Functions of Alternative ClpP Subunits in *Pseudomonas aeruginosa*

by

Gina D. Mawla

B.S. Molecular, Cell and Developmental Biology
University of California, Santa Cruz, 2012

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN BIOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MAY 2020

Signature of the Author: _____

Department of Biology
29 April 2020

Certified by: _____

Tania A. Baker
E. C. Whitehead Professor of Biology
Thesis Supervisor

Accepted by: _____

Stephen P. Bell
Uncas and Helen Whitaker Professor of Biology
Investigator, Howard Hughes Medical Institute
Co-Director, Biology Graduate Committee

Functions of Alternative ClpP Subunits in *Pseudomonas aeruginosa*

by

Gina D. Mawla

Submitted to the Department of Biology
On April 29, 2020, in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in Biology

ABSTRACT

Proteolysis is the process by which proteins are broken down, or hydrolyzed, into small peptides or amino acids by enzymes. Cells from all forms of life carry out regulated protein degradation as a way to control cellular physiology and regulate stress responses. Clp proteases, containing a AAA+ (ATPases Associated with various cellular Activities) unfoldase stacked with a compartmentalized peptidase, are central to bacterial proteolysis, and use the energy of ATP hydrolysis to unfold and translocate protein substrates into the peptidase chamber for their destruction. The opportunistic pathogen *Pseudomonas aeruginosa* is unusual in that it contains two isoforms of the subunits that form the ClpP peptidase chamber. These isoforms, PaClpP1 and PaClpP2, have not been well characterized previously and their specific functions are largely elusive. This work examines the structures and functions of PaClpP1 and PaClpP2 and proposes a model for functional peptides generated by these enzymes in *P. aeruginosa* development. Biochemical analysis establishes that PaClpP2 is only active as a peptidase when it is part of a $PaClpP1_7P2_7$ heterocomplex. Furthermore, multiple lines of evidence support that *P. aeruginosa* cells have two distinct ClpP peptidase assemblies: $PaClpP1_{14}$ and $PaClpP1_7P2_7$. Importantly, peptidase and protease analyses establish that these two ClpP assemblies exhibit distinct peptide cleavage specificities and interact differentially with the AAA+ unfoldases, ClpX and ClpA. Finally, the PaClpP2 peptide-cleavage active site uniquely contributes to *P. aeruginosa* biofilm development. Therefore, results presented in this thesis suggest that within AAA+ proteases, the specificity of the peptidase subunits, not only the recognition properties of the AAA+ unfoldase, control the biological outcome(s) of proteolysis.

Thesis Supervisor: Tania A. Baker
Title: E. C. Whitehead Professor of Biology

Acknowledgements

There are many people who I would like to thank for helping me get to where I am today.

Foremost I would like to thank my thesis advisor, Tania Baker, for cultivating a welcoming and scientifically rigorous lab environment and for allowing me to flourish in that space. I am forever grateful for the ways in which Tania has supported me both scientifically and personally. As a graduate student, it is not easy to feel empowered at times, but Tania has always taken my scientific contributions and ideas seriously, and lifted me up when I was down. Tania, thank you for being a fantastic mentor and person.

I would not be where I am today without my undergraduate advisor, John Tamkun. My first lab experience was in John's lab as a college sophomore. It was through this experience working alongside then-post-doc Giorgia Siriaco that I decided to pursue graduate research. As a 19-year-old, I remember feeling thrilled and deeply moved every time I looked at fluorescently-labeled chromosomes of fruit flies through a confocal microscope (which was daily for three years). This thrill for science has not left me. John has graciously spent many hours giving me advice, encouragement, and guidance. Most of all, what kept me going in graduate school was John's belief in me. John, thank you for being my mentor and friend, in science and in life. Your presence in my life has been indispensable to both my personal and professional development.

Thank you to my fellow graduate students who have taught me so much. To Chikdu Shivalila and Albert Wu for teaching me how to clone. To Yarden Katz for advice and inspiration. To Mary Kay (MK) Thompson for advice, laughs, and for teaching me how to work with yeast. To my classmates Jennifer Cherone, Justin Chen, and Eleina England, thank you for your solidarity and friendship. You are forever friends. To Tristan Bell, who selflessly spent hours helping me study for my prelim, and for all of the scientific chats. Thank you.

There are countless post-docs to whom I am indebted for teaching me experiments, techniques, and ways of analyzing and thinking about data. In particular, I am grateful to have been bay-mates with Thomas Carlile and Julia Kardon. Thomas, thank you for your endless support, sharing of life hacks, and for teaching me so much about science and life. Your mentorship and friendship are invaluable. To Julia Kardon, thank you for teaching me how to do biochemical research, for instrumental advice and perspective, and also for sharing life hacks with me. Thank you for being a phenomenal role model of a strong woman in science.

To my best friends from college: Desirae Karmazin, Liza Pomerantz, Patricia Meier, Stephanie Pengilley. Thank you for sticking by my side through thick and thin, for your support and encouragement, for the adventures and laughter, for the tears and for the friendship. I look forward

to making more memories with you in the future. To Jason Hogin, my best friend from high school. Thank you for always knowing how to make me laugh, and for always making sure not too much time has passed since we last talked.

I must acknowledge Dr. Nina Geller for single-handedly pulling me out of a funk during my third year of graduate school. Thank you for all of the Shabbat dinners and for sharing many stories.

Joshua Goldford, thank you for your love, eternal friendship, and everlasting support. It has been an honor to go through graduate school together. You are a constant source of happiness to me. Thank you to Josh's parents, Mr. and Mrs. Goldford, for raising such a caring and genuine son.

Finally, thank you to my family: my parents, Shamil and Carol Mawla, and my sisters, Jacqueline and Pauline Mawla. Thank you for supporting me through the achievement of my dream, even though it meant me moving across the country for countless years to do so. To my sister, Pauline, for all of the heart-to-heart chats, encouragement, and food shipments. Especially, thank you to my mother, who made education a priority for her children.

Table of Contents

Chapter 1:

Introduction

Thesis overview and goals

This thesis describes my research towards the discovery and characterization of a new example of developmentally-regulated proteolysis in the opportunistic pathogen *Pseudomonas aeruginosa* (*P. aeruginosa*, Chapter 2). This experimental work is contextualized by current knowledge of protein metabolism, including pathways of protein synthesis, enzymes that modulate protein folding and conformation (molecular chaperones), and enzymes that destroy them (ATP-dependent proteases). Therefore, a brief background of these topics is provided, covering protein synthesis, folding and misfolding, degradation, well as the chaperones that assist proper folding *in vivo* (Chapter 1).

ATP-dependent intracellular protein degradation in all branches of life is carried out by enzymes of the AAA+ protease family. These macromolecular machines are composed of a AAA+ unfoldase enzyme coupled to a chamber-shaped peptidase. The AAA+ unfoldases are part of a much larger protein family that uses ATP hydrolysis to do work on other cellular constituents, are described in the first part of the introduction. The second part of the introduction provides a more extensive background on ClpP peptidases from the model system *E. coli* and various bacteria. The final part of the introduction describes the dynamic lifestyle and communication strategies of *P. aeruginosa*.

Cellular proteostasis

Overview of protein folding

All cells must actively establish and maintain their proteome to achieve proper form and function. The process by which cells control the composition, identity, concentration, and localization of their protein repertoire at any given time is called cellular proteostasis (Balch et al., 2008). The global network of pathways that shape the proteome include protein synthesis, protein folding, quaternary interactions, post-translational modifications, subcellular localization, and decay; these processes can be highly regulated (Belle et al., 2006; Alveraz-Castelao et al., 2012). Although proteins begin to fold co-translationally based on their intrinsic biochemical properties, many require the help of molecular chaperones to overcome folding challenges, especially in a crowded cellular environment. There is an ongoing balance between protein folding and protein aggregation, as proteins are thought to sample multiple regions of their free-energy landscape on the path to the native state. Molecular chaperones aid in the formation of the thermodynamically favorable native species by decreasing free-energy barriers and preventing aberrant intermolecular interactions (Figure 1-1, Anfinsen, 1973; Hartl et al., 2011).

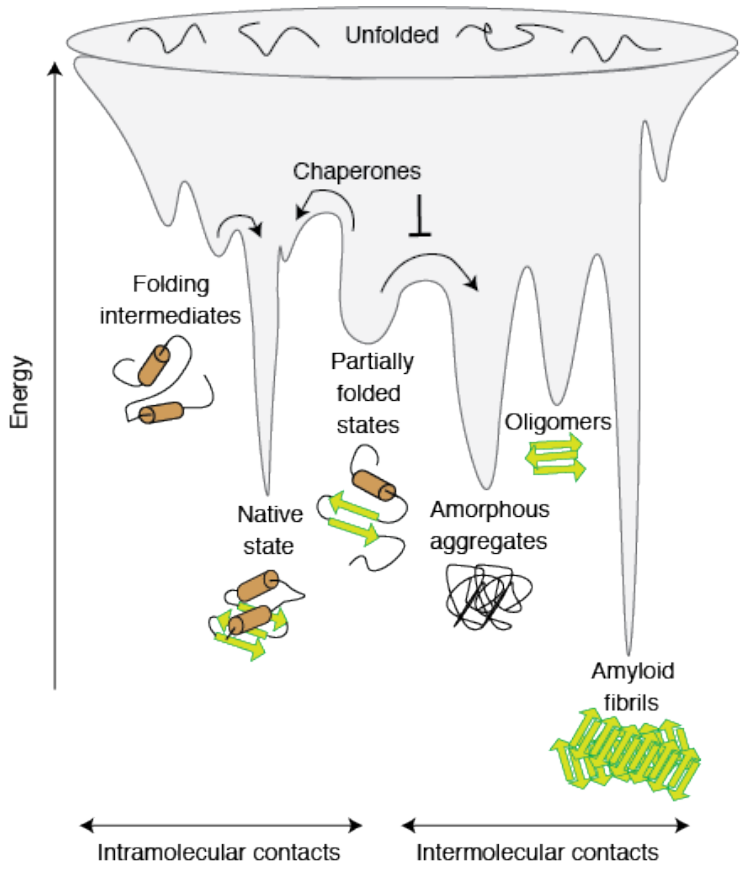

Figure 1-1: Protein folding and aggregation.
Folding of newly-synthesized polypeptides occurs by sampling various conformations and burying hydrophobic residues. Partially folded states can become kinetically trapped in localized energy minima. These intermediates are prone to aggregation through formation of undesired intermolecular contacts. Molecular chaperones promote the formation of the thermodynamically favorable native species by lowering free-energy barriers and preventing aberrant intermolecular interactions. Arrows represent the lowering of free-energy barriers by chaperones. Adapted from Hartl et al., (2011), Balchin et al., (2016) and Klaips et al., (2018).

Chaperones, unfoldases, and proteases monitor the proteome

Hsp70 and Hsp60 folding chaperones

Two classes of molecular chaperones that assist newly-synthesized proteins to fold into their native state are the heat shock protein Hsp70 and Hsp60 families (Hartl et al., 1992). The Hsp70 and Hsp60 families of molecular chaperones use two distinct mechanisms of action to promote proper protein folding. Generally, the Hsp70 and Hsp60 chaperone machines bind otherwise solvent-exposed hydrophobic regions on nascent proteins and assist the newly synthesized proteins in attaining the defined structure specified by their amino acid sequence. Through repeated cycles of binding and release of hydrophobic segments in substrate polypeptides, Hsp70 induces local conformational changes in the substrate that aid in protein folding (Rüdiger et al., 1997). In

contrast, Hsp60 proteins are large, cylindrical complexes that promote the proper folding of non-native polypeptides through a mechanism that involves internalization into a central chamber. By binding and releasing (Hsp70), or by binding and sequestering (Hsp60) misfolded polypeptides, these chaperones prevent the aggregation of non-native proteins and facilitate their proper folding. A prime example of an Hsp70 chaperone is the bacterial DnaK protein, which works alongside DnaJ (Hsp40) and the nucleotide-exchange factor, GrpE, whereas the best understood Hsp60 chaperonin is *E. coli* GroEL (Hsp60)/GroES (Hsp10) (Bukau and Horwich, 1998).

Hsp90 conformational chaperones

The Hsp90 family of molecular chaperones is not directly involved in the *de novo* folding of nascent proteins (Nathan et al., 1997). Instead of promoting early folding events, Hsp90 enzymes work as dimers with a variety of co-chaperones to remodel and activate proteins at the later stages in their folding pathway. Hsp90s function differently from Hsp60/Hsp70 in that Hsp90s hold onto client proteins and maintain their conformation in a poised state to accept an activation signal. Unlike Hsp60/70, which bind to unfolded polypeptides, Hsp90 enzymes bind to and maintain substrate proteins in a nearly-completely folded conformation. More than 200 Hsp90 client proteins have been identified in yeast and mammalian cell extracts. The client proteins have a broad array of cellular functions including protein kinases, steroid hormone receptors, and transcription factors (Li and Buchner 2012; Hartson et al., 2012). Together, Hsp60, Hsp70, and Hsp90 chaperones regulate the proteome by assisting in protein folding and/or maturation.

Clp/Hsp100 unfoldases, disaggregases, and proteases

The proteome is also maintained by enzymes that help clear the cell of damaged, aggregated or unwanted proteins by proteolysis. A Clp/Hsp100 ATPase coupled to a compartmentalized peptidase (both enzyme families are described in detail below) is responsible for protein turnover in bacteria and eukaryotic organelles. Additionally, some members of the Clp/Hsp100 family act peptidase-independently to promote the disaggregation and reactivation of aggregated proteins. For example, ClpB in *E. coli* and Hsp104 in *S. cerevisiae* resolubilize stress-induced protein aggregates either on their own or in conjunction with Hsp70/Hsp40 partners (Doyle and Wickner, 2008). Other Clp/Hsp100 proteins, like bacterial ClpX, ClpA/C, and HslU, principally work by linking their ATP-dependent unfoldase activity to the degradation activity of ClpP or HslV resulting in irreversible protein clearance (Sauer and Baker, 2011). Clp proteases are part of a larger superfamily and will be described in detail below.

AAA+ superfamily of ATPases

The AAA+ (ATPases associated with various cellular activities) superfamily is a large group of diverse enzymes that contain a homologous ATPase module and is conserved in all kingdoms of life. AAA+ ATPases use the chemical energy of ATP hydrolysis to perform mechanical work, and thus are key players in many cellular functions including DNA replication, recombination, proteolysis, cellular trafficking, protein disaggregation and disassembly (Table 1-1, Neuwald et al., 1999; Miller and Enemark, 2016; Sauer and Baker, 2001; Sousa, 2014; Monroe and Hill, 2016). As discussed below, the domain organization and mechanism of ATP hydrolysis share common features among many members of this family.

Function	AAA+ ATPase	Reference
Membrane fusion/transport	NSF/Sec18p	Zhao et al., 2015
Reconstitution of ER/Golgi	p97/VCP/Cdc48, NSF/Sec18p	Ye et al., 2017
Proteolysis	Proteosomal ATPases, FtsH, ClpA, ClpX, HslU, Lon	Sauer and Baker, 2011
Protein disaggregation/refolding	Hsp104/Hsp78/ClpB	Liberek et al., 2008
Molecular motors	Dyneins	Roberts et al., 2014
DNA replication	Orc1,4,5, Mcm2-7, Rfc1-5, SV40 T-antigen, DnaA	Duderstadt and Berger, 2008
DNA recombination/repair	RuvB, Rad17, Rfc2-5	Ye et al., 2004
Transcription	Tip49a.b, Ntrc	De Carlo et al., 2006
Mitosis/meiosis, apoptosis, severing of microtubules	Cdc48p/VCP, katanin/MEI-1	Meyer and Popp, 2008
Endosome function	Vps4p/SKD1	Yoshimori et al., 2000
Peroxisome function	Pex1p, Pex6p	Platta et al., 2008

Table 1-1: Various cellular activities of AAA+ ATPases
Adapted from Ogura and Wilkinson, 2001.

General structural features

The AAA+ family is part of the larger P-loop ATPase superfamily. AAA+ ATPases contain characteristic structural and functional features, including assembling into oligomeric rings with a central pore. The large domain of AAA+ proteins is largely composed of an α/β Rossman fold, a β-sheet of parallel strands in a β5−β1−β4−β3−β2 conformation. At the C-terminus, AAA+ proteins contain an α-helical bundle, called the small AAA+ domain (Iyer et al., 2014, Hanson and Whiteheart, 2005). Hallmark motifs of AAA+ proteins include the Walker A, Walker B, Sensor 1, and Sensor 2 motifs, in addition to a conserved Arg residue, known as an arginine finger, and a pore 1 loop (Figures 1-2 and 1-3).

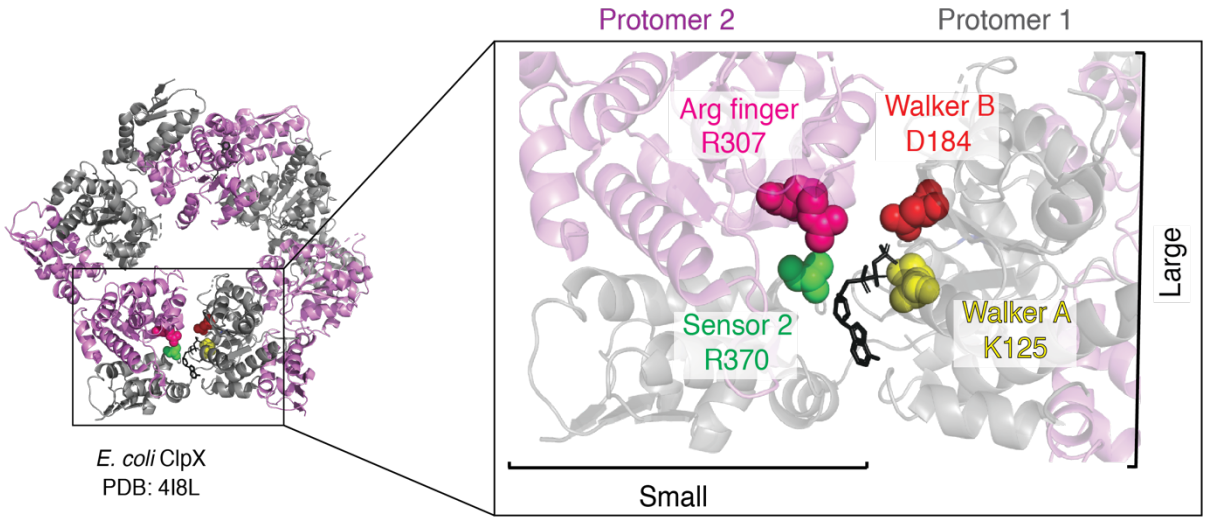

Figure 1-2: Model of ClpX showing the nucleotide binding site.
A close-up view of two adjacent monomers of *E. coli* ClpX (PDB: 4I8L; Stinson et al., 2013) is shown. Amino acid residues from conserved motifs are shown as spheres. The Walker A (yellow), and Walker B (red) motifs reside in the large AAA+ domain of protomer 1 (gray), whereas the sensor 2 motif (green) is located in the small AAA+ domain. The arginine finger (pink) is provided by the neighboring promoter 2 (violet). ATPγS is shown in black as a stick model. The sensor 1 motif is notably absent in ClpX but present in other AAA+ ATPases.

The Walker A and B motifs promote ATP binding and hydrolysis. The ATP-binding pocket that the Walker motifs help form is usually positioned at the interface between adjacent protomers in the hexamer (Figure 1-2). Residues of the Walker A, Walker B, and Sensor 1 motifs work together to coordinate Mg^{2+} and the phosphate groups of ATP and ADP. The Sensor 2 motif is located within the small domain, between α6 and α7, and is important for ATP hydrolysis, and to a lesser extent, ATP binding, likely by facilitating the movement of the C-domain relative to the N-domain of the AAA+ module during ATP hydrolysis. The arginine finger from one protomer makes up part of the nucleotide binding site on an adjacent protomer, and is thought to communicate ATPase activity from one subunit into the neighboring subunit of the hexamer. Finally, a conserved Tyr residue in the pore 1 loop plays key roles in substrate engagement and translocation (Siddiqui et

al., 2004). These regions provide the structural and functional attributes that define AAA+ ATPases into a superfamily of enzymes.

A.

