA system modulating intracellular protein levels and activities. *Appl Microbiol Biotechnol*, 2020. 104(12): p. 5229–5241. [PubMed: 32342145]
23. Flynn JM, et al., Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. *Genes Dev*, 2004. 18(18): p. 2292–301. [PubMed: 15371343]
24. Levchenko I, et al., A specificity-enhancing factor for the ClpXP degradation machine. *Science*, 2000. 289(5488): p. 2354–6. [PubMed: 11009422]
25. Thibault G, et al., Specificity in substrate and cofactor recognition by the N-terminal domain of the chaperone ClpX. *Proc Natl Acad Sci U S A*, 2006. 103(47): p. 17724–9. [PubMed: 17090685]
26. Bolon DN, et al., Nucleotide-dependent substrate handoff from the SspB adaptor to the AAA+ ClpXP protease. *Mol Cell*, 2004. 16(3): p. 343–50. [PubMed: 15525508]
27. Park EY, et al., Structural basis of SspB-tail recognition by the zinc binding domain of ClpX. *J Mol Biol*, 2007. 367(2): p. 514–26. [PubMed: 17258768]
28. Levchenko I, et al., Versatile modes of peptide recognition by the AAA+ adaptor protein SspB. *Nat Struct Mol Biol*, 2005. 12(6): p. 520–5. [PubMed: 15880122]
- 29•. Micevski D, et al., Insight into the RssB-Mediated Recognition and Delivery of sigma(s) to the AAA+ Protease, ClpXP. *Biomolecules*, 2020. 10(4). Crystal structures of the N- and C-termini of RssB were solved and, together with biochemical analysis, a model unifying the repressive action of anti-adaptors and mechanism of RpoS delivery to ClpXP was proposed.
30. Dougan DA, Weber-Ban E, and Bukau B, Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX. *Mol Cell*, 2003. 12(2): p. 373–80. [PubMed: 14536077]
- 31••. Dorich V, et al., Structural basis for inhibition of a response regulator of sigma(S) stability by a ClpXP antiadaptor. *Genes Dev*, 2019. 33(11–12): p. 718–732. [PubMed: 30975721] First crystal structure of anti-adaptor bound RssB revealed a closed architecture of contacts between distal residues and key structural features for modulating this unique response-regulator.
32. Joshi KK, et al., An Adaptor Hierarchy Regulates Proteolysis during a Bacterial Cell Cycle. *Cell*, 2015. 163(2): p. 419–31. [PubMed: 26451486]
- 33•. Joshi KK, Battle CM, and Chien P, Polar Localization Hub Protein PopZ Restrains Adaptor-Dependent ClpXP Proteolysis in *Caulobacter crescentus*. *J Bacteriol*, 2018. 200(20). The first model of subcellular compartmentalization as a model for ClpXP regulation.
34. Wang T, Darwin KH, and Li H, Binding-induced folding of prokaryotic ubiquitin-like protein on the Mycobacterium proteasomal ATPase targets substrates for degradation. *Nat Struct Mol Biol*, 2010. 17(11): p. 1352–7. [PubMed: 20953180]
35. Wu Y, et al., Mycobacterium tuberculosis proteasomal ATPase Mpa has a beta-grasp domain that hinders docking with the proteasome core protease. *Mol Microbiol*, 2017. 105(2): p. 227–241. [PubMed: 28419599]
36. Pearce MJ, et al., Ubiquitin-Like Protein Involved in the Proteasome Pathway of Mycobacterium tuberculosis. *Science*, 2008. 322(5904): p. 1104–1107. [PubMed: 18832610]
37. Burns KE, et al., Proteasomal protein degradation in Mycobacteria is dependent upon a prokaryotic ubiquitin-like protein. *J Biol Chem*, 2009. 284(5): p. 3069–75. [PubMed: 19028679]
38. Festa RA, et al., Prokaryotic ubiquitin-like protein (Pup) proteome of Mycobacterium tuberculosis [corrected]. *PLoS One*, 2010. 5(1): p. e8589. [PubMed: 20066036]
39. Watrous J, et al., Expansion of the mycobacterial “PUPylome”. *Mol Biosyst*, 2010. 6(2): p. 376–85. [PubMed: 20094657]
40. Guth E, Thommen M, and Weber-Ban E, Mycobacterial ubiquitin-like protein ligase PafA follows a two-step reaction pathway with a phosphorylated pup intermediate. *J Biol Chem*, 2011. 286(6): p. 4412–9. [PubMed: 21081505]