Motif	Conserved residues	Function
Walker A	GXXXXGK[T/S]	Nucleotide binding
Walker B	ϕϕϕϕDE	ATP hydrolysis
Sensor 1	Polar residue	ATP hydrolysis
Sensor 2	GAR	ATP binding and hydrolysis
Arginine finger	R	ATP hydrolysis
Pore loop	YVG	Substrate binding and translocation

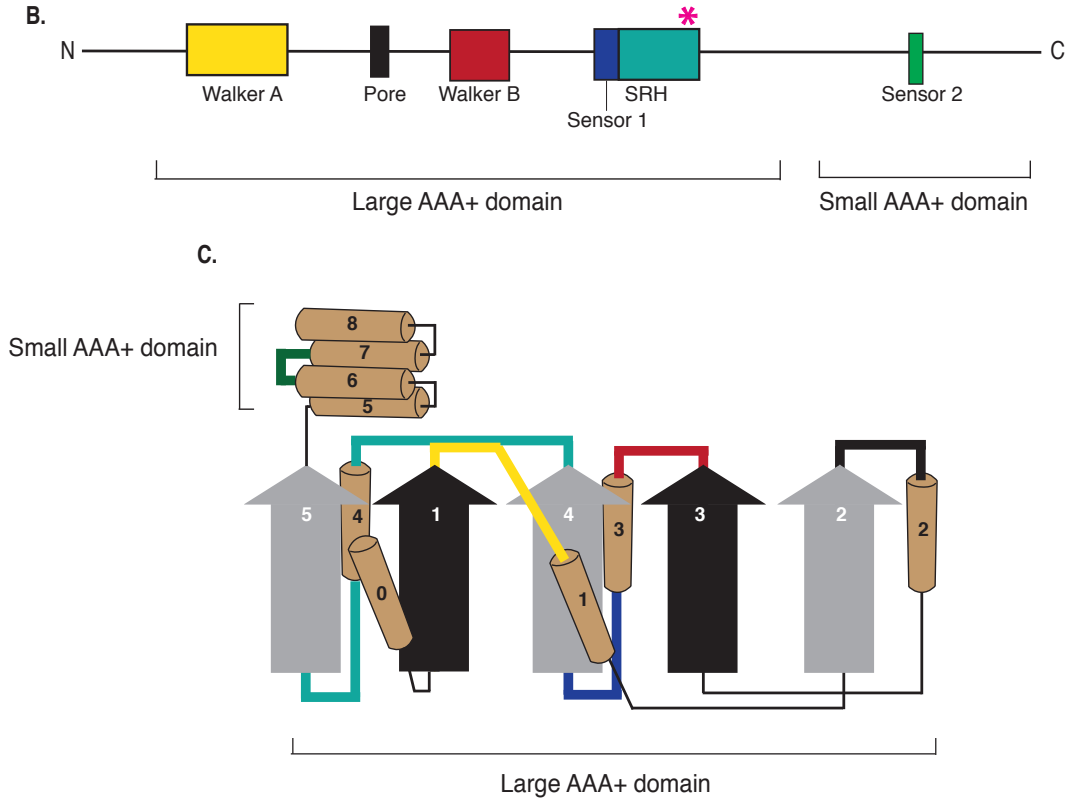

Figure 1-3: Secondary structure and characteristic elements of AAA+ ATPase proteins.
(A) A table of the conserved motifs of AAA+ proteins and their respective functions. X = any amino acid, ϕ = hydrophobic amino acid. B) Schematic of the relative location of key AAA+ motifs in the ATPase primary sequence. Pink asterisk indicates the Arg finger. C) Secondary structure schematic of AAA+ proteins. Coloring system is the same as in (A) and (B). Adapted from Hanson and Whiteheart, 2005.

AAA+ enzyme-powered reactions drive various cellular processes

AAA+ proteins are implicated in a wide range of cellular processes. Through cycles of ATP binding, hydrolysis, and ADP + Pi release, they are able to perform a variety of mechanical functions, such as unwinding, pulling apart, prying apart, or threading of substrates (Figure 1-4).

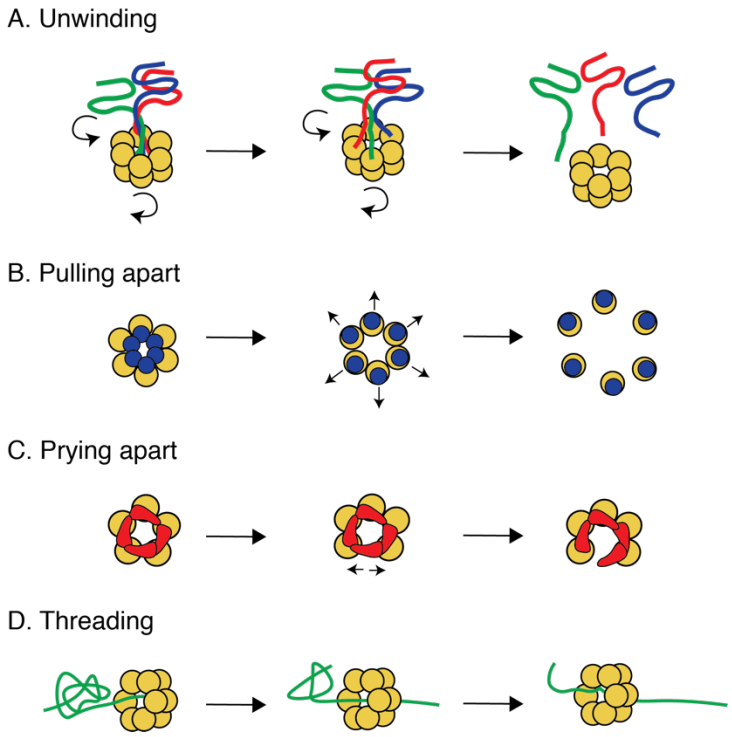

Figure 1-4: Examples of work done by AAA+ proteins.
Yellow circles represent AAA+ proteins. A) Twisting motions promote the disassembly of helical substrate bundles, as proposed for SNARE disassembly. B) Pulling apart mechanism, as proposed for microtubule disassembly. C) Prying apart of subunits can lead to partial ring opening, as is the case for DNA clamp loaders. D) Threading occurs by directional translocation, as seen by substrates of Clp ATPases. Adapted from Maurizi and Li, 2001.

To highlight the diversity of AAA+ remodeling enzymes in cellular reactions, two examples of non-proteolytic AAA+-catalyzed reactions will be discussed: the disassembly of SNARE complexes in eukaryotic vesicular transport, and the loading of the bacterial β–clamp during DNA replication.

Membrane fusion in eukaryotes is a fundamental final step in all vesicular trafficking events. Specialized proteins called SNAREs (soluble *N*-ethylmaleimide sensitive factor attachment protein receptors) are essential for membrane fusion, and must be disassembled into individual protein components following fusion reactions (Kaiser and Shekman, 1990). The disassembly of SNARE proteins is achieved by the AAA+ enzyme NSF (*N*-ethylmaleimide sensitive factor) with the help of SNAP (soluble NSF attachment proteins) cofactors. ATP hydrolysis on NSF provides the torque (unwinding force) required to separate SNARE complexes into their individual proteins (Figure 1-4A; Yoon and Munson, 2018; Zhao et al., 2015; Hong, 2005; Hay and Scheller, 1997).

Another example illustrating the broad cellular functions of AAA+ ATPases, is the role of the *E. coli* γ complex in ATP-dependent assembly of the β-clamp onto DNA during DNA replication. The γ complex is a pentameric AAA+ ATPase that loads the β sliding clamp (PCNA in eukaryotes) onto primed DNA templates in a process that first involves "prying apart" the β sliding clamp (Figure 1-4C). Binding of ATP triggers the conformational change of the γ complex and promotes the formation of a γ complex•open β ring•DNA intermediate, and subsequent ATP hydrolysis catalyzes ring closing around the DNA template and clamp loader release (Hingorani and O'Donnell, 1998; Kelch et al., 2012). Once loaded onto DNA, the sliding clamp tethers DNA Polymerase III to the template and enables the replicase to synthesize DNA at speeds greater than 500 nucleotides per second (Stukenberg et al., 1991; Stukenberg et al., 1994; Naktinis et al., 1994; Jeruzalmi et al., 2001).

AAA+ unfoldases in compartmentalized proteases

In all forms of life, the major pathways for intracellular protein degradation occur via energy-dependent proteolysis. The burden of protein degradation is shared by several large cylindrical proteases, including ClpAP, ClpXP, HslUV, Lon, FtsH in bacteria, PAN/20S in archaea, and the 26S proteasome in eukaryotes (Gottesman et al., 2003). The diverse group of enzymes share a general architecture that is critical for their function: they assemble into a set of stacked rings composed of AAA+ unfoldases paired with a chamber-shaped peptidase (Figure 1-5).

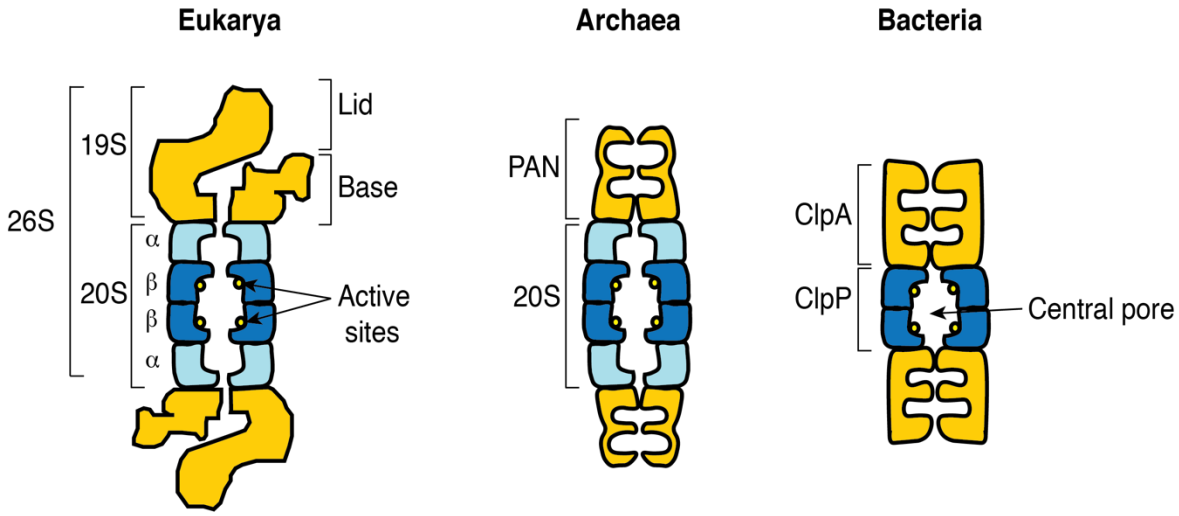

Figure 1-5: Representative AAA+ proteases from eukaryotes, archaea, and bacteria.
Schematic of the eukaryotic 26S proteasome, the archaeal PAN/20S proteasome, and ClpAP protease from bacteria. Note the shared structural similarity among proteases of all kingdoms of life. Unfoldases are colored in yellow and peptidase components are colored in blue. Adapted from Sorokin et al., 2009 and Striebel et al., 2009.

Eukaryotic and archaeal proteasomes

Eukaryotic proteasomes are present in both the nucleus and cytoplasm of eukaryotic cells. They are composed of a 20S core peptidase complex bound to one or two regulatory 19S particles (Fig. 1-5). As a complex, they make up the 26S proteasome. The 20S core consists of four stacked

heptameric rings, two composed of α subunits and two composed of β subunits; together these rings assemble in an $\alpha_7\beta_7\beta_7\alpha_7$ arrangement. Whereas the outer α rings provide structural support and interact with the 19S regulatory particle (RP), the polypeptide cleavage activity of the proteasome is derived from three subunits of the β rings: caspase-like β1, trypsin-like β2, and chymotrypsin-like β5 (Tamura et al., 1995, Tanaka, 2009). The 19S regulatory particle is composed of two subcomplexes: the base and the lid. The base contains four non-ATPase subunits (Rpn1-2, Rpn10, and Rpn 13) and six ATPase subunits (Rpt1-6) which is the AAA+ unfoldase of the 26S. The ATPase portion of the base directly contacts the 20S core particle. In contrast, the lid is a multicomponent scaffold that bridges the base to the proteasome's many substrates (Bard et al., 2018).

Protein substrates are flagged for proteolysis by the 26S proteasome via ubiquitination. Ubiquitination is the process by which the 76 amino acid ubiquitin protein is covalently attached to lysine residues on target proteins (Ciechanover and Schwartz, 1994). Although ubiquitin is necessary for proteolysis by the 26S proteasome, it is not sufficient. The substrate must also contain a loosely-folded region. There are three main steps in proteasomal degradation: (i) binding of ubiquitylated proteins, (ii) deubiquitylation, carried out by components of the lid, and (iii) ATP-dependent unfolding of the deubiquitylated substrate by the Rpt1-6 unfoldase of the deubiquitylated substrate and its translocation into the 20S chamber for destruction (Collins and Goldberg 2017). The initial binding of substrate involves recognition of both ubiquitin and a loosely-folded region within the substrate by 19S subunits Rpn1, Rpn10, and Rpn13 of the 19S complex. The 19S complex also contains deubiquitylation enzymes (DUBs), which remove ubiquitin chains from the substrate protein prior to translocation into the 20S proteolytic chamber.

The DUB subunit of the 19S complex, Rpn11, is positioned just above the substrate entry channel of the 20S chamber. This positioning provides a surveillance mechanism to ensure that ubiquitin chains are removed prior to entrance into the ATPase ring (Beck et al., 2012, Lander et al., 2012). This critical role for Rpn11 is supported by mutational studies showing that inactivation of Rpn11 DUB activity leads to a buildup of ubiquitylated proteins on the 26S complex (Matyskiela et al., 2013). Deubiquitylation of substrates on the proteasome is therefore an essential determinant of degradation.

Once substrates are deubiquitylated, the hetero-hexameric AAA+ Rpt1-6 ring triggers the widening of the 19S axial gate by docking C-terminal tails onto pockets on the α ring (Smith et al., 2007, Eisele et al., 2018). Cycles of ATP binding and hydrolysis stimulate translocation of substrates through the central channel into the core 20S particle to bring the unfolded substrate into close proximity to the peptidase's internal catalytic sites for cleavage (de la Peña, 2018). The evolutionary addition of many 19S subunits to a AAA+ unfoldase ring generated a single multi-functional complex that is competent to recognize substrates, recycle ubiquitin, translocate and degrade proteins.

Despite their overall similarity, there are key structural differences between eukaryotic and archaeal proteasomes. First, unlike eukaryotes, archaea do not contain a 19S base-lid complex, but instead contain a simple homo-hexameric AAA+ unfoldase, called PAN (proteasome-activating nucleotidase), which is thought to be the evolutionary precursor to the eukaryotic 19S particle (Zwickl et al., 1999; Barthelme et al., 2014). Second, archaea, most notably, lack both ubiquitin and ubiquitin-conjugating enzymes. This fundamental difference in substrate recognition

compared to eukaryotes alleviates the evolutionary need for DUB enzymes in the archaeal proteasome complex. Finally, whereas the eukaryotic proteasome contains seven distinct but homologous α subunits and seven distinct but homologous β subunits, the 20S proteasome from archaebacteria contains only one type of α subunit and one type of β subunit that form four stacked homo-heptameric rings; this lack of specialized peptidases is another indicator that archaeal proteasomes, although having clear similarities to the eukaryotic enzymes, are much simpler, being of similar complexity to the major bacterial AAA+ proteases (described below, Baumeister et al., 1998).

Bacterial AAA+ proteases

In bacteria, five major AAA+ protease sub-families carry out the bulk of cellular degradation: ClpXP, ClpAP, Lon, FtsH, and HslUV (Figure 1-6). Each family contains at least one AAA+ module, which consists of a large and small AAA+ domain. FtsH and Lon are distinct from the multi-subunit AAA+ proteases in that the protease domains and AAA+ modules are contained in a single polypeptide, and therefore are hexamers in their final functional form. The remaining protease sub-families have distinct AAA+ ATPase and protease subunits. The ATPase HslU forms a complex with the hexameric protease HslV, whereas the ClpX and ClpA/ClpC ATPases couple to the ClpP peptidase. In the case of Clp proteases, the coupling of hexameric ClpX, ClpA, or ClpC to the tetradecameric ClpP requires a symmetry mismatch question, as complex formation requires docking of a sixfold symmetric ATPase with the sevenfold symmetric ClpP (Grimaud et al., 1998).

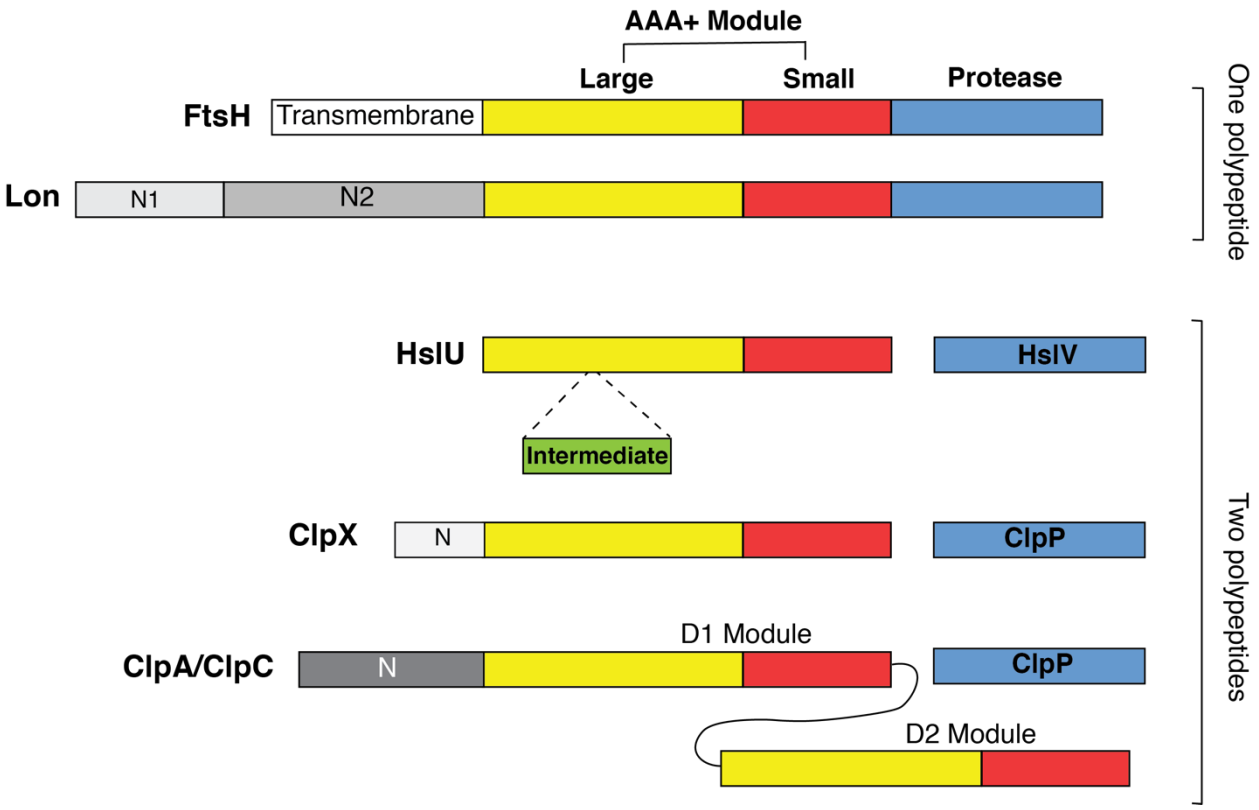

Figure 1-6: Domain architecture of bacterial AAA+ protease families.
FtsH, Lon, HslU, ClpX, and ClpA/ClpC share a homologous AAA+ module (small and large domains coupled to a compartmentalized peptidase (represented in blue) to undergo ATP-dependent proteolysis. Note that the ATPase and proteolytic domains of FtsH and Lon are encoded on a single polypeptide. ClpA/ClpC contain two AAA+ modules. N stands for N-terminal domain. Adapted from Sauer and Baker, 2011.

Unlike the eukaryotic and archaeal proteasomes, bacterial proteases (except in a few cases where bacteria 20S-like complexes exist, see below) do not contain structural alpha subunits; instead, the unfoldases bind directly to the proteolytically active subunits of the peptidase complex (Figure 1-5). Bacteria from the phylum Actinobacteria are distinct in that they encode a 20S peptidase and a AAA+ ATPase Mpa/Arc that contains the unfoldase activity. (Striebel et al., 2014). Whereas 20S peptidase complexes contain distinct β subunits with distinct classes of cleavage active sites (all β subunits perform a threonine-dependent nucleophilic attack; differences in cleavage specificity

arise from the amino acids that make up the S1 binding pocket), homomeric FtsH, Lon, HslV, and ClpP complexes each employ a single catalytic mechanism to achieve peptide bond hydrolysis. Interestingly, the catalytic mechanism of peptide bond hydrolysis is distinct between the different bacterial peptidase families. For example, the active site of the membrane-bound metalloprotease FtsH contains a catalytic Asp-Zn^{2+}, Lon contains a Ser-Lys dyad, HslV uses its N-terminal threonine as the active site residue, and ClpP employs a Ser-His-Asp catalytic triad.

Importantly, for enzyme function, there is a clear division of labor between the peptidases and the AAA+ unfoldases; substrate recognition is carried out by the unfoldase component, sometimes with the help of adaptor proteins. Bacterial unfoldases recognize specific peptide tags on substrates marked for degradation, termed "degrons." Degron-containing proteins that are destined for degradation fall into one of two classes: they are either damaged or irregular and must be removed by the cell, or they are natively (or conditionally) unstable and their turnover is part of a regulatory or developmental process (Mizusawa and Gottesman, 1983; Jenal and Fuchs, 1998; Meyer and Baker, 2011; Gur and Sauer, 2008; Moore and Sauer, 2011; Flynn et al., 2004; Zhou and Gottesman, 1998; Ingmer and Bronsdsted, 2009; Konovalova et al., 2014). Degrons are generally located at or near the N- or C- termini of proteins, but in some cases, are present in the interior of proteins and are then exposed by protein unfolding or processing (Hoskins et al 2002; Sauer and Baker 2001). The most well-studied C-terminal degron is the ssrA (small stable RNA A) tag. During abortive translation events, bacteria have evolved a ribosomal rescue system involving the co-translational appendage of the ssrA degron, a short amino acid sequence (11 amino acids in *E. coli*) (Keiler et al., 1996; Karzai et al., 2000). In *E. coli*, ClpXP is primarily responsible for the clearance of ssrA-tagged proteins, although ClpAP, FtsH, and Lon can also aid in their degradation

(Lies and Maurizi 2008; Levchenko et al., 2000; Farrell et al., 2005). The main function of the compartmentalized peptidase component of each enzyme family is to provide the catalytic activity required for proteolysis (described below).

ClpP peptidase

Key structural elements

ClpP (caseinolytic protease P) is a highly conserved serine peptidase present in all kingdoms of life. ClpP was first purified and characterized in *E. coli*, and *E. coli* ClpP (EcClpP) remains the best-studied model of the ClpP family (Katayama-Fujimura et al., 1987; Katayama et al., 1988). EcClpP is expressed as a proenzyme and undergoes autoproteolysis to remove the N-terminal 14 amino acids. The resulting product is the mature, 193 amino acid ClpP protein of ~21 kDa (Maurizi et al., 1990). ClpP forms a tetradecamer of two stacked heptameric rings with an internal proteolytic chamber of ~50Å in diameter. Entry into this chamber is gated by two axial entrance pores, which are ~10Å wide in the closed conformation. The ClpP monomer contains three subdomains known as the "handle," "head," and the N-terminal region (Figure 1-7). The handle region is composed of alpha helix E and beta sheet 9 and forms the area of interaction between the two heptameric rings. The globular head domain is largely invariable among ClpP orthologs and contains the catalytic triad resides: Ser97, His122, and Asp171 as well as the residues that make up the hydrophobic binding pocket where the unfoldases dock. Importantly, the active site residues are positioned in the interior of the barrel, while the unfoldase binding pockets are situated on the top of the barrel, near the entrance pores (Wang et al., 1997). Finally, the third key structural elements of ClpP are its N-terminal loops. These loops are highly flexible

and are often unresolved in ClpP crystal structures. The N-terminal loops of ClpP line the axial pore and serve as a "gating mechanism" for substrate entry; these loops can adopt either a

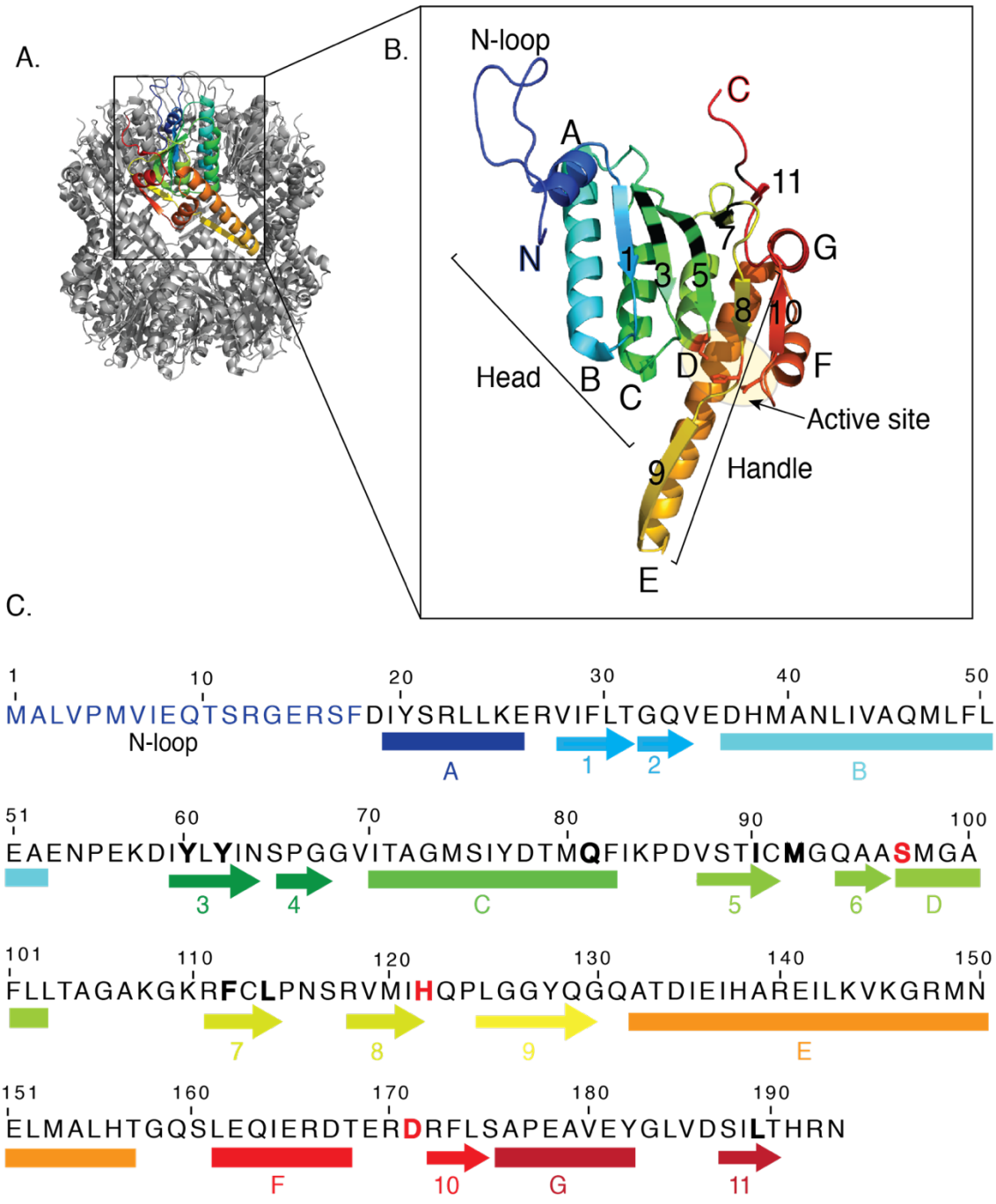

Figure 1-7: Structure of the ClpP monomer.
A) Crystal structure of *E. coli* $ClpP_{14}$ (PDB: 1YG6) with a single monomer colored in rainbow. B) Close-up of a ClpP monomer. Key domains and secondary structure nomenclature are indicated. C) Primary sequence of mature *E. coli* ClpP. Amino acids in red are the catalytic residues, those in black make up the

hydrophobic binding pocket for unfoldases. Secondary structure elements are colored as in (B). Adapted from Wang et al., 1997.

"closed" or "open" conformation. The importance of these loop elements is emphasized by deletion studies that demonstrate that removal of various N-terminal segments from ClpP results in a constitutively "open" ClpP that permits the unregulated degradation of unfolded proteins in the absence of a partner AAA+ unfoldase (ClpA or ClpX) (Bewley et al., 2009, Jennings et al., 2008).

Determinants in ClpP for unfoldase binding

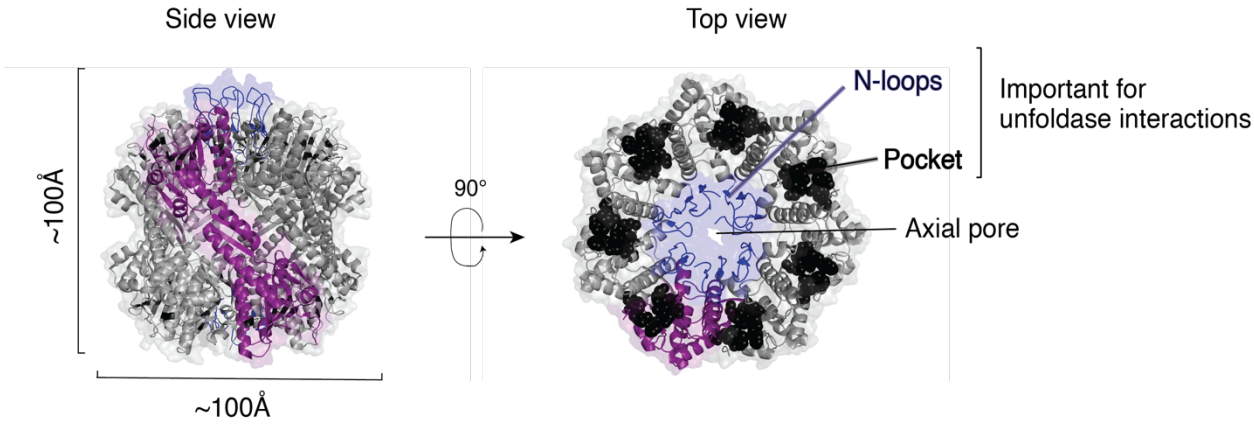

Figure 1-8: Key ClpP structural elements for binding to AAA+ unfoldases.
Side and top views of the *E. coli* ClpP tetradecamer are shown (PDB: 1YG6). In the left panel, one monomer from each heptamer is shown in purple. Note the region of interaction between heptameric rings. A 90º rotation reveals a top-down view of ClpP. The axial pore is gated by N loops (blue). The residues that make up the hydrophobic binding pockets for unfoldases are shown as black spheres.

On its own, ClpP has peptidase activity against short, unstructured peptides that can presumably diffuse into the central chamber via the N-terminally-gated axial pores. For the degradation of folded proteins, ClpP requires direct coupling to an unfoldase, namely, ClpX or ClpA (or ClpC in

some bacteria) to power ATP-dependent substrate unfolding and translocation. The two structural elements of ClpP that are critical for complex formation with unfoldases are (i) the hydrophobic binding pockets and (ii) the N-terminal loops. The binding pockets, located in the head region of the ClpP monomers (Figure 1-8), are the docking sites for the highly conserved ClpX and ClpA IGL/IGF loops (Kim et al., 2001; Joshi et al., 2004; Bewley et al., 2006). The N-terminal region of ClpP interacts with the pore-2 loops of unfoldases, which is important for complex formation, making a ring-like structure between the ATPase and ClpP that aligns the central pores of the two enzyme complexes (Figures 1-8, 1-9, Gribun et al., 2005, Bewley et al., 2006; Martin et al., 2007; Alexopoulos et al. 2012).

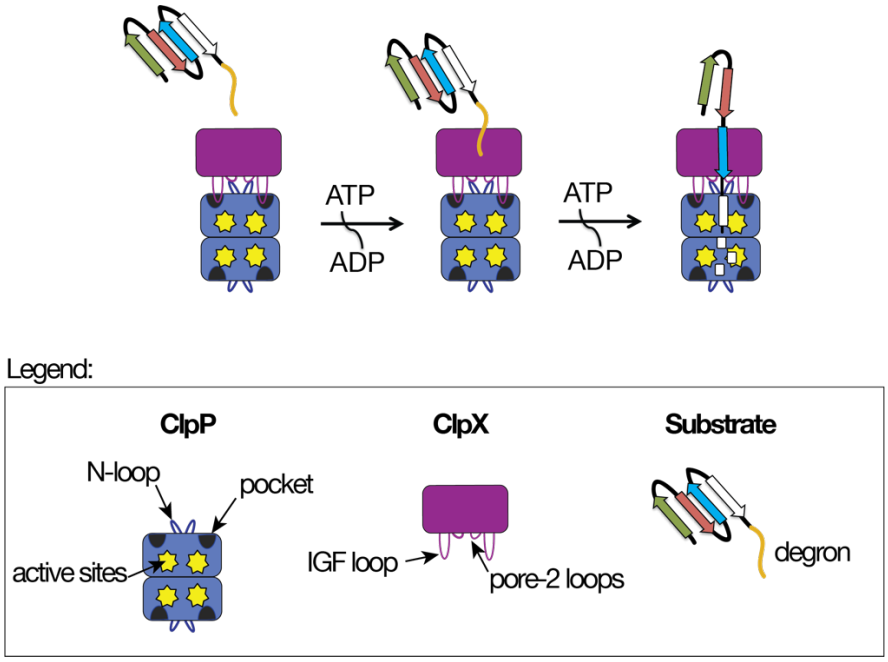

Figure 1-9: Schematic of ATP-dependent proteolysis by the ClpXP protease. Upon degron recognition, AAA+ unfoldases undergo cycles of ATP hydrolysis that leads to the unfolding and translocation of substrates into the central chamber of ClpP, where they are then degraded into shorter peptides. Key elements on ClpP, ClpX, and the substrate are indicated in the legend.

Roles of multiple *clpP* genes in bacteria

ClpP from *E. coli* has been the experimental paradigm to which all other ClpP isoforms are compared. It is a tetradecamer composed of 14 identical ~21 kDa protomers. Many organisms, like *E. coli*, encode a single ClpP variant. However, bacteria from some other specific phyla commonly

encode multiple ClpP isoforms. Some well-characterized examples include Cyanobacteria, Actinobacteria, Firmicutes, and Proteobacteria (Figure 1-10). In these organisms, ClpP tetradecamers can be found as homomeric or heteromeric assemblies. Representative examples of organisms with multiple ClpPs and their significance are discussed below.

Cyanobacteria

Cyanobacteria perform oxygenic photosynthesis and are thought to be the progenitors of plant chloroplasts (Martin et al., 1998). The model Cyanobacterium *Synechococcus elongatus* (*S. elongatus*) encodes three distinct ClpP paralogs: ClpP1, ClpP2, and ClpP3, and one ClpP-like protein, called ClpR. Despite having sequence similarity to ClpP, ClpR lacks the Ser-His-Asp catalytic triad and therefore is proteolytically inactive, analogous to some of the inactive β subunits in the eukaryotic 20S core peptidase. Two functional ClpP protease complexes form in *S. elongatus*, each with mixed subunits: ClpP1P2, and ClpP3R. In both complexes, the subunit arrangement within each heptameric ring is mixed. In the case of ClpP3R, the two stacked rings contain three ClpP3 and four ClpR subunits in an alternating arrangement (Andersson et al., 2009). The two ClpP heterocomplexes each interact with a different AAA+ unfoldase partner: ClpC binds to ClpP3R, and ClpX binds to ClpP1P2 (Stanne et al., 2007). The partitioning of distinct ATPase partners to distinct ClpP heterocomplexes is reflected in their genome organization. Whereas *clpP1* is located in a monocistronic operon, *clpP3* and *clpC* share an operon. Similarly, *clpP2* is in an operon with *clpX* (Schelin et al., 2002).

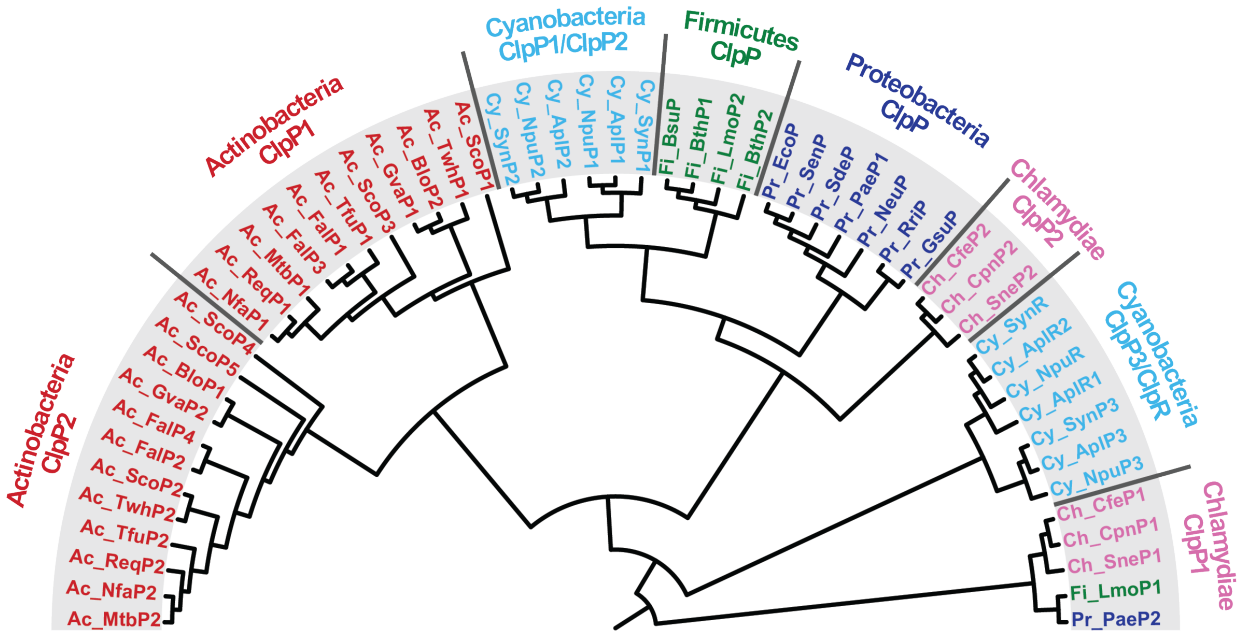

Figure 1-10: Phylogenetic tree of ClpP1 and ClpP2 peptidases from Actinobacteria, Cyanobacteria, Firmicutes, Proteobacteria, and Chlamydiae.
Phylogentic tree showing the evolutionary relationship between ClpP1 and ClpP2 peptidases from various clades. Adapted from Schmitz and Sauer, 2014.

Actinobacteria

Two or more ClpP paralogs are characteristic of the phyla Actinobacteria (Figure 1-9). The importance of the Clp system in the Actinobacteria *Mycobacterium tuberculosis* (*M. tuberculosis*) is highlighted by the fact that *clpP1*, *clpP2*, *clpX* and *clpC* genes are all essential for growth and virulence (Ollinger et al., 2012; Personne et al., 2013). *clpP1* and *clpP2* are co-transcribed in a single operon. The active form of ClpP in *M. tuberculosis* is a 14-mer composed of discrete ClpP1 and ClpP2 homo-heptameric rings. The AAA+ unfoldases ClpX and ClpC bind asymmetrically to the ClpP1P2 complex, interacting exclusively with the ClpP2 ring (Leodolter et al., 2015). Interestingly, the in *vitro* activity of ClpP1P2 requires the binding of N-blocked dipeptide activators (peptide agonists) to the peptidase active sites (Akopian et al., 2012; Schmitz et al.,

2014; Li et al., 2016). Recent cryo-EM and NMR studies have shown that binding of peptide agonists results in a concerted conformational change that appears to result in a more active conformation of ClpP1P2 (Vahidi et al., 2020).