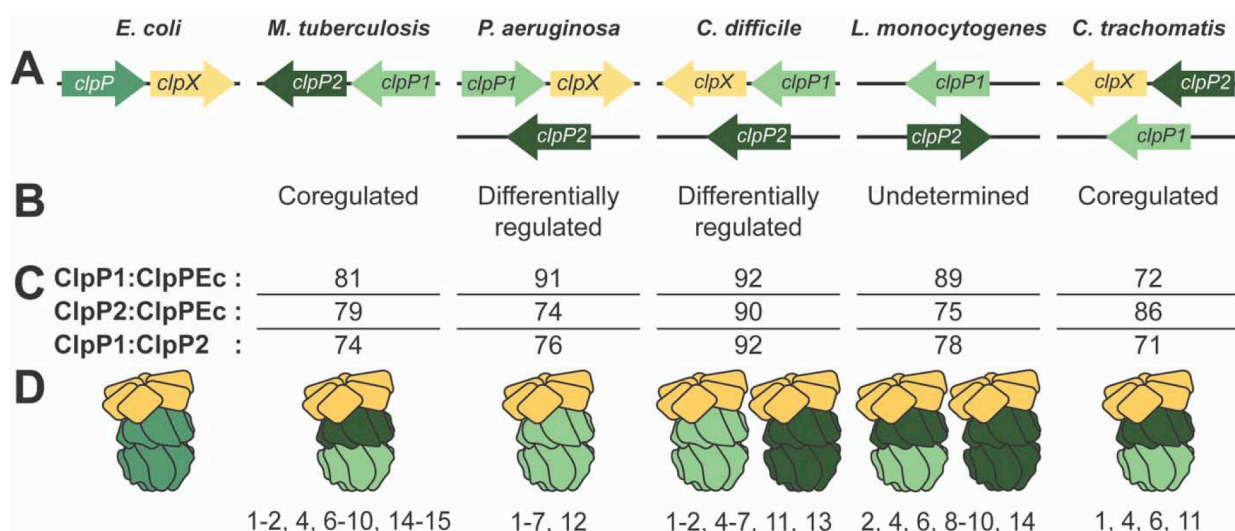


41. Striebel F, et al., Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. *Nat Struct Mol Biol*, 2009. 16(6): p. 647–51. [PubMed: 19448618]
42. Burns KE, et al., “Depupylation” of prokaryotic ubiquitin-like protein from mycobacterial proteasome substrates. *Mol Cell*, 2010. 39(5): p. 821–7. [PubMed: 20705495]
43. Imkamp F, et al., Dop functions as a depupylase in the prokaryotic ubiquitin-like modification pathway. *EMBO Rep*, 2010. 11(10): p. 791–7. [PubMed: 20798673]
44. Hecht N, et al., Inter- and intramolecular regulation of protein depupylation in *Mycobacterium smegmatis*. *FEBS J*, 2020.
- 45•. Becker SH, et al., The *Mycobacterium tuberculosis* Pup-proteasome system regulates nitrate metabolism through an essential protein quality control pathway. *Proc Natl Acad Sci U S A*, 2019. 116(8): p. 3202–3210. [PubMed: 30723150] Discovery of the first condition in which global pupylome abundance is altered in *M. tuberculosis*, providing a phenotype for determining regulators of Pup-proteasome enzyme activities.
46. Delley CL, et al., Bacterial proteasome activator bpa (rv3780) is a novel ring-shaped interactor of the mycobacterial proteasome. *PLoS One*, 2014. 9(12): p. e114348. [PubMed: 25469515]
47. Jastrab JB, et al., An adenosine triphosphate-independent proteasome activator contributes to the virulence of *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A*, 2015. 112(14): p. E1763–72. [PubMed: 25831519]
- 48•. Hu K, et al., Proteasome substrate capture and gate opening by the accessory factor PafE from *Mycobacterium tuberculosis*. *J Biol Chem*, 2018. 293(13): p. 4713–4723. [PubMed: 29414791] Crystallization of the ATP-independent proteasome activator revealed an unusual dodecameric ring, with extended tails containing GQYL motifs necessary and sufficient for proteasome activation *in vitro*, and necessary for both *in vivo*.
49. Ziemski M, et al., Cdc48-like protein of actinobacteria (Cpa) is a novel proteasome interactor in mycobacteria and related organisms. *Elife*, 2018. 7.
50. Alexopoulos JA, Guarne A, and Ortega J, ClpP: a structurally dynamic protease regulated by AAA + proteins. *J Struct Biol*, 2012. 179(2): p. 202–10. [PubMed: 22595189]