Firmicutes

The Firmicute *Listeria monocytogenes* (*L. monocytogenes*) also expresses two ClpP proteins, ClpP1 and ClpP2. ClpP2 shares 66% homology with *E. coli* ClpP, whereas ClpP1 shares only 44%. ClpP2 has been shown to be an important stress-induced protein that promotes virulence in mouse models (Galliot et al., 2000). Conversely, the intracellular roles of ClpP1 are less understood. There are two functional ClpP assemblies in *L. monocytogenes*: ClpP2 homo-tetradecamers, and ClpP1P2 hetero-tetradecamers. The heterocomplex is made up of two discrete homo-heptameric rings of ClpP1 and ClpP2. On its own, ClpP1 is only partly assembled and inactive, and requires co-oligomerization with ClpP2 to gain activity. Interestingly, instead of a Ser-His-Asp catalytic triad, ClpP1 contains a Ser-His-Asn catalytic triad. Mutation of the active site Asn (residue 170) to Asp results in the assembly of active ClpP1 tetradecamers. In wild-type ClpP, it has been proposed that the Asp170 of one heptamer can make favorable contacts with Arg171 on the other heptamer. These residues are located in the ClpP handle region, in an area known as the oligomerization sensor. The Asp170-Arg171 favorable interaction may therefore trigger proper re-alignment of multiple residues that stabilize the tetradecamer (Zeiler et al., 2013; Gersch et al., 2012).

LmClpP1P2 interacts asymmetrically with the AAA+ unfoldase, ClpX, such that ClpX binds exclusively to the ClpP2 ring of the ClpP1P2 complex (Dahmen et al., 2015; Balogh et al., 2017,

Fux et al., 2019; Gatsogiannis et al., 2019). An apparently unique feature of the *L. monocytogenes* system is that it has been reported that ClpX binds to ClpP1P2 with a sevenfold tighter affinity than it does to ClpP2 homocomplexes (Balogh et al., 2017). Thus, heterocomplexes bound to ClpX experience a "boost" in proteolytic activity compared to their homomeric counterparts. It is unclear how shifting the intracellular equilibrium towards formation of ClpX•ClpP1P2 or this higher activity might impact proteostasis *in vivo*.

Phyla	Species	No. of ClpPs	Functional ClpP complexes	Subunit arrangement in heptameric rings	Interaction with unfoldases	Selected references
Cyanobacteria	*S. elongatus*	3, plus 1 related	$ClpP1_{6\text{ or }8}P2_{8\text{ or }6}$	Mixed	Binds to ClpX	Stanne et al., 2017
			$ClpP3_6R_8$	Mixed	Binds to ClpC	
Actinobacteria	*M. tuberculosis*	2	$ClpP1_7P2_7$	Pure	ClpP2 binds to ClpX/C	Leodolter et al., 2015
	M. smegmatis	2	$ClpP1_7P2_7$	Pure	ClpP2 binds to ClpX/C	Nagpal et al., 2019
Firmicutes	*L. monocytogenes*	2	$ClpP1_7P2_7$	Pure	ClpP2 binds to ClpX	Dahmen et al., 2015
			$ClpP2_{14}$	Pure	Binds to ClpX	
Chlamydiae	*C. trachomatis*	2	$ClpP1_7P2_7$	Pure	ClpP2 binds to ClpX	Pan et al., 2019
Proteobacteria	*E. coli*	1	$ClpP_{14}$	Pure	Binds to ClpX/A	Maurizi et al., 1990
	B. subtilus	1	$ClpP_{14}$	Pure	Binds to ClpX/C	Kock et al., 2004
	S. aureus	1	$ClpP_{14}$	Pure	Binds to ClpX/C	Gersch et al., 2011
	*P. aeruginosa**	2	$ClpP1_7P2_7$	Pure	ClpP1 binds to ClpX/A	This thesis
			$ClpP1_{14}$	Pure	Binds to ClpX/A	

Table 1-2: Table summarizing ClpPs from organisms with more than one isoform.
The number of ClpPs from various phyla are listed along with their subunit arrangement in functional complexes. Interaction(s) with specific unfoldases and selected references are also listed.

Pseudomonas aeruginosa: an opportunistic pathogen

The Gram-negative bacterium *P. aeruginosa* is an opportunistic pathogen that can survive in a broad range of natural and artificial settings, including soil, rivers, and surfaces in medical facilities. Because of its high intrinsic antibiotic resistance and ability to thrive in healthcare institutions, *P. aeruginosa* is one of the dominant pathogens causing airway infections of hospitalized patients. The success of this organism as a pathogen stems from its ability to produce a range of virulence factors, grow as a biofilm, and acquire adaptive resistance (Gellatly and

Hancock, 2013; Balasubramanian et al., 2013). Below I provide a brief introduction to *P. aeruginosa*, including overviews of quorum sensing, biofilm formation, and virulence strategies.

Quorum sensing

Communication among organisms in bacterial populations, dubbed quorum sensing (QS), is characterized by the secretion and recognition of small molecules within a bacterial community for the purpose of intercellular signaling. QS systems promote the coordination of bacterial group-like behaviors, such as regulating processes like biofilm development and virulence factor synthesis (Williams et al., 2000; Parsek et al., 2005). QS is exemplified by the observation that cell-free supernatants from high-density bacterial cultures are sufficient to cause gene expression changes when added to isogenic cells of the same strain grown at a lower density; thus mimicking the common phenotypes associated with QS (Fuqua et al., 1994). A large portion – up to 12% - of the *P. aeruginosa* transcriptome is regulated by quorum sensing (Schuster et al., 2003; Déziel et al., 2005). There are two acyl-homoserine lactone (AHL) QS systems in *P. aeruginosa*, the Las Inducer-Las Receptor (LasI-LasR) and Rhl Inducer-Rhl Receptor (RhlI-RhlR). LasI and RhlI stimulate the production of specific diffusible QS signals: *N*-3-oxo-dodecanoyl homoserine lactone (C12-HSL), and butyryl-HSL (C4-HSL), respectively. The binding of these diffusible molecules to cognate receptors (LasR or RhlR) activates QS-dependent changes in gene expression, including the upregulation of *lasI*, *rhlI*, and *rhlR*, due to the direct actions of the signal-bound LasR and RhlR transcription factors (Kostylev et al., 2019; Waters and Bassler, 2005). Accordingly, deletion of either *lasR* or *lasI* in *P. aeruginosa* abrogates the expression of QS-induced genes and causes diminished virulence in acute infection animal models (Tang et al., 1996; Rumbaugh et al., 1999).

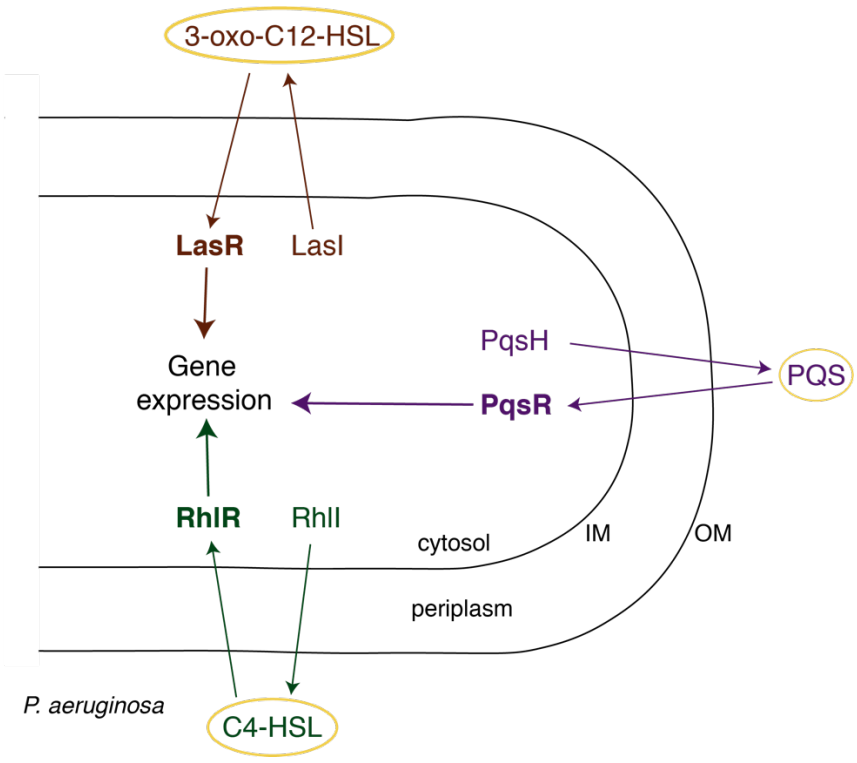

Figure 1-11: Quorum sensing systems in *P. aeruginosa*.
P. aeruginosa produces two homoserine lactone autoinducers, *N*-3-oxo-dodecanoyl homoserine lactone (3-oxo-C12-HSL), and butyryl-HSL (C4-HSL). These diffusible signals bind to the LasR and RhlR receptors, respectively. HSL binding to these receptors triggers their activation as transcription factors, leading to upregulation of genes involved in quorum-sensing are upregulated. A third system is also depicted. Unlike C12-HSL and C4-HSL, the Pseudomonas quinolone signal (PQS) is not membrane-diffusible and is transported via vesicles in and out of cells. Adapted from Schertzer et al., 2009. IM: Inner membrane. OM: Outer membrane.

The two AHL QS pathways interact with a third system that uses the signal 2-heptyl-3-hydroxy-4-quinolone (*Pseudomonas* quinolone signal (PQS)), and its biosynthetic enzyme, PqsH. Analogous to the AHL system, PQS binds to the transcriptional regulator PqsR, which leads to the expression of virulence genes. Altogether, these three QS systems are chiefly responsible for driving cell-density-dependent changes in gene expression that can lead to the coordination of group bacterial activities (Figure 1-11).

Biofilm formation and virulence strategies

QS is intricately-linked to virulence; the majority of genes upregulated by QS are either direct virulence factors or biosynthetic precursors of virulence factors (phenazines, proteases, toxins, rhamnolipids, hydrogen cyanide), genes that promote the motile-to-sessile switch (flagella, type 4 pili), or that regulate biofilm development (exopolysaccharides (Psl, Pel, and alginate), extracellular DNA (eDNA), lipids, cyclic di-GMP). Other QS genes encode factors that are involved in antibiotic resistance mechanisms or adjust metabolic pathways for stress responses (Venturi, 2006; Barr et al., 2015; Moradali et al., 2017). Thus, the formation of biofilms is one of the physiological outcomes of QS.

Biofilm formation involves the developmental transition from freely-swimming cells to surface-immobilized layers and occurs in three phases: attachment, maturation, and dispersion (Figure 1-12). Flagella are particularly important for attachment to surfaces, and type IV pili-mediated twitching is necessary for microcolony formation (O'Toole and Kolter, 1998). Once surface-attached, bacterial cells begin producing components that make up an extracellular polymeric substance (EPS) matrix (i.e.: polysaccharides, DNA, and proteins). In *P. aeruginosa*, alginate is an abundant EPS that contributes to the biofilm's mucoidity and pathogenicity (May et al., 1991). Biofilms grow and mature into thick "mushroom-like" shapes that consist of more than a hundred layers of bacterial cells. The cells are arranged in the mushroom-like structures according to their aerobic state and metabolism. For example, anaerobic bacteria avoid oxygen exposure by colonizing the deeper layers of the biofilm. Finally, when mature biofilms reach a certain density, signals are released to trigger bacteria dispersion into the environment which can seed new biofilms at other sites (Rabin et al., 2015).

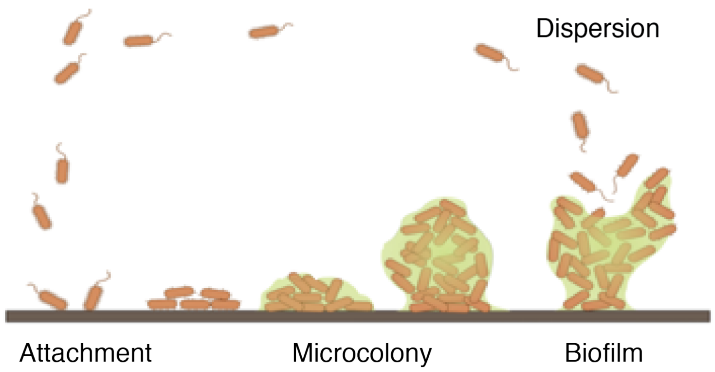

Figure 1-12: Biofilm formation in *P. aeruginosa*.
Biofilms develop when freely-swimming cells attach to surfaces and express extracellular polysaccharides that encapsulate the growing biofilm (green outline). After attachment, microcolonies form, followed by mature biofilms. Once mature biofilms reach a certain cell density threshold, cells undergo dispersion and are released from biofilms to populate other surfaces. Adapted from Maunders and Welch, 2017.

In this thesis, I explore the biophysical and biochemical characteristics of a new ClpP heterocomplex, ClpP1P2 from *P. aeruginosa* (PaClpP1$_7$P2$_7$, Chapter 2). First, I show that PaClpP1 and PaClpP2 physically interact in *P. aeruginosa* cell cultures. Next, I developed a strategy for the purification of PaClpP1$_7$P2$_7$ from *E. coli* overexpression cells, which lead to the discovery that the PaClpP1 and PaClpP2 active sites in the co-complex exhibit different polypeptide cleavage-site specificities, and that the AAA+ unfoldase PaClpX binds exclusively to the PaClpP1 face of PaClpP1$_7$P2$_7$. Additionally, I found that the serine active site residue of PaClpP2 within the heterocomplex is important for promoting biofilm thickness *in vivo*. These results point towards a function for PaClpP1$_7$P2$_7$ in the regulation of *P. aeruginosa* biofilm development and growth. In Chapter 3, I discuss the implications of my work and raise questions that will be important to address.

References

Akopian, T., Kandror, O., Raju, R. M., Unnikrishnan, M., Rubin, E. J., & Goldberg, A. L. (2012). The active ClpP protease from M. tuberculosis is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO Journal*, *31*(6), 1529–1541. https://doi.org/10.1038/emboj.2012.5

Arevalo-Ferro, C., Hentzer, M., Reil, G., Görg, A., Kjelleberg, S., Givskov, M., … Eberl, L. (2003). Identification of quorum-sensing regulated proteins in the opportunistic pathogen Pseudomonas aeruginosa by proteomics. *Environmental Microbiology*, *5*(12), 1350–1369. https://doi.org/10.1046/j.1462-2920.2003.00532.x

Arnold, I., Wagner-Ecker, M., Ansorge, W., & Langer, T. (2006). Evidence for a novel mitochondria-to-nucleus signalling pathway in respiring cells lacking i-AAA protease and the ABC-transporter Mdl1. *Gene*, *367*(1–2), 74–88. https://doi.org/10.1016/j.gene.2005.09.044

Augustin, S., Nolden, M., Müller, S., Hardt, O., Arnold, L., & Langer, T. (2005). Characterization of peptides released from mitochondria: Evidence for constant proteolysis and peptide efflux. *Journal of Biological Chemistry*, *280*(4), 2691–2699. https://doi.org/10.1074/jbc.M410609200

Baker, T. A., & Sauer, R. T. (2012). ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1823*(1), 15–28. https://doi.org/10.1016/j.bbamcr.2011.06.007

Chua, S. L., Hultqvist, L. D., Yuan, M., Rybtke, M., Nielsen, T. E., Givskov, M., … Yang, L. (2015). In vitro and in vivo generation and characterization of Pseudomonas aeruginosa biofilm-dispersed cells via c-di-GMP manipulation. *Nat Protoc*, *10*(8), 1165–1180. https://doi.org/10.1038/nprot.2015.067

E, V. (2014). *Microarray Analysis of Pseudomonas aeruginosa Quorum-Sensing Regulons : Effects of Growth Phase and Environment Microarray Analysis of Pseudomonas aeruginosa Quorum-Sensing Regulons : Effects of Growth Phase and Environment †*. *185*(7), 2080–2095. https://doi.org/10.1128/JB.185.7.2080

Ferrington, D. A., & Gregerson, D. S. (2012). Immunoproteasomes: Structure, function, and antigen presentation. In *Progress in Molecular Biology and Translational Science* (Vol. 109). https://doi.org/10.1016/B978-0-12-397863-9.00003-1

Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T., & Baker, T. A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Molecular Cell*, *11*(3), 671–683. https://doi.org/10.1016/S1097-2765(03)00060-1

Gersch, M., Stahl, M., Poreba, M., Dahmen, M., Dziedzic, A., Drag, M., & Sieber, S. A. (2016). Barrel-shaped ClpP Proteases Display Attenuated Cleavage Specificities. *ACS Chemical Biology*, *11*(2), 389–399. https://doi.org/10.1021/acschembio.5b00757

Gilbert, K. B., Kim, T. H., Gupta, R., Greenberg, E. P., & Schuster, M. (2009). Global position analysis of the Pseudomonas aeruginosa quorum-sensing transcription factor LasR. *Molecular Microbiology*, *73*(6), 1072–1085. https://doi.org/10.1111/j.1365-2958.2009.06832.x

Hall, B. M., Breidenstein, E. B. M., de la Fuente-Núñez, C., Reffuveille, F., Mawla, G. D., Hancock, R. E. W., & Baker, T. A. (2017). Two isoforms of Clp peptidase in Pseudomonas aeruginosa control distinct aspects of cellular physiology. *Journal of Bacteriology*, *199*(3), 1–15. https://doi.org/10.1128/JB.00568-16

Harris, J. L., Alper, P. B., Li, J., Rechsteiner, M., & Backes, B. J. (2001). Substrate specificity of the human proteasome. *Chemistry and Biology*, *8*(12), 1131–1141. https://doi.org/10.1016/S1074-5521(01)00080-1

Haynes, C. M., Petrova, K., Benedetti, C., Yang, Y., & Ron, D. (2007). ClpP Mediates Activation of a Mitochondrial Unfolded Protein Response in C. elegans. *Developmental Cell*, *13*(4), 467–480. https://doi.org/10.1016/j.devcel.2007.07.016

Haynes, C. M., Yang, Y., Blais, S. P., Neubert, T. A., & Ron, D. (2010). The Matrix Peptide Exporter HAF-1 Signals a Mitochondrial UPR by Activating the Transcription Factor ZC376.7 in C. elegans. *Molecular Cell*, *37*(4), 529–540. https://doi.org/10.1016/j.molcel.2010.01.015

Herget, M., & Tampé, R. (2007). Intracellular peptide transporters in human - Compartmentalization of the "peptidome." *Pflugers Archiv European Journal of Physiology*, *453*(5), 591–600. https://doi.org/10.1007/s00424-006-0083-4

Hong, Z., Bolard, A., Giraud, C., Prévost, S., Genta-Jouve, G., Deregnaucourt, C., … Li, Y. (2019). Azetidine-Containing Alkaloids Produced by a Quorum-Sensing Regulated Nonribosomal Peptide Synthetase Pathway in Pseudomonas aeruginosa. *Angewandte Chemie - International Edition*, *58*(10), 3178–3182. https://doi.org/10.1002/anie.201809981

Kim, Y. I., Burton, R. E., Burton, B. M., Sauer, R. T., & Baker, T. A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. *Molecular Cell*, *5*(4), 639–648. https://doi.org/10.1016/S1097-2765(00)80243-9

Kisselev, A. F., Garcia-Calvo, M., Overkleeft, H. S., Peterson, E., Pennington, M. W., Ploegh, H. L., … Goldberg, A. L. (2003). The caspase-like sites of proteasomes, their substrate specificity, new inhibitors and substrates, and allosteric interactions with the trypsin-like sites. *Journal of Biological Chemistry*, *278*(38), 35869–35877. https://doi.org/10.1074/jbc.M303725200

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., … Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, *47*(W1), W636–W641. https://doi.org/10.1093/nar/gkz268

Maunders, E., & Welch, M. (2017). Matrix exopolysaccharides; the sticky side of biofilm formation. *FEMS Microbiology Letters*, *364*(13), 1–10. https://doi.org/10.1093/femsle/fnx120

Melber, A., & Haynes, C. M. (2018). UPR mt regulation and output: A stress response mediated by mitochondrial-nuclear communication. *Cell Research*, *28*(3), 281–295. https://doi.org/10.1038/cr.2018.16

Patteson, J. B., Lescallette, A. R., & Li, B. (2019). Discovery and Biosynthesis of Azabicyclene, a Conserved Nonribosomal Peptide in Pseudomonas aeruginosa. *Organic Letters*, *21*(13), 4955–4959. https://doi.org/10.1021/acs.orglett.9b01383

Rutherford, S. T., & Bassler, B. L. (2012). Bacterial quorum sensing: Its role in virulence and possibilities for its control. *Cold Spring Harbor Perspectives in Medicine*, *2*(11), 1–25. https://doi.org/10.1101/cshperspect.a012427

Schmitza, K. R., Carneyb, D. W., Sellob, J. K., & Sauera, R. T. (2014). Crystal structure of mycobacterium tuberculosis ClpP1p2 suggests a model for peptidase activation by aaa+ partner binding and substrate delivery. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(43), E4587–E4595. https://doi.org/10.1073/pnas.1417120111

Schuster, M., Lostroh, C. P., Ogi, T., & Greenberg, E. P. (2003). (4916) Identification, Timing,

and Signal Speci city of. *Society*, *185*(7), 2066–2079. https://doi.org/10.1128/JB.185.7.2066
Sheps, J. A., Ralph, S., Zhao, Z., Baillie, D. L., & Ling, V. (2004). The ABC transporter gene family of Caenorhabditis elegans has implications for the evolutionary dynamics of multidrug resistance in eukaryotes. *Genome Biology*, *5*(3), 1–17. https://doi.org/10.1186/gb-2004-5-3-r15
St-Pierre, C., Morgand, E., Benhammadi, M., Rouette, A., Hardy, M. P., Gaboury, L., & Perreault, C. (2017). Immunoproteasomes Control the Homeostasis of Medullary Thymic Epithelial Cells by Alleviating Proteotoxic Stress. *Cell Reports*, *21*(9), 2558–2570. https://doi.org/10.1016/j.celrep.2017.10.121
Thompson, M. W., & Maurizi, M. R. (1994). Activity and specificity of Escherichia coli ClpAP protease in cleaving model peptide substrates. *Journal of Biological Chemistry*, *269*(27), 18201–18208.
Vahidi, S., Ripstein, Z. A., Juravsky, J. B., Rennella, E., Goldberg, A. L., Mittermaier, A. K., … Kay, L. E. (2020). An allosteric switch regulates Mycobacterium tuberculosis ClpP1P2 protease function as established by cryo-EM and methyl-TROSY NMR. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(11), 5895–5906. https://doi.org/10.1073/pnas.1921630117
Yewdell, J. W. (2005). Immunoproteasomes: Regulating the regulator. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(26), 9089–9090. https://doi.org/10.1073/pnas.0504018102
Young, L., Leonhard, K., Tatsuta, T., Trowsdale, J., & Langer, T. (2001). Role of the ABC transporter Mdl1 in peptide export from mitochondria. *Science*, *291*(5511), 2135–2138. https://doi.org/10.1126/science.1056957
Yu, A. Y. H., & Houry, W. A. (2007). ClpP: A distinctive family of cylindrical energy-dependent serine proteases. *FEBS Letters*, *581*(19), 3749–3757. https://doi.org/10.1016/j.febslet.2007.04.076

Chapter 2:

ClpP1P2 peptidase activity promotes biofilm formation in *P. aeruginosa*

This chapter has been written as a manuscript for publication. A draft is currently being reviewed by collaborators. I performed all the experiments reported in this chapter, with the exception of Figure 2-11, which was contributed by G. Cárcamo-Oyarce. J. Zhang assisted with peptidase/protease experiments in Figures 2-12 and 2-13. The crystal structures of *P. aeruginosa* ClpP1 and ClpP2 were solved by R. Grant and B. Hall.

Summary

Clp proteases, containing a AAA+ unfoldase stacked with a compartmentalized peptidase, are central to bacterial proteolysis, control cellular physiology and regulate stress responses. The opportunistic pathogen *P. aeruginosa* is unusual in that it contains two isoforms of ClpP peptidase: PaClpP1 and PaClpP2; the specific functions of the PaClpP isoforms are, however, largely elusive. Here, we report that the active form of PaClpP2 is a part of a heteromeric $PaClpP1_7P2_7$ tetradecamer, and that this enzyme is required for proper biofilm development. $PaClpP1_{14}$ and $PaClpP1_7P2_7$ complexes exhibit distinct peptide cleavage specificities and interact differentially with the *P. aeruginosa* AAA+ unfoldases, PaClpX and PaClpA. Crystal structures reveal that PaClpP2 has non-canonical features in its N- and C-terminal regions that explain its poor interaction with unfoldases. However, *in vivo* experiments indicate that the PaClpP2 peptide-cleavage active site uniquely contributes to biofilm development. These data strongly suggest that the specificity of PaClpP, not only that of the AAA+ unfoldase, can contribute to the biological outcome of proteolysis. Furthermore, by defining this unique role of PaClpP2, we highlight it as an attractive target for antimicrobial agents that could interfere specifically with late-stage *P. aeruginosa* biofilm development.

Introduction

Regulated protein degradation is central to bacterial physiology and development, playing critical roles in establishing and maintaining appropriate levels of intracellular proteins. AAA+ proteases of the Clp family are macromolecular complexes conserved in bacteria and in specific organelles in eukaryotes. Clp protease substrates include proteins that regulate cellular stress responses, cell division, antibiotic resistance, motility, and virulence (Gur et al., 2011; Culp and Wright, 2017). The Clp proteases are composed of a hexameric AAA+ (ATPases associated with diverse cellular activities) unfoldase (ClpA or ClpX in *E. coli* and *P. aeruginosa*) complexed with a self-compartmentalized tetradecameric peptidase, ClpP. Substrates destined for degradation are recognized and unfolded in an energy-dependent process by the unfoldase and then translocated into the central chamber of ClpP where they are degraded into short peptides (Sauer and Baker, 2011; Baker and Sauer, 2012) (Fig. 2-1 A,B).

Most proteobacteria, such as *E. coli*, contain a single *clpP* gene, which is transcribed, translated and processed to remove a pro-peptide, before subunits assemble into a functional homomeric tetradecamer. However, organisms from phyla including *actinobacteria*, *cyanobacteria*, *firmicutes*, and *chlamydiae* frequently harbor two or more *clpP* genes that encode subunits that in some cases assemble into active heteromeric tetradecamers (Akopian et al., 2012; Stanne et al., 2007; Dahmen et al., 2015; Pan et al., 2019). The specific biological advantage(s) provided by heteromeric ClpP complexes, however, is not well-understood. Here we investigate the assembly, structures, activities, and *in vivo* roles of ClpP1 and ClpP2 from *P. aeruginosa* and gain insight into potential benefits that having two ClpP isoforms may impart to an organism.

Clp proteases achieve specificity for their protein substrates by their associated AAA+ unfoldase; this unfoldase recognizes specific degron tags (short peptide sequences) on cognate substrates. Unfoldases in proteolytic complexes are thus multifunctional, being responsible for substrate recognition, unfolding, and translocation of unfolded substrate into ClpP (Baker and Sauer 2012). Importantly, within unfoldase•ClpP complexes, the ClpP axial pores expand to allow substrate entry. Binding of unfoldases to ClpP involves two key interactions: (i) docking of the IGF/IGL loops of ClpX/A within specific hydrophobic pockets on the ClpP surface, and (ii) establishment of contacts between ClpP N-terminal axial loops and pore-2 loops of the unfoldase (Kim et al., 2001; Joshi et al., 2004; Bewley et al., 2006; Gribun et al., 2005; Alexopoulos et al., 2012). In the chamber, unfolded proteins are cleaved into peptides of ~8-10 amino acids by ClpP's 14 Ser-His-Asp active sites (Thompson et al., 1994). Thus, through coordinated action, the unfoldase•peptidase complexes choose substrates, unfold them, and cleave them into short peptides (again, Fig. 2-1 A,B).

The opportunistic pathogen *P. aeruginosa* is a Gram-negative bacterium that can cause serious chronic infections in immunocompromised patients and is a major factor in the mortality of patients with cystic fibrosis (Lyczak et al., 2000; Gellatly and Hancock, 2013). The bacterium's ability to switch from free-swimming planktonic cells to surface-attached biofilm communities contributes to both antibiotic resistance and pathogenicity (Drenkard and Ausubel, 2002; Mah et al., 2003). *P. aeruginosa* contains two *clpP* genes located at distinct genomic loci: *clpP1* (PA1801) is adjacent to *clpX*, as is common in many bacteria, whereas *clpP2* (PA3326) is at a different locus within a cluster of genes that appear to form a mobile genetic island. The *clpP1* and *clpP2* genes also exhibit distinct expression patterns. The *clpP1* gene is expressed throughout growth whereas

clpP2 is tightly controlled by quorum sensing machinery, including the LasR transcription factor (Wagner et al., 2003; Schuster et al., 2003; Arevalo-Ferro et al., 2003; Hentzer et al., 2003; Gilbert et al., 2009). Genetic studies implicate both *clpP1* and *clpP2* in virulence, and deletion of either gene results in distinct cellular phenotypes (Qiu et al., 2008, Hall et al., 2017). Initial characterization of the purified *P. aeruginosa* ClpP isoforms reveals ClpP1 to be a homo-tetradecamer that is active both as a stand-alone peptidase and as a protease when combined with PaClpX or PaClpA. ClpP2, however, purifies as a heptamer with no observed catalytic activity, and thus its active form has remained mysterious (Hall et al., 2017).

Here, we establish that PaClpP2 is active within a heterocomplex with PaClpP1. Using biochemical, crystallographic, and microbiological assays, we provide evidence that PaClpP1P2 hetero- and PaClpP1 homocomplexes have unique peptidase and protease activities, and form different protein-protein complexes, and, furthermore, PaClpP2 peptidase activity is specifically required for assembly of robust, thick biofilms. These results are the first to report the existence and biological significance of a PaClpP1P2 hetero-tetradecamer in *P. aeruginosa*. We discuss models for how PaClpP1 and PaClpP2 contribute to pathogenicity and poor patient prognosis during *P. aeruginosa* infections.

Results

P. aeruginosa ClpP1 and ClpP2 form a heteromeric complex *in vivo.*

In a previous study characterizing the two ClpP isoforms from *P. aeruginosa*, (genes *clpP1* (PA1801) and *clpP2* (PA3326)), PaClpP1 and PaClpP2 were purified from *E. coli* expression strains. Initial results revealed that, whereas PaClpP1 purifies as an active tetradecamer, PaClpP2 is an inactive heptamer (Hall et al., 2017). Thus, we sought to determine the active form of PaClpP2 in *P. aeruginosa* cells using a strategy to probe for native protein-protein complexes.

One hypothesis was that PaClpP2 exists in a co-complex with PaClpP1, as similar heterocomplexes have been characterized from other bacterial species that carry multiple *clpP* genes (Akopian et al., 2012; Stanne et al., 2007; Dahmen et al., 2015; Pan et al., 2019). To probe for interactions between PaClpP1and PaClpP2 in *P. aeruginosa* strain PAO1, chromosomally-encoded *clpP2-flag* was introduced to enable rapid PaClpP2 pull-down experiments from fresh cell extracts with anti-flag antibodies. The cellular extract was made from stationary-phase PAO1 and captured protein was observed by stained SDS-PAGE (Fig. 2-1C). Two similarly sized protein bands were evident, one that co-migrated with PaClpP2-flag, and the other that co-migrated with native PaClpP1.The identity of PaClpP1 was verified by Western blot using antibody specific for PaClpP1 and mass spectrometry (Fig. 2-1D and Fig. 3-3). Thus, we conclude that PaClpP1 and PaClpP2 interact *in vivo* in *P. aeruginosa*.

To purify larger amounts of the PaClpP1P2 co-complex for biochemical analysis, a dual expression plasmid encoding both affinity-tagged PaClpP1 (His_6) and PaClpP2 (StrepII) was used for co-expression in *E. coli*. The co-complex was purified using tandem Ni^{2+}-NTA and streptavidin

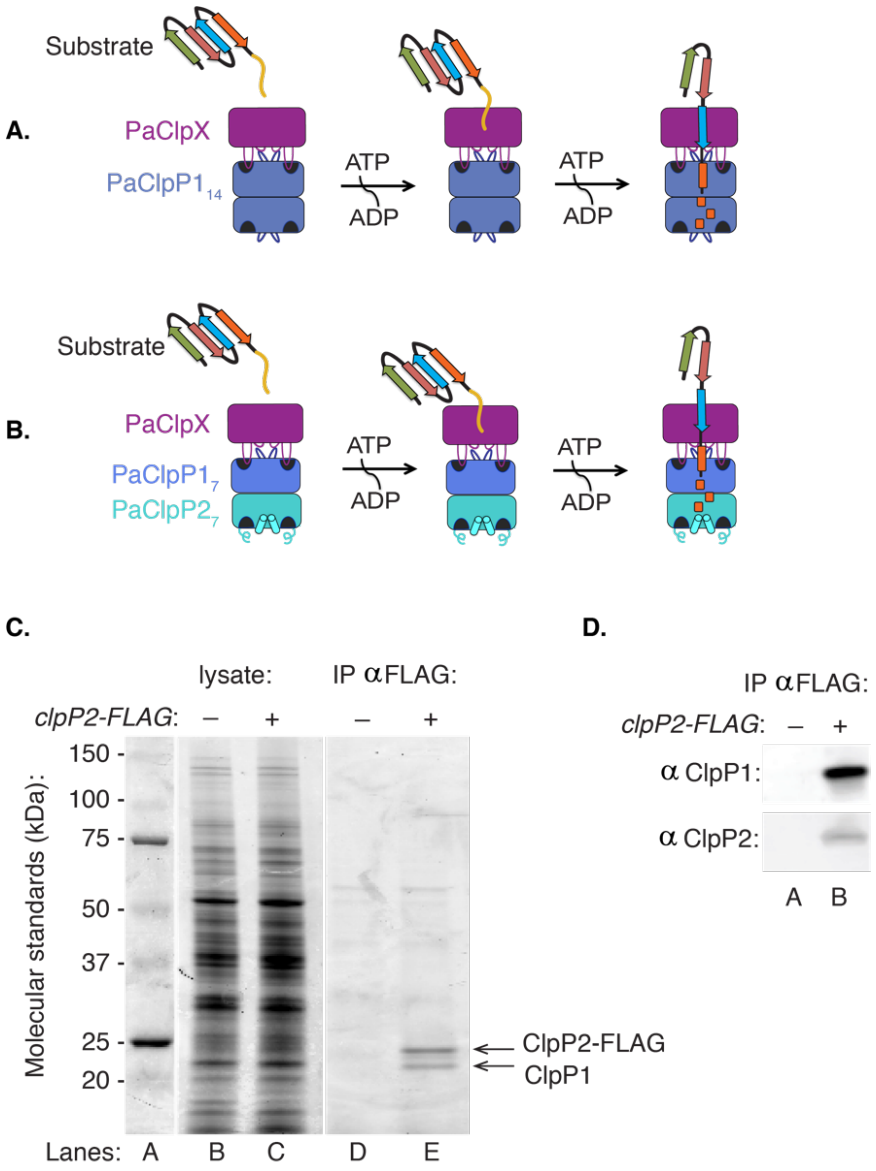

Figure 2-1: PaClpP1 and PaClpP2 co-immunoprecipate from PAO1 cells.
A) Schematic of ATP-dependent proteolysis by PaClpXP114. PaClpX is shown in purple, PaClpP114 is shown in blue. B) Schematic of ATP-dependent proteolysis by PaClpP1P27. Color scheme is the same as in A), except PaClpP2 is shown in cyan. C) ClpP2-FLAG was isolated from isolated from cell lysate from PAO1 strains containing either *clpP2-1xFLAG* (lane C, +) or untagged clpP2 (lane B, -) alleles at the genomic loci using anti-FLAG antibody-conjugated beads and elution with 3xFLAG peptide (Sigma-Aldrich). Eluate was separated by SDS-PAGE and proteins were stained with SYPRO orange (lanes D and E, Sigma-Aldrich). Molecular weight standards are shown in lane A. Molecular weights of monomers: ClpP2-FLAG: 23,237 Da. ClpP1: 21,879 Da. D) The eluate was analyzed by Western blot for PaClpP1 (1:15,000 anti-PaClp1) and PaClpP2 (1:7,500 anti-PaClpP2). Secondary antibody: 1:10,000 goat anti-rabbit.

columns, followed by size-exclusion chromatography (Fig. 2-2). The final column yielded a single peak, which eluted at the size expected for a tetradecamer (~320 kDa) and contained both PaClpP1-His_6 and PaClpP2-StrepII (Fig. 2-3). These peak fractions were pooled and concentrated. Quantification of bands in stained gels of the purified co-complex indicated that the two ClpP isoforms were present in close to a 1:1 ratio. Importantly, formation of PaClpP1P2 complexes strictly depended upon co-expression and co-purification, as separately purified PaClpP1 and PaClpP2 did not mix to form heterocomplexes (Hall et al., 2017). Previous data also established that PaClpP1 elutes at ~320 kDa upon gel filtration, consistent with a fully assembled tetradecamer, whereas PaClpP2 elutes at ~160 kDa, consistent with a heptamer (Hall et al., 2017). However, control experiments established that these two homomeric complexes were separated from the co-complex by the tandem affinity steps of our new purification method (Fig. 2-2).

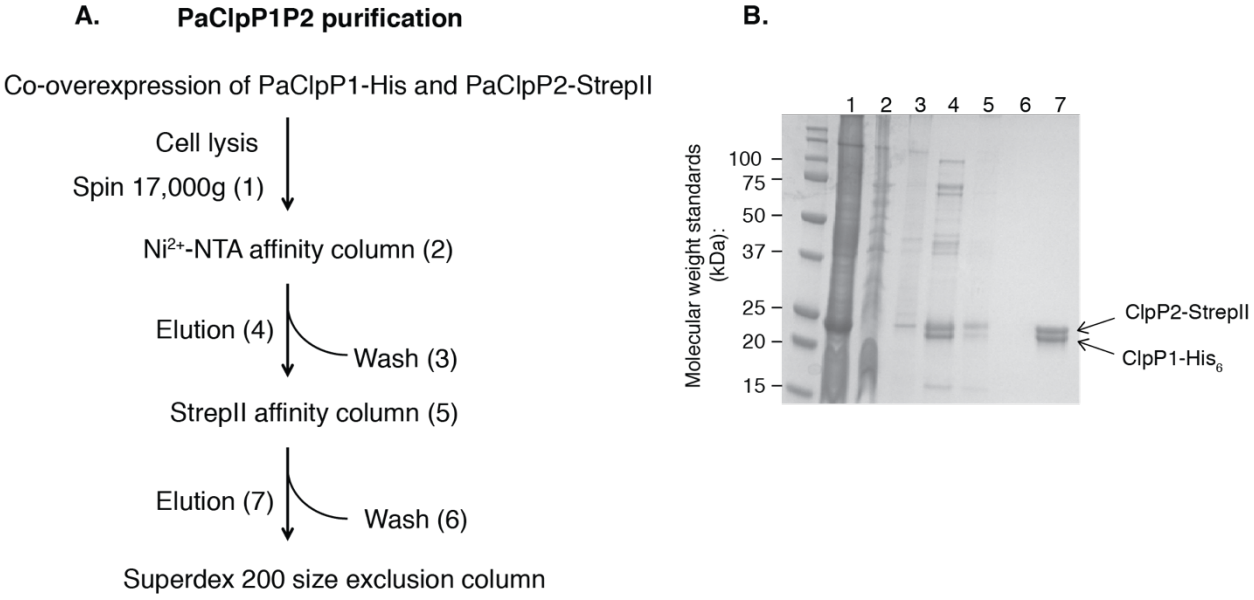

Figure 2-2: PaClpP1$_7$P2$_7$ purification workflow.
A) Overview of PaClpP1$_7$P2$_7$ purification (refer to Methods for details). PaClpP1P2 was purified from *E. coli* cells co-overexpressing PaClpP1-His and PaClpP2-StrepII. Briefly, cell lysate was applied to tandem affinity columns and PaClpP1$_7$P2$_7$ co-complexes were further purified by gel filtration chromatography.

Numbers refer to purification intermediates analyzed by SDS-PAGE in (B). B) SDS-PAGE analysis of purification intermediates. Lanes on the gel are numbered according to the purification workflow described in (A). Specifically, (1): supernatant (2): flow-through (3): wash (4): elution (5): flow-through (6): wash (7): elution. Calculated molecular weights of monomers are: ClpP1-6xHis: 23,063 Da (mature). ClpP2-StrepII: 23,326 Da.