- 51••. Gatsogiannis C, et al., Cryo-EM structure of the ClpXP protein degradation machinery. *Nat Struct Mol Biol*, 2019. 26(10): p. 946–954. [PubMed: 31582852] The first high resolution cryo-EM structure of *Listeria monocytogenes* ClpXP1P2, revealing offset binding of ClpX to accommodate asymmetry at the ClpX-ClpP2 interface.
- 52••. Fei X, et al., Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. *Elife*, 2020. 9. Greatest resolution of the substrate-bound *E. coli* ClpXP protease to date, including a comprehensive discussion of structural and biochemical data challenging different models for how ATP-hydrolysis is coupled to substrate processing across AAA+ machines.
- 53••. Ripstein ZA, et al., A processive rotary mechanism couples substrate unfolding and proteolysis in the ClpXP degradation machinery. *Elife*, 2020. 9. Determination of the substrate-bound ClpXP structure revealed two ClpX conformations, improved resolution of details at the ClpX-ClpP2 interface, ClpX substrate contacts, and relationship between nucleotide binding and structure in *Neisseria meningitidis*.
54. Huang X and Miller W, A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math*, 1991. 12: p. 373–381.



**Figure 1: Structural features of ClpP, the bacterial proteasome, and their activators.**

(A) ClpX docks to ClpP via interactions between extended IGF-loops and hydrophobic clefts between ClpP subunits, illustrated by the open circles. Because the ClpP surface has seven-fold symmetry, one hydrophobic cleft is unoccupied in ClpXP, indicated by the open circle with a dashed outline. To accommodate the asymmetry of this interface, ClpX is laterally shifted and tilted relative to the ClpP central axis ( $11^\circ$  in *L. monocytogenes*[51]••,  $15^\circ$  in *N. meningitidis*[53]••), generating a bend in the channel through which substrates are extruded into the proteolytic machine. (B) The bacterial proteasome core particle (CP) is composed of two heptameric rings of  $\beta$ -subunits sandwiched between two heptameric rings of  $\alpha$ -subunits.  $\alpha$ -subunits restructure upon interaction with GQYL motifs in the tails of proteasome activators. The crystal structure of Mpa revealed the GQYL tails are tucked under the structure, hindering docking with the proteasome *in vitro*. PafE is a dodecameric ring that opens the proteasome using extended GQYL tails, enabling degradation *in vitro*.