To explore the subunit arrangement in PaClpP1P2 complexes we carried out a time course of glutaraldehyde crosslinking with three separate samples: (i) PaClpP1$_{14}$, (ii) PaClpP2$_7$ (inactive, see above) and (iii) freshly purified PaClpP1P2 complexes. Crosslinked samples were then analyzed by coomassie-stained SDS-PAGE. Time-dependent crosslinking of PaClpP1P2 generated a gel a pattern that looked like the sum of the two gel patterns observed for crosslinked PaClpP1$_{14}$ (rapid but incomplete dimer formation) and crosslinked PaClpP2$_7$ (a ladder of cross-linked species that continually formed over ~15 min) (Fig. 2-4). These data, together with the apparent 1 to 1 stoichiometry of the two ClpP isoforms in the ~320 kDa complex (see Fig. 2-1, above), support a configuration of the hetero-complex composed of two stacked homo-heptameric rings: one of PaClpP1 and one of PaClpP2 (calculated MW of 325 kDa); hereafter this complex is referred to as PaClpP1$_7$P2$_7$. Thus, we conclude that *P. aeruginosa* makes two forms of ClpP tetradecamers during its life cycle: PaClpP1$_{14}$ and PaClpP1$_7$P2$_7$. Furthermore, as described below, both PaClpP1$_{14}$ and PaClpP1$_7$P2$_7$ are active peptidases on their own, and, when coupled with PaClpX or PaClpA, they each form ATP-dependent proteases *in vitro*, and very likely *in vivo*.

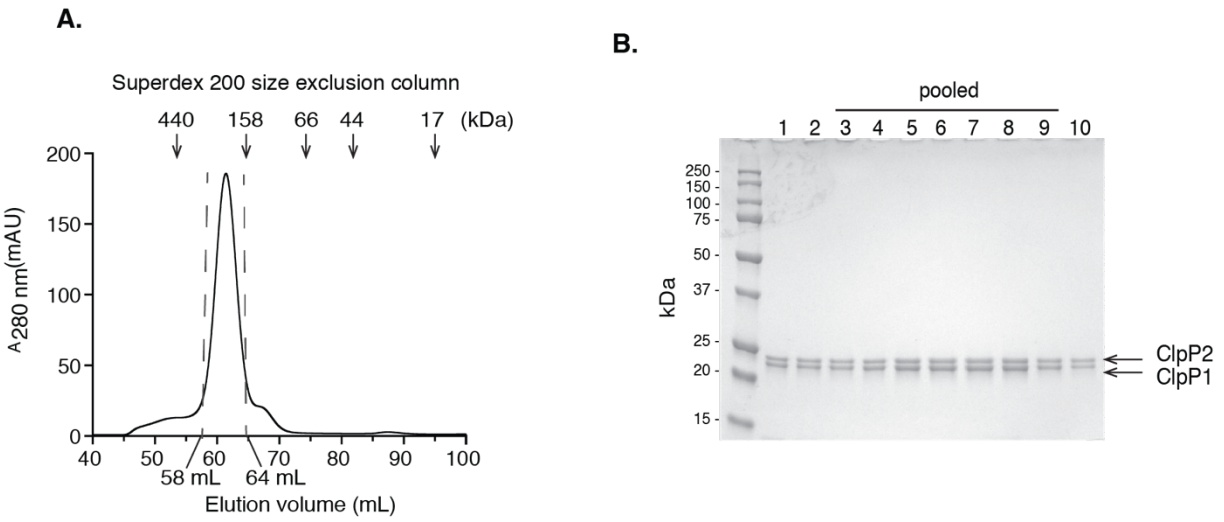

Figure 2-3: PaClpP1$_7$P2$_7$ is a tetradecamer.
A) Gel filtration profile of PaClpP1$_7$P2$_7$. Peak fractions corresponding to tetradecamers (58-64 mL, as indicated by dotted lines) were pooled and concentrated for biochemical analysis. B) ClpP1 and ClpP2 co-associate during purification. SDS-PAGE of elution fractions. Lanes 1-10 correspond to samples taken from 1 mL fractions eluting from 56-66 mL. Elution fractions corresponding to lanes 3-9 were pooled and concentrated for further characterization. Molecular weights of monomers: ClpP1-6xHis: 23,063 Da. ClpP2-StrepII: 23,326 Da.

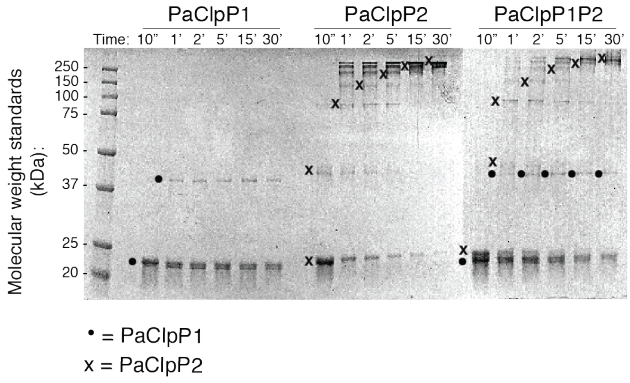

Figure 2-4: Chemical crosslinking suggests PaClpP1P2 is composed of two pure heptameric rings.
Time-dependent glutaraldehyde-crosslinking of purified PaClpP1$_{14}$, PaClpP2$_7$, and PaClpP1$_7$P2$_7$ visualized by SDS-PAGE. ClpP complexes were incubated with 0.125% glutaraldehyde and samples were withdrawn at 10 seconds, 1 min, 2 min, 5 min, 15 min, and 30 min. The crosslinking pattern of PaClpP1$_7$P2$_7$ is a combination of the crosslinking patterns of pure ClpP1$_7$ and pure ClpP2$_7$.

Crystal structures of PaClpP1 and PaClpP2 reveal key distinguishing features.

PaClpP1 shares 74% sequence identity with *E. coli* ClpP. In contrast, although the sequence analysis of PaClpP2 indicates that the protein has the features required to form an active peptidase, it shares only 40% identity with *E. coli* ClpP (Fig. 2-5). To better understand the molecular differences between the PaClpP1and PaClpP2 peptidase subunits we solved crystal structures of PaClpP1 and PaClpP2 to 2.6 Å and 2.0 Å resolutions, respectively (Table 2-1, PDB: 5BQV and 5BQW). PaClpP1 was crystalized in one of its biologically relevant assemblies, a homo-tetradecamer. PaClpP2, in contrast, crystalized as $ClpP2_7$, the structure it forms when purified in the absence of co-expressed with PaClpP1; it did however form a 14-mer by crystallographic symmetry (although we have no evidence that stable $PaClpP2_{14}$ forms or is functional either *in vitro* or *in vivo*). These crystal structures revealed important structural differences between the two ClpP isoforms.

Figure 2-5: Sequence alignment of ClpPs from *E. coli* and *P. aeruginosa*.
Numbering refers to the mature protein sequence. Residues boxed in red make up the catalytic triad. Residues boxed in blue with a Φ symbol above form the hydrophobic binding pocket that unfoldases dock to. The regions highlighted in yellow emphasizes the divergent N- and C-terminal sequences of PaClpP2.

As expected, PaClpP1 shared high structural similarity to *E. coli* ClpP (RMSD = 0.426); it adopted a more "extended" ClpP conformation (~98-100 Å in height) that has been described previously (Kimber et al., 2010; Liu et al., 2014). PaClpP2, in contrast, adopted a more compressed cylindrical barrel form with height by width dimensions of ~90 x 102Å (Fig. 2-6). Although the overall conformations of the PaClpP1 and PaClpP2 subunits and rings were largely similar, three distinct structural differences were evident. First, while the most N-terminal modeled residues of $PaClpP1_{14}$ formed an unstructured axial loop (residues 4-8, 18-20), the analogous visible N-terminal segment of $PaClpP2_{7}$ was an alpha helix that pointed downwards into the axial pore (residues 14-20, Fig. 2-7A, Liu et al. 2014). The dynamic N-terminal loops of numerous ClpP isoforms make important contacts with AAA+ unfoldases ClpX and ClpA, and work to modulate entry of substrates into the ClpP degradation chamber (Kang et al., 2004; Gribun et al., 2005; Bewley et al., 2006). A second structural difference was observed near the hydrophobic pockets located on the top and bottom surfaces of the two PaClpP isoforms. These pockets are critical for unfoldase•ClpP interactions as they provide docking sites for the ClpX and ClpA hydrophobic loops (sequence of docked residues is IGF for ClpX and IGL for ClpA), and these are the major sites of structural and functional interaction between Clp unfoldases and ClpPs (Amor et al., 2019).

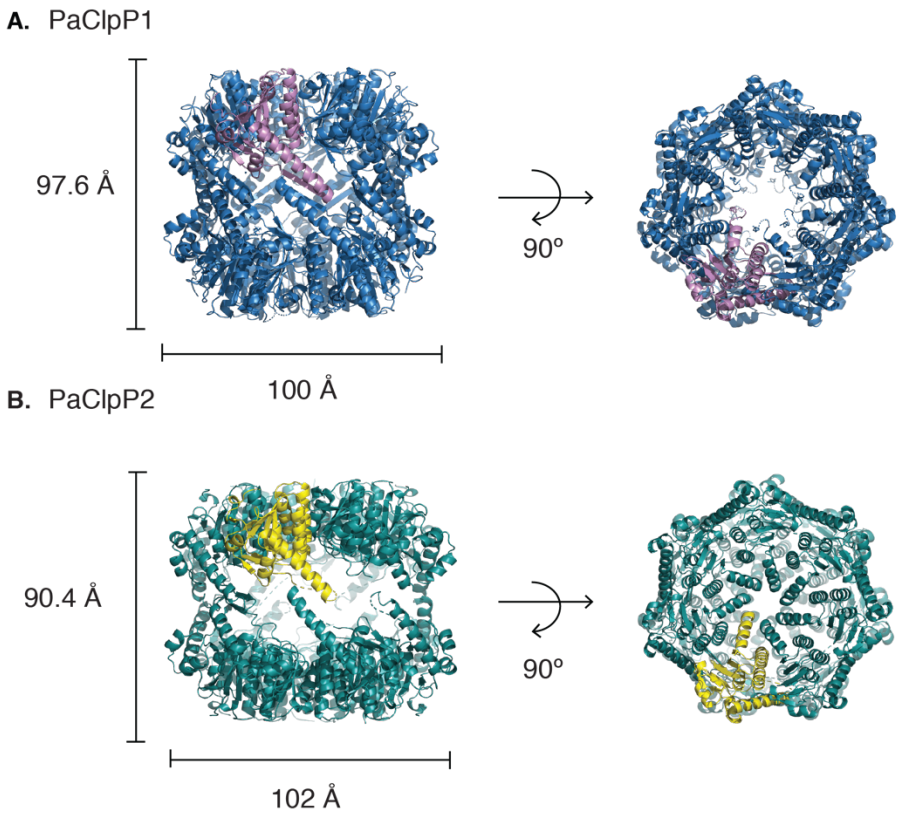

Figure 2-6: Crystal structures of PaClpP1$_{14}$ and PaClpP2$_7$ complexes.
A) Side and top-down views of homomeric PaClpP1 (PDB code: 5BQV, blue, with a single monomer in purple). B) Side and top-down views of homomeric PaClpP2 (PDB code: 5BQW, deep teal, with a single monomer in yellow).

Table 2-1: Crystallographic scaling and refinement statistics		
	PaClpP1	**PaClpP2**
Space group	P1	C121
Unit cell	96.97 97.11 108.34 66.99 85.93 77.16	124.37 147.43 99.33 90 121.73 90
Resolution (Å)	2.6	2.0
R_{merge}	0.131 (0.553)	0.073 (0.931)
R_{pim}	0.09309 (0.401)	0.044 (0.566)
No. of reflections	280114 (27232)	383187
No. of unique reflections	103265 (10111)	102045 (5093)
Completeness (%)	94.60 (92.38)	99.56 (99.53)
Mean I/sigma(I)	6.31 (1.70)	12.1
R_{work}	0.2071 (0.3121)	0.1969 (0.325)
R_{free}	0.2497 (0.353)	0.2175 (0.388)
Clash score	6.50	4.36
Rotamer outliers (%)	0.70	0.28
Ramachandran outliers (%)	0	0
Ramachandran favored (%)	98.04	97.36
RMS (bonds/angles)	0.005/0.79	0.009/.93

Table 2-1: Crystallographic scaling and refinement statistics.
Values in parentheses refer to the highest resolution shell. PaClpP1: PDB code 5BQV. PaClpP2: PDB code 5BQW.

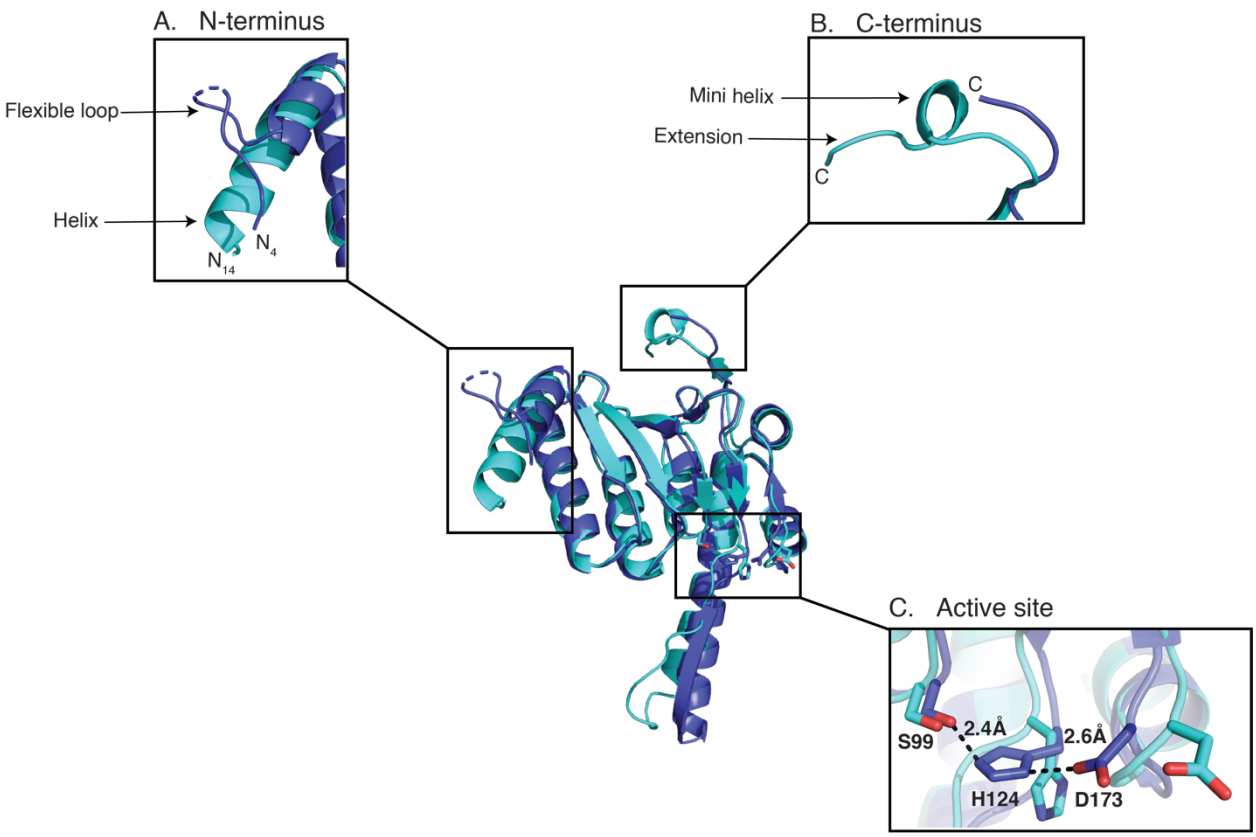

Figure 2-7: Overlay of PaClpP1$_{14}$ and PaClpP2$_7$ monomers.
Overlay of ClpP1 (blue) and PaClpP2 (cyan) monomers with insets showing select structural differences: (A) N-termini (Note: the first 13 residues of PaClpP2 and the first four residues of PaClpP1 are missing electron density.) (B) C-termini and (C) active sites. The active site residues of PaClpP1 are properly aligned (distances in angstroms are labeled on the figure), while the active site of ClpP2 is not (the distances between PaClpP2 S98 to H123, and H123 to D172 are 6.5Å and 10.1Å, respectively). Residue numbering in (C) refers to the mature amino acid sequence of PaClpP1.

Interestingly, the most C-terminal segment of PaClpP2 adopts a "mini helix" that protrudes over this critical hydrophobic binding pocket and which may therefore make unfoldase docking unfavorable (Fig. 2-7B). Thus, differences near both the N-terminus and C-terminus of PaClpP2 suggest it differs from PaClpP1 and other canonical ClpPs in its interactions with partner Clp unfoldases.

Finally, the structures revealed differences in the relative positioning of active site residues. In the PaClpP1 structure, the critical Ser-His-Asp residues largely appeared properly spaced and positioned for peptidase activity. In contrast, those in the PaClpP2 structure were not optimally positioned (compare the positions of the active-site His and Asp residues in Fig. 2-7C). In all subunits of PaClpP2, the active site D172 is too far (~10 Å) from S98 and H123 to form an active catalytic triad (Fig. 2-7C). This misalignment is consistent with crystallization of PaClpP2 as an inactive heptamer. We address the activity of PaClpP2 catalytic sites in the PaClpP1$_7$ClpP2$_7$ heterocomplex biochemically below. Despite significant effort, we were unable to grow high-quality crystals of the PaClpP1$_7$ClpP2$_7$ complex.

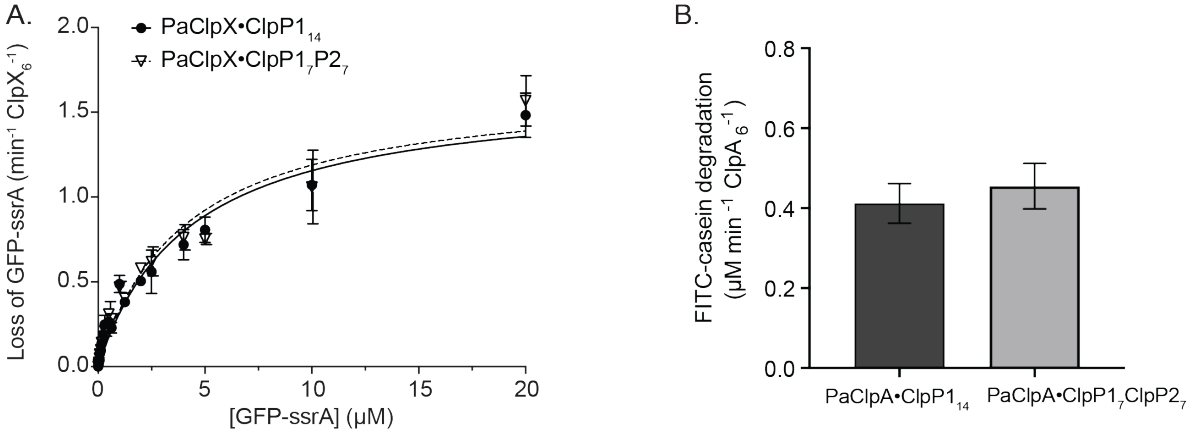

Figure 2-8: PaClpX productively interacts with the PaClpP1$_7$ side of PaClpP1$_7$P2$_7$.
A) Degradation of GFP-ssrA by PaClpX$_6$ (0.1 μM) and PaClpP1$_{14}$ (filled circles) or PaClpP1$_7$P2$_7$ (open triangles) (0.54 μM). Curves are fit to the Michaelis-Menton equation (solid line refers to PaClpX•ClpP1$_{14}$: K_M = 4.2 ± 0.5 μM, V_{max} = 1.6 ± 0.08 min^{-1} ClpX$_6^{-1}$ and dotted lines refers to PaClpX•ClpP1$_7$P2$_7$ K_M = 4.1 ± 0.52 μM, V_{max} = 1.7 ± 0.09 min^{-1} ClpX$_6^{-1}$). B) Degradation of 50 μM FITC-casein (Casein fluorescein isothiocyanate) by 0.4 μM PaClpP1$_{14}$ or PaClpP1$_7$P2$_7$ and 0.2 μM PaClpA$_6$. For both (A) and (B), data points are averages of three independent replicates performed in duplicate ± SD.