**Figure 2: Structural and biochemical data reveal varied ClpP architectures in pathogenic bacterial species encoding two *clpP* genes.**

(A) Genomic organization of *clpP* genes. *clpX* is shown only when co-expressed with a *clpP* gene; when encoded elsewhere in the genome, *clpX* is not shown for clarity. Arrow direction indicates coding strand. (B) Summary of ClpP1/P2 proteomic and *clpP1/P2* expression data supporting either coregulation or differential regulation. (C) From top to bottom: pairwise amino acid similarities between *E. coli* ClpP (ClpPEc) and ClpP1 or ClpP2 from other species, as well as between ClpP1 and ClpP2 from the same species. Alignments were performed with LALIGN [54] (BLOSUM50 matrix, open gap penalty  $-12$ , gap extension penalty  $-2$ ). Notably, *C. difficile* ClpP proteins are more similar to ClpPEc, and to each other, than most other ClpP proteins. (D) Active ClpP machines identified in each species. Numbers listed below describe techniques used to validate these structures and rule out alternatives: (1) *in vitro* reconstitution with purified proteins, (2) size exclusion chromatography, (3) size exclusion chromatography-multiangle light scattering, (4) peptidase assay, (5) ClpX ATP hydrolysis rate, (6) *ssrA*-GFP degradation, (7) serine hydrolase probe, (8) negative-stain EM, (9) x-ray crystallography, (10) cryogenic EM, (11) native polyacrylamide gel electrophoresis, (12) sedimentation velocity-analytical ultracentrifugation, (13) dynamic light scattering, (14) pull down from bacteria and immunoblotting, (15) crosslinking and immunoblotting. Note that high-resolution structures (techniques 9 and 10) have only been obtained for *M. tuberculosis* and *L. monocytogenes* ClpP proteases. Because ClpX binding has been empirically tested for all structures illustrated, ClpX is shown in complex with the active tetradecamers.

**Table 1:**  
**Gated proteolytic machines and ClpP activators are found in diverse bacterial species.**

Tabular lists of genomic features were downloaded from the NCBI Genome base and annotations verified by BLAST, UniProt entries, and citations in PubMed. These annotations are based on the following strains: *E. coli* str. K-12 substr. MG1655, *N. meningitidis* MC58, *Caulobacter vibrioides* (*crenscentus*) CB15 *Staphylococcus aureus* subsp. aureus NCTC8325, *Bacillus subtilis* subsp. Subtilis str.168, *Pseudomonas aeruginosa* PAO1, *C. trachomatis* 434/Bu, *L. monocytogenes* EGD-e, *Clostridioides difficile* 630, *M. smegmatis* mc<sup>2</sup>155, *M. tuberculosis* H37Rv, and *M. leprae* TN. Gated proteases are in shaded columns, activators are unshaded.

Bacterial Species	FtsH	Lon	HsIVU	ClpP	ClpA	ClpX	ClpC	ClpE	Proteasome	Mpa	PafE
<i>Escherichia coli</i>	+	+	+	+	+	+	-	-	-		
<i>Neisseria meningitidis</i>	+	+	-	+	+	+	-	-	-		
<i>Caulobacter crescentus</i>	+	+	+	+	+	+	-	-	-		
<i>Staphylococcus aureus</i>	+	-	+	+	-	+	+	-	-		
<i>Bacillus subtilis</i>	+	2	+	+	-	+	+	+	-		
<i>Pseudomonas aeruginosa</i>	+	+	+	2	+	+	-	-	-		
<i>Chlamydia trachomatis</i>	+	+	-	2	-	+	-	-	-		
<i>Listeria monocytogenes</i>	+	-	+	2	-	+	+	+	-		
<i>Clostridium difficile</i>	+	+	-	2	-	+	+	-	-		
<i>Mycobacterium smegmatis</i>	+	+	-	2	-	+	+	-	+	+	+
<i>Mycobacterium tuberculosis</i>	+	-	-	2	-	+	+	-	+	+	+
<i>Mycobacterium leprae</i>	+	-	-	2	-	+	+	-	+	+	+