ClpX engages the PaClpP1$_7$P2$_7$ complex through interactions with PaClpP1

Having analyzed the structural differences between PaClpP1$_{14}$ and PaClpP2$_7$, we performed kinetic experiments to compare the ability of PaClpP1$_{14}$ and PaClpP1$_7$P2$_7$ heterocomplexes to degrade a degron-tagged model substrate, GFP-ssrA, when paired with PaClpX. The K_M and V_{max} values for GFP-ssrA degradation by PaClpX•P1$_7$P2$_7$ were indistinguishable from those of PaClpX•P1$_{14}$ (~4.1 μM and ~1.7 min^{-1} ClpX$_6^{-1}$), revealing that both hetero- and homocomplexes functioned equivalently with PaClpX to degrade GFP-ssrA (Fig. 2-8A). Likewise, the AAA+ unfoldase PaClpA functioned similarly in complex with either ClpP1$_{14}$ and or ClpP1$_7$P2$_7$ to degrade the model ClpAP substrate FITC-casein (Fig. 2-8B).

AAA+ unfoldases can dock to either one or both sides of the ClpP barrel of canonical isoforms, producing singly-capped (1:1) or doubly-capped (2:1) proteolytic complexes (Grimaud et al., 1998). However, the crystal structure of PaClpP2$_7$ suggests that, unless the N- and C- terminal regions of this protein undergo substantial conformational changes during co-complex formation, productive interactions between ClpX/A and PaClpP2 would be disfavored. To directly visualize the stoichiometry of PaClpX binding to PaClpP1$_7$P2$_7$, we mixed PaClpX with PaClpP1$_{14}$ or PaClpP1$_7$P2$_7$ (in the presence of ATPγS which stabilizes complexes) and scored the number of singly-capped and doubly-capped ClpP barrels by electron microscopy. Approximately half of the PaClpX• ClpP1$_{14}$ particles had PaClpX bound on both sides the ClpP core (Fig. 2-9). This observation was in stark contrast with what was observed with PaClpX and PaClpP1$_7$P2$_7$ complexes where only 2% of the particles were observed to have PaClpX on both faces and 98% of the complexes were singly-capped by the ATPase. Importantly, in these experiments, PaClpX was present at more than two-fold molar excess to either ClpP$_{14}$, conditions expected to favor

doubly-capped complexes. Thus, we conclude that doubly-ATPase-capped PaClpP1$_7$P2$_7$ is a rare species (Fig. 2-9).

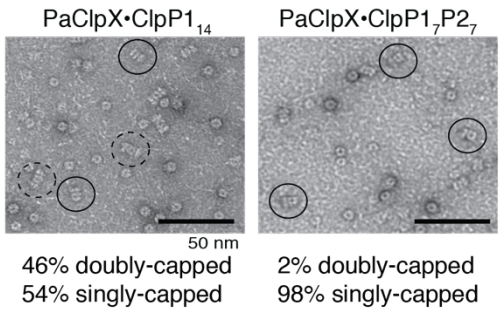

Figure 2-9:
Negative-stain microscopy images of PaClpX$_6$ (0.5 μM) with PaClpP1$_{14}$ or PaClpP1$_7$P2$_7$ (0.2 μM). Approximately half of the PaClpX•P1$_{14}$ complexes were doubly-bound by PaClpX$_6$ (left; 46% doubly-bound, 54% singly-bound), whereas most of PaClpX$_6$•ClpP1$_7$P2$_7$ complexes were singly-bound (right; 98% singly-bound). Representative doubly- and singly- bound complexes are circled with a dotted and filled line, respectively.

A critical residue for formation of ClpX•P complexes is Arg192 of EcClpP (Fei et al., 2020). We carried out enzymatic degradation assays using either wild-type or variant ClpPs to establish that the equivalent arginine (Arg194) in PaClpP1 is similarly important for PaClpX•ClpP assembly into an active protease (Fig. 2-10). Furthermore, no PaClpX + PaClpP1$^{R194K}{}_{14}$ co-complexes were observed by microscopy, reinforcing the conclusion that Arg194 of PaClpP1$_{14}$ is essential for assembly of the PaClpX•P1$_{14}$ complex. To assess biochemically whether the PaClpP2$_7$ face of PaClpP1$_7$P2$_7$ interacts with PaClpX, we purified PaClpP1$^{R194K}{}_7$P2$_7$ heterocomplexes (in which PaClpP2$_7$ was wild-type and PaClpP1$_7$ was a variant with the R194K mutation). GFP-ssrA degradation was not observed in reactions containing PaClpP1$^{R194K}{}_7$P2$_7$ and PaClpX (Fig. 2-10A). Interestingly, the Arg194 residue that is important for ClpX interaction is conserved in PaClpP2, although we have substantial evidence that the PaClpP2 in the PaClpP1$_7$P2$_7$ peptidase does not functionally interact with PaClpX.

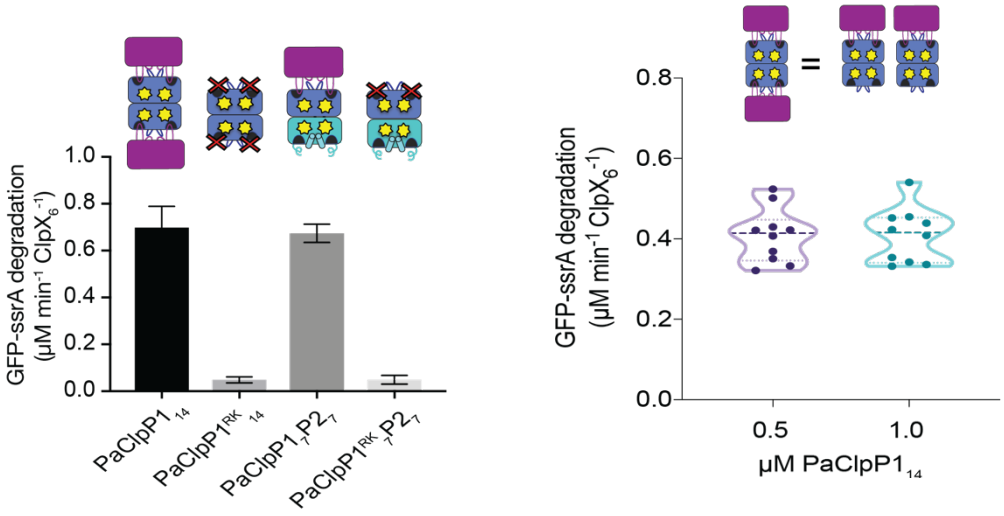

Figure 2-10: PaClpX engages the PaClpP1$_7$P2$_7$ complex through interactions with PaClpP1.
A) Efficient degradation of GFP-ssrA (8 μM) requires productive interaction of PaClpX$_6$ (0.5 μM) with the ClpP1$_7$ ring of ClpP1$_7$P2$_7$. Wild-type PaClpP1$_{14}$ or PaClpP1$_7$P2$_7$ (2 μM) with PaClpX degrades GFP-ssrA, whereas PaClpX-binding-defective variants of PaClpP1$_{14}$ (R194K) were not able to support degradation of GFP-ssrA in a heterocomplex with PaClpP2$_7$ (PaClpP1$^{RK}{}_7$P2$_7$). Data points are averages of three independent replicates performed in triplicate ± SD. B) Violin plots of GFP-ssrA degradation by PaClpX•ClpP1$_{14}$. The rate of GFP-ssrA (8 μM) degradation is dependent upon the amount of ClpP$_7$-bound-ClpX$_6$. 2:1 complex formations were favored in conditions containing 1.0 μM ClpX$_6$ and 0.5 μM of ClpP1$_{14}$, while 1:1 stoichiometries were favored to form conditions containing 1.0 μM ClpX$_6$ and either with 1.0 μM ClpP1$_{14}$ or 0.5 and 1.0 μM ClpP1$_7$P2$_7$.

To address the relative activities of PaClpP$_{14}$ complexes loaded with either one or two Clp ATPases (here PaClpX), we compared the degradation rates of proteases made with PaClpP1 and either one or two equivalents of PaClpX hexamer. These enzyme mixtures were used to degrade GFP-ssrA and the degradation rates were measured and compared under conditions that either favored formation of (i) 2:1 PaClpX$_6$•PaClpP1$_{14}$ or (ii) twice as many 1:1 PaClpX$_6$•PaClpP1$_{14}$ complexes. The degradation rates observed using these two experimental conditions were identical (Fig. 2-10B). This observation strongly suggests that doubly-PaClpX$_6$-bound PaClpP1$_{14}$ complexes can carry out proteolysis just as effectively as two singly-bound PaClpP1$_{14}$ complexes. Very similar results have been reported before for *E. coli* ClpAP and ClpXP (Ortega et al., 2002, Maglica et al.,

2009). Because PaClpX•ClpP1$_7$P2$_7$ only forms singly-ATPase-capped complexes, this "doubly active" protease phenomenon is not observed with the PaClpP1$_7$P2$_7$ heterocomplex, revealing a functional distinction between singly-and-doubly ClpX-capped complexes in *P. aeruginosa*.

ClpP2 active sites are essential for normal biofilm development in P. aeruginosa

Deletion of *clpP1* in *P. aeruginosa* results in several characterized cellular phenotypes including defects in: (i) in cell attachment, (ii) motility, (iii) pyocyanin and pyoverdine production, and (iv) pathogenicity in a mouse model (Fernández et al., 2012; Shanks et al., 2006; Hall et al., 2017; Zhao et al., 2016). We have previously shown that *clpP2* deletion specifically impairs the microcolony-formation stage of biofilm development in a flow-cell assay (Hall et al., 2017). Defects in biofilm assembly/maturation were also observed under static (non-flow) growth conditions with Δ*clpP2* strains. As shown in Figure 2-11, biofilms formed by Δ*clpP2* cells were thinner than those formed by wild-type PAO1, as measured by confocal microscopy using GFP-expressing strains (Fig. 2-11A,B). These data reveal that the *clpP2* deletion changes the three-dimensional architecture of *P. aeruginosa* biofilms, perhaps through alteration of the biofilm extracellular matrix (as regulation of alginate biosynthesis is under control of ClpP-dependent proteolysis in some strains). We further tested whether the peptide cleavage activity of PaClpP2 (presumably within a PaClpP1$_7$P2$_7$ complex) was important for the observed role of PaClpP2 in biofilm growth/morphology. In this case, rather than using a strain with a *clpP2* deletion, the *clpP2*^{S98A} allele, which encodes protein lacking the critical active site serine, was integrated into the PAO1 chromosome and the biofilm growth assay was repeated (Fig. 2-11 A,B). Importantly, this *clpP2*^{S98A} strain exhibited a very similar, even a slightly more severe biofilm-growth defect than did the Δ*clpP2* strain. These data lead us to propose that cleavage of a peptide or protein by

the PaClpP2 active site residues influences biofilm thickness. In contrast to the inhibitory effect of *clpP2* mutations, deletion of *clpP1* slightly stimulated biofilm thickness, as reported previously (Shanks et al., 2006). Importantly, the C-terminal flag tag on the *clpP2*^{S98A} allele was unlikely to influence biofilm formation, because *clpP2-flag* strains developed biofilms of normal thickness (Fig. 2-11C).

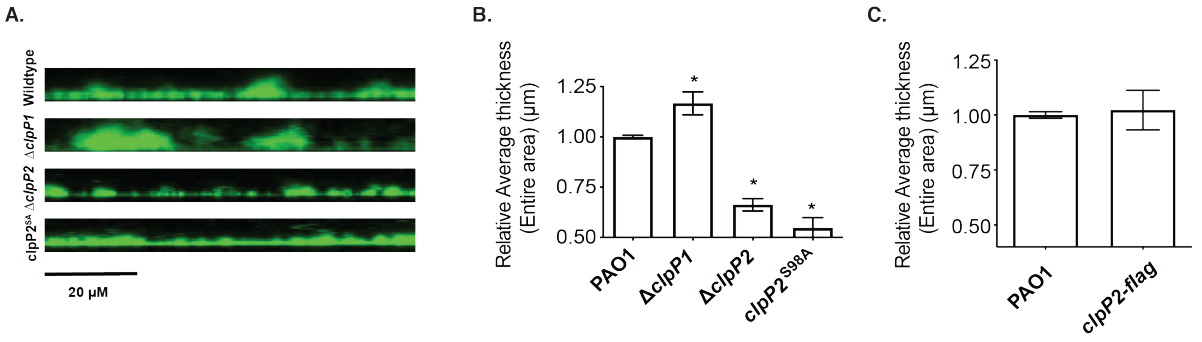

Figure 2-11: The active site of PaClpP2 is essential for normal biofilm development in *P. aeruginosa* PAO1.
A) Representative side-view confocal images of GFP-labeled PAO1 biofilms grown in 96-well glass-bottom plates under static conditions for 48 hours. Strains are either wild type, Δ*clpP1*, Δ*clpP2*, or contain a flag-tagged catalytic-null allele of *clpP2* in lieu of the wild-type gene (*clpP2*^{S98A}). B-C) Quantification of biofilms are averages of three independent wells, with five z-stacks taken per well. Error bars are ± SEM. Biofilm quantification was performed using COMSTAT software (Heydorn et al., 2000). * $p < 0.0001$ as analyzed by One-way ANOVA. C) The C-terminal flag-tag on *clpP2* does not interfere ClpP2 function in terms of regulating biofilm thickness.

PaClpP2 has a distinct cleavage specificity that contributes to ClpP1$_7$P2$_7$ activity

To characterize the peptidase activity of PaClpP1$_7$P2$_7$ and compare it to that of PaClpP1$_{14}$, purified peptidase complexes were first assayed for activity against a panel of 18 small peptides each linked to a C-terminal fluorogenic leaving group 7-amino-4-methylcoumarin (AMC) moiety; cleavage between the C-terminal most amino acid and the AMC in the peptide results in a fluorescent signal that is easily quantified. Notably, ClpP tetradecamers without an associated AAA+ unfoldase are

only active against small peptides that can diffuse into the ClpP degradation chamber through the barrel's axial pores (Maurizi et al., 1994).

Both PaClpP1$_{14}$, and PaClpP1$_7$P2$_7$ heterocomplexes had substantial and generally similar cleavage activity against seven of the eighteen tested peptide substrates, and exhibited weak cleavage activity against the remaining peptides (Fig. 2-12A). However, their activities against some sequences indicated that the two complexes had significantly different substrate preferences. Whereas most sequences were hydrolyzed at a faster rate by PaClpP1$_{14}$ than PaClpP1$_7$P2$_7$, two of the "well-cleaved" peptides, GGL-AMC and LLL-AMC, were preferentially cleaved by the heterocomplex (Fig. 2-12 A,B); likewise, one of the weakly-cleaved substrates, LEHD-AMC was also preferentially cleaved by PaClpP1$_7$P2$_7$.

To confirm this faster rate was due to the PaClpP2$_7$ active sites, we purified and assayed PaClpP1$^{SA}{}_7$P2$_7$, a variant with all PaClpP1$_7$'s active-site serines in the heterocomplex replaced by alanine. PaClpP1$^{SA}{}_7$P2$_7$ degraded GGL-AMC and LLL-AMC, confirming the PaClpP2$_7$ active sites within the heterocomplex that contribute substantially to the higher cleavage rates of these substrates (Fig. 2-12B). Conversely, PaClpP1$^{SA}{}_7$P2$_7$ displayed poor activity against AMC-modified IETD, AAN, and LLE, and was nearly inactive against the AAA, DEVD, and IIW substrates. (Note however, the basal activities of both wild-type complexes were also low against IETD, LLE, and DEVD, Fig. 2-12 A,B). Importantly, these hydrolysis experiments suggest that (i) both the PaClpP1 and PaClpP2 active sites contribute to the activity of PaClpP1$_7$P2$_7$ and (ii) PaClpP1 and PaClpP2 exhibit differences in their preferred cleavage sites.

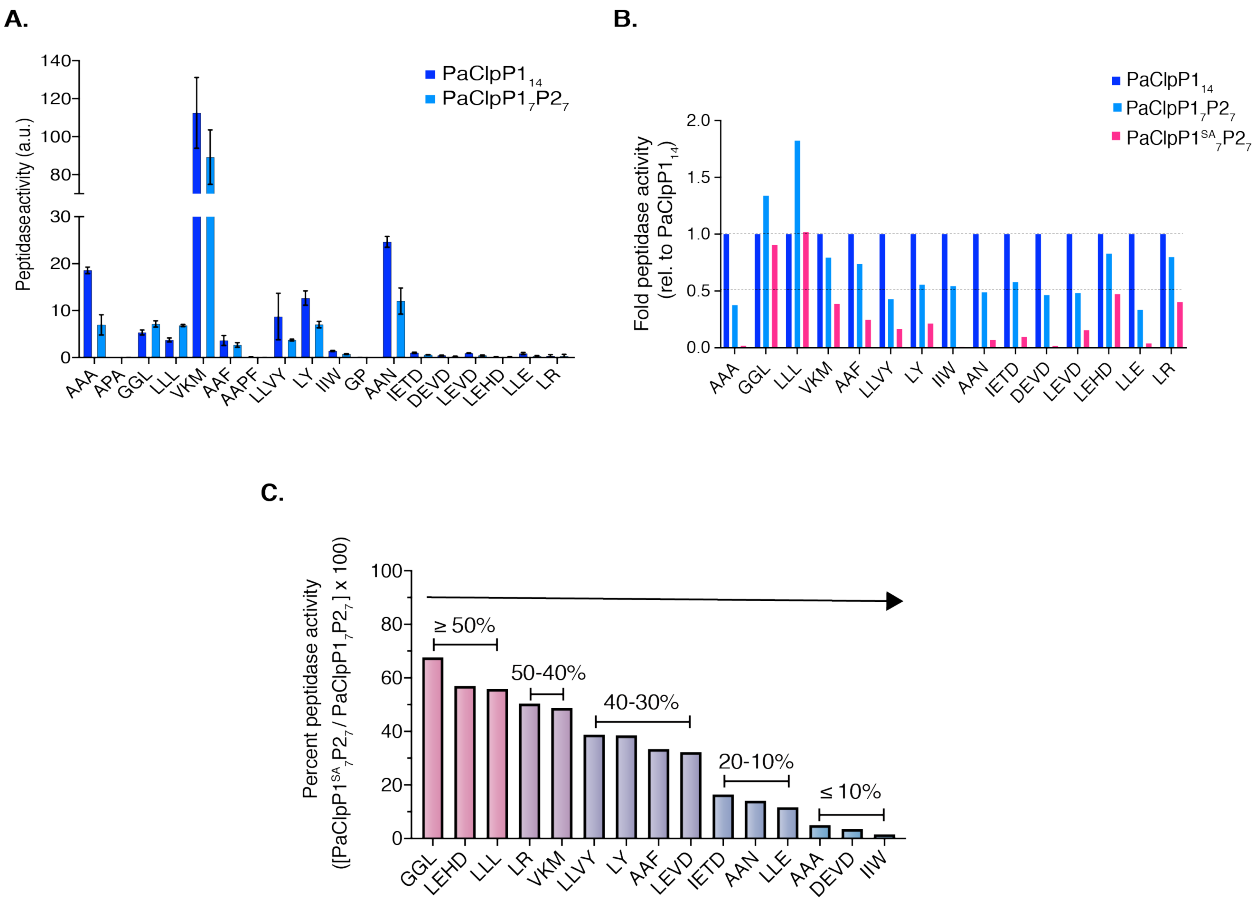

Fig. 2-12. The PaClpP1$_7$ and PaClpP2$_7$ active sites of PaClpP1$_7$P2$_7$ contribute uniquely to peptide cleavage.
A) Cleavage of a panel of 18 fluorescent peptide substrates (100 µM) by PaClpP1$_{14}$ or PaClpP1$_7$P2$_7$ (0.4 µM) at 30ºC. Substrates are ordered by the properties of the amino acid at position P1′ (from left to right: small to large, and non-polar to charged). Data points are averages of three independent replicates performed in duplicate ± SD. B) The PaClpP2$_7$ active sites contribute to peptide hydrolysis. PaClpP1$^{SA}{}_7$P2$_7$ is a partially-inactive variant of PaClpP1$_7$P2$_7$ in which the active serine of PaClpP1 is mutated to alanine. Peptide substrates are ordered as in (A). Substrates containing proline were tested but are not because of lack of activity against them. Peptidase activity is normalized to PaClpP1$_{14}$. C) Percent contribution of PaClpP2$_7$ activity to total peptidase activity in PaClpP1$_7$P2$_7$. Substrates are grouped from left to right based on their percent contribution to PaClpP1$_7$P2$_7$ peptidase activity, such that percent peptidase activity is equal to the quotient of normalized peptidase rates relative to PaClpP1$_{14}$ of PaClpP1$^{SA}{}_7$P2$_7$ by PaClpP1$_7$P2$_7$ (high to low).

PaClpP1$_{14}$ and PaClpP1$_7$P2$_7$ display distinct cleavage patterns within AAA+ proteases.

Having confirmed that the PaClpP2 subunits are catalytically active within the PaClpP1$_7$P2$_7$ heterocomplex, we next sought to determine the protease cleavage specificities of the different

ClpP complexes in the context of the ClpXP proteases. Protein degradation assays were carried out with four *P. aeruginosa* protease complexes: ClpX•$P1_{14}$, ClpX•$P1_7P2_7$, ClpX•$ClpP1^{SA}{}_7P2_7$ and ClpX•$ClpP1_7P2^{SA}{}_7$. Each protease was used to degrade four different substrates with C-terminal ssrA degradation tags (Table 2-2). More than 13,000 cleaved peptides were identified by LC-MS/MS for each protease complex; these results allowed us to determine product peptide lengths and cleavage sequences. Although all four complexes generated peptides that averaged 8-11 amino acids long (Fig. 2-13), the two complexes that each contained one inactivated ClpP ring (PaClpX•$ClpP1^{SA}{}_7P2_7$ and PaClpX•$ClpP1_7P2^{SA}{}_7$) generated a larger proportion of longer peptides (ranging from 12-28 amino acids). This observation is consistent with previous reports of the product size distribution of partially inactivated ClpP 14-mers (Thompson et al., 1994).

Table 2-2: Protease complexes and substrates used to determine protease cleavage specificity

Proteases tested	Substrates tested
PaClpX•$P1_{14}$	GFP-ssrA
PaClpX•$P1_7P2_7$	FNR-ssrA
PaClpX•$P1^{SA}{}_7P2_7$	arc-GCN4p1-st11-ssrA
PaClpX•$ClpP1_7P2^{SA}{}_7$	Halo-(V15P titin^{I27})$_4$-ssrA

Note: The C-terminal ssrA tag on these model substrates is from the *E. coli* sequence (-AANDENYALAA). The *P. aeruginosa* ssrA sequence is otherwise identical except it contains an Asp instead of a Glu in a position of the sequence not shown to be critical for interaction with ClpX or ClpA (-AANDDNYALAA) (Flynn et al., 2001). GFP-ssrA: Green fluorescent protein-ssrA, FNR-ssrA: *E. coli* fumarate and nitrate reductase-ssrA, arc-GCN4p1-st11-ssrA: bacteriophage P22 arc-repressor fused to the p1 peptide of the yeast transcription factor GCN4-ssrA, Halo-(V15P titinI^{27})$_4$-ssrA: four titin domains fused to an N-terminal HaloTag domain.

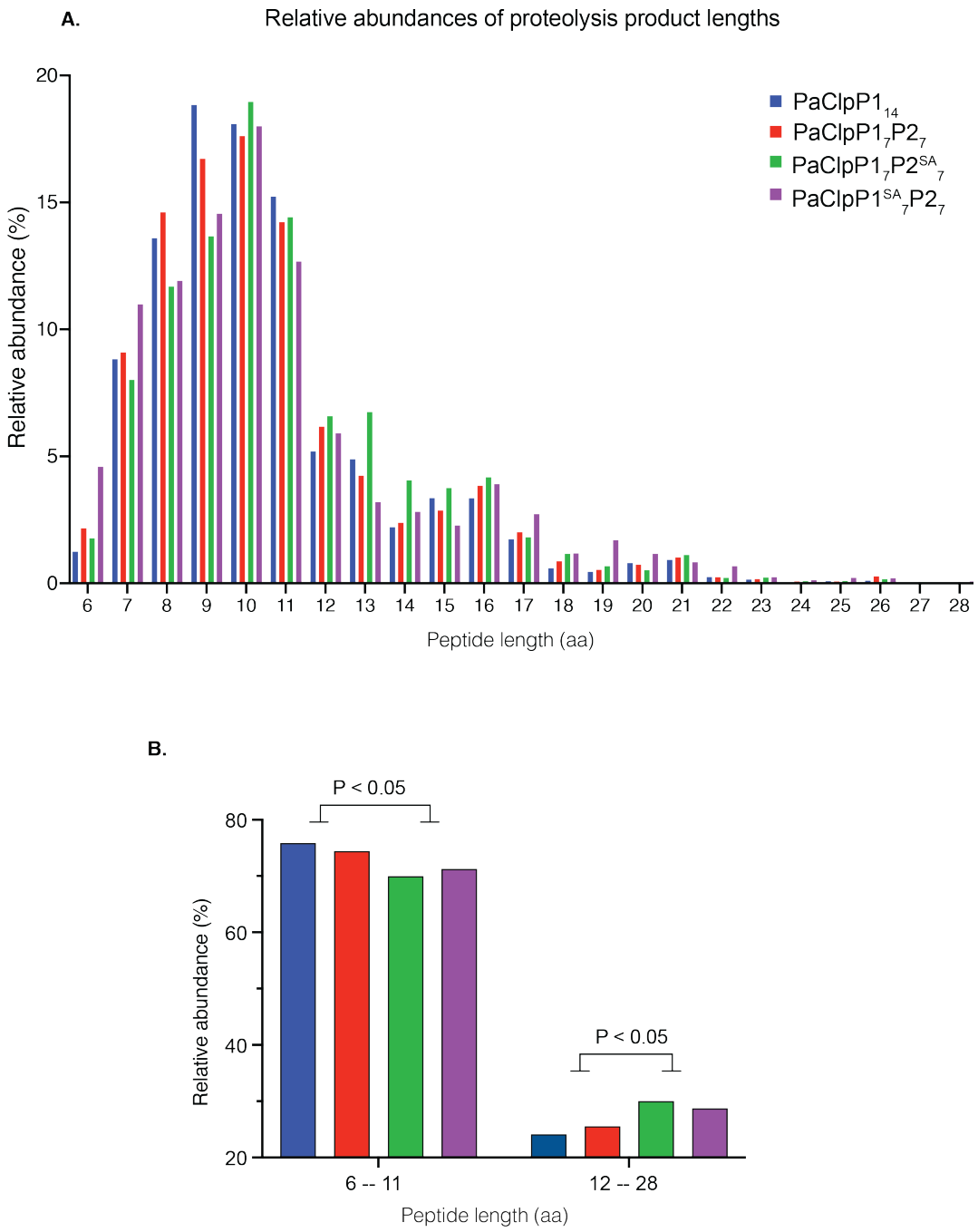

Fig. 2-13. Distribution of lengths of proteolytic products generated by *Pa*ClpX•P complexes.
A) Peptide length versus relative abundance of proteolytic products generated by PaClpX•ClpP1$_{14}$ (blue), PaClpX•ClpP1$_7$P2$_7$ (red), PaClpX•ClpP1$^{SA}{}_7$P$_2$ (purple), and PaClpX•ClpP1$_7$P2$^{SA}{}_7$ (green). More than 13,000 peptides per degradation reaction were identified by LC-MS/MS. B) Substitution of either PaClpP1 or PaClpP2 serine active site to alanine results in longer proteolytic cleavage products. Peptides were grouped into two classes according to length: 6-11 amino acids and 12-18 amino acids long. $p < 0.05$ as analyzed by a two-tailed t-test.

Preferred sequences surrounding the cleavages sites produced by each enzyme are displayed as a WebLogo (Crooks et al., 2004). Overall, wild-type PaClpX•ClpP1$_{14}$ and PaClpX•ClpP1$_7$P2$_7$ display generally similar protease cleavage preferences, favoring Gly, Ala, Leu, and Phe at the P1 position (Fig. 2-14). The nomenclature for labeling the residues the substrate, N to C terminal surrounding the cleavage site, is P3-P2-P1-P1′-P2′-P3′, with the scissile bond between P1-P1′ (Schechter and Berger, 1967). The proteolytic products from PaClpX•P1$^{SA}{}_7$P2$_7$ and PaClpX•P1$_7$P2$^{SA}{}_7$ analyzed by WebLogo revealed that the small residues Gly and Ala were notably de-enriched at position P1 from complexes containing only PaClpP2 active site. The large hydrophobic Leu and Phe residues remained enriched at P1, and notably, Arg also appears to be

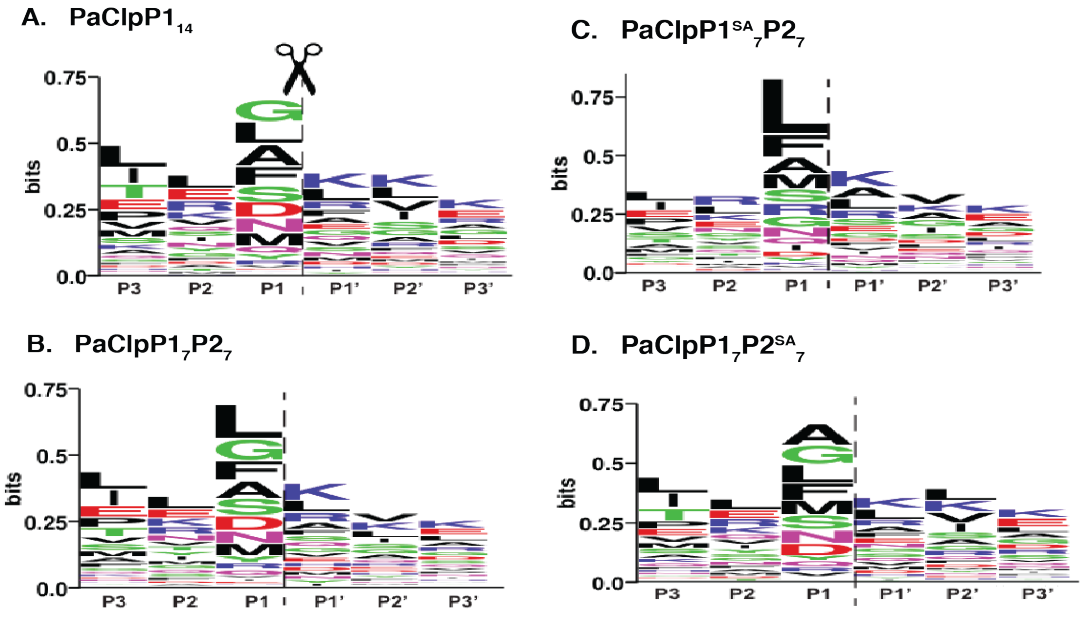

Figure 2-14: Protease cleavage consensus motifs by PaClpX•ClpP complexes.
Data are from Fig. 2-13. More than 13,000 proteolytic products spanning four substrates (GFP-ssrA, FNR-ssrA, arc-GCN4p1-st11-ssrA, and Halo-(V15P titin^{I27})$_4$-ssrA) (see Table 2-1) were sequenced by mass spectrometry and compiled to create WebLogos (Crooks et al., 2004). A) Motif generated from PaClpX•ClpP1$_{14}$ complexes. (B) Motif generated by PaClpX•ClpP1$_7$P2$_7$ complexes. (C) Motif generated by PaClpX•ClpP1$^{SA}{}_7$P$_2$ complexes. (D) Motif generated by PaClpX•ClpP1$_7$P2$^{SA}{}_7$.

substantially more preferred at P1 in cleavages catalyzed by only PaClpP2. Note however that, for example, this preference for Arg at the P1 position was not evident in the wild-type co-complexes, demonstrating that specificity at the P1 site by PaClpX•ClpP1$_7$P2$_7$ is apparently not equally partitioned between ClpP1 and ClpP2, but rather is more heavily influenced by the ClpP1 active site (compare Figs. 2-14B and C).

Whereas cleavage by PaClpP1$_7$P2$_7$ at the P1 position appears to be dominated by PaClpP1$_7$, the most notable difference in cleavage motifs between wild-type PaClpP1$_{14}$ and wild-type PaClpP1$_7$P2$_7$ were at positions P2 and P1′. The preference for specific amino acids at these positions for cleavage by the PaClpP1 or PaClpP2 active sites was further supported by protease assays using active site variants of the PaClpP heterocomplex. For example, at position P2, Asn was enriched in reactions that contained wild-type PaClpP2 subunits. In addition, at the P1′ position, PaClpX•ClpP1$_7$P2$_7$, but not PaClpX•ClpP1$_{14}$, displayed a clear preference for substrates containing Ser, and PaClpX•ClpP1$_{14}$ prefers substrates with Phe at this position (Fig. 2-15). These results establish that the sequence determinants within substrates that specify preferential cleavage by the PaClpP1 sites in the homocomplex versus the PaClpP1 and PaClpP2 active sites in PaClpP1$_7$P2$_7$ heterocomplex are multifaceted, and include the identity of residues at positions P2, P1, and P1′.

Thus, these final experiments lead to two key observations: (i) that PaClpP1$_{14}$ and PaClpP1$_7$P2$_7$ have distinct (though partially overlapping) cleavage preferences, and (ii) there is a unique requirement for functional ClpP2 actives site for proper biofilm development. These data therefore strongly support a model involving a peptide(s) or protein(s) that is preferentially cleaved by ClpP2

within the heterocomplex that goes on to play a substantial positive role in the properly structured and clinically dangerous developmental of biofilm during the late stages of *P. aeruginosa* growth.

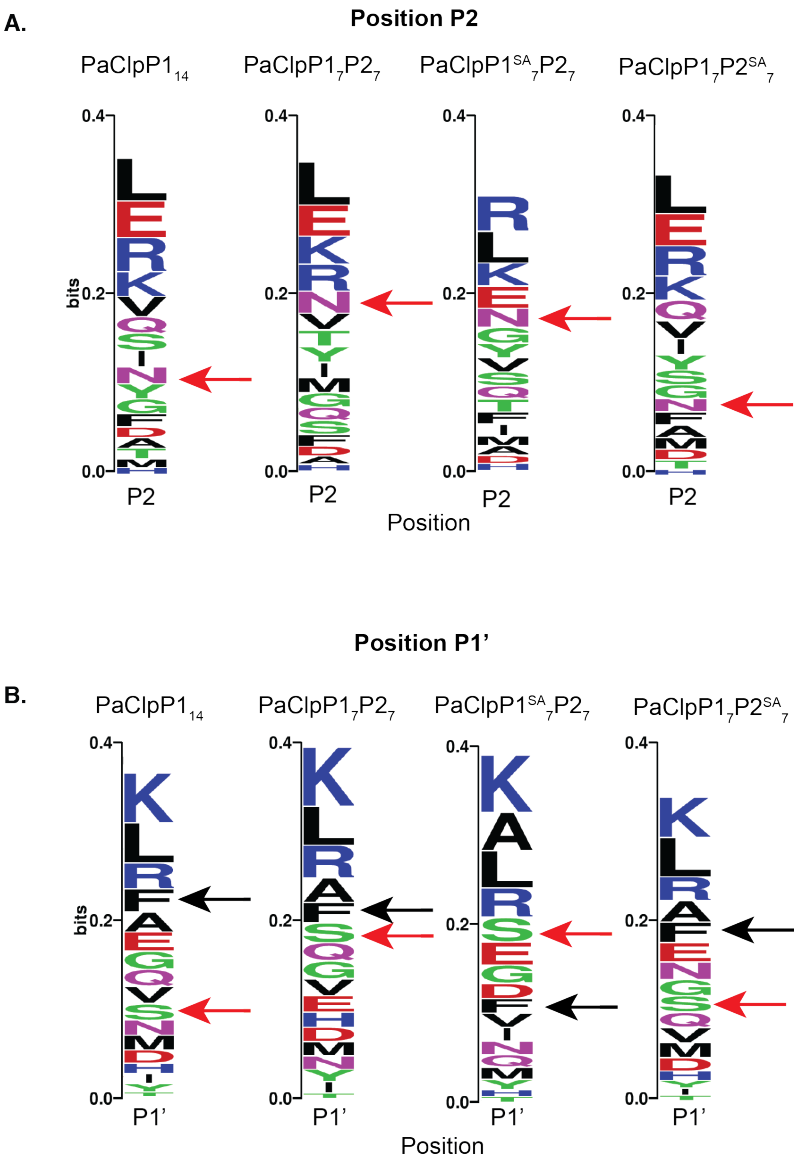

Figure 2-15: Positions P2 and P1′ of substrates confer differential cleavage preference by PaClpX•ClpP$_{14}$ complexes.
A) Zoom-in of position P2 as shown in Fig. 2-14. B) Zoom-in of position P1′ as shown in Fig. 2-14. In (B), black arrows point to Phe$_{P1}$ which is favored by the PaClpP1 active sites (i.e.: its presence in the motif shifts down upon inactivation of PaClpP1), and in (A) and (B), red arrows point to residues that are favored by the PaClpP2 active sites (i.e.: their presence in the motif is enriched only when wild-type PaClpP2 is present in the heterocomplex; Asn$_{P2}$ in (A), and Ser$_{P1'}$ in (B)).

Discussion

P. aeruginosa makes two distinct functional ClpP peptidases during development.

Previous studies established that the opportunistic pathogen *P. aeruginosa* encodes and expresses two distinct isoforms of ClpP peptidase subunits. ClpPs form barrel-shaped peptidases constructed of two seven-subunit rings that function with Clp-family unfoldases to generate AAA+ proteases. In a prior study, we established that PaClpP1 is expressed throughout the bacteria's different growth phases, assembles into functional tetradecamers, is an active peptidase, and forms a critical part of functional proteases when partnered with the PaClpX or PaClpA AAA+ unfoldases (Hall et al., 2017). Here we add to this work by purifying PaClpP2 from native *P. aeruginosa* cultures and from *E. coli* expression strains that co-overproduce *clpP1* and *clpP2*, and identifying the active form of PaClpP2 as half of a $PaClpP1_7ClpP2_7$ hetero-tetradecameric enzyme. Based on this work, as well as expression studies showing *clpP2* (but not *clpP1*) is induced approximately 10-fold during stationary phase and biofilm development, we conclude that *P. aeruginosa* makes two distinct forms of ClpP peptidase, $PaClpP1_{14}$ and $PaClpP_7ClpP2_7$, during its life cycle, and these different peptidases play substantial roles in determining developmental outcomes during bacterial growth (Waite et al., 2006).

PaClpP2 has distinct structural features and is related to L. monocytogenes ClpP1.

PaClpP1 is clearly a canonical member of the ClpP peptidase family, being 73% identical to *E. coli* ClpP and exhibiting very similar biochemical properties (Hall et al., 2017). Curiously, however, PaClpP2 has a substantially different sequence, and in phylogenetic trees of *clpP* genes, PaClpP2 is on a branch distinct from that carrying PaClpP1, but rather with ClpP2s from close relatives of *P. aeruginosa* (not shown) and surprisingly with ClpP1 from the firmicute *Listeria*

monocytogenes (*L. monocytogenes*). This observation suggests that PaClpP2 and LmClpP1 were likely acquired via horizontal gene transfer from a common ancestor (Fig. 2-16, note: LmClpP2 is the "canonical" ClpP isoform in *L. monocytogenes*, whereas LmClp1 is the unusual isoform, similar to PaClpP2; genes were named historically, not by homology or function). Indeed, PaClpP2 shares 50% identity with LmClpP1, 42% identity with *E. coli* ClpP, and only 40% identity with PaClpP1 (Fig. 2-17). Interestingly *L. monocytogenes*, like *P. aeruginosa*, makes two types of ClpP enzymes: a canonical ClpP tetradecamer ($LmClpP2_{14}$) and a $LmClpP1_7P2_7$ heterocomplex (Dahmen et al., 2015).

Figure 2-16: Phylogenetic tree of ClpP from representative species of Firmicutes, Proteobacteria, and Chlamydiae.
Branches and nodes belonging to the phylum Proteobacteria are in blue, Firmicutes in green, and Chlamydiae in purple. Select species representing each phyla were chosen for simplicity. ClpP1 and ClpP2 from *P. aeruginosa* are in bold-type font and circled. ClpP from *E. coli* (ECOLI) is labeled in blue. Note that ClpP2 from the species *P. aeruginosa* (PSEAE) and ClpP1 from *L. monocytogenes* (LISMO) share a

common node from the branch of Chlamydiae. Alignments were made using Clustal Omega, and phylogenic tree was constructed using the Interactive Tree of Life (iTOL, Sievers et al., 2011; Letunic et al., 2019).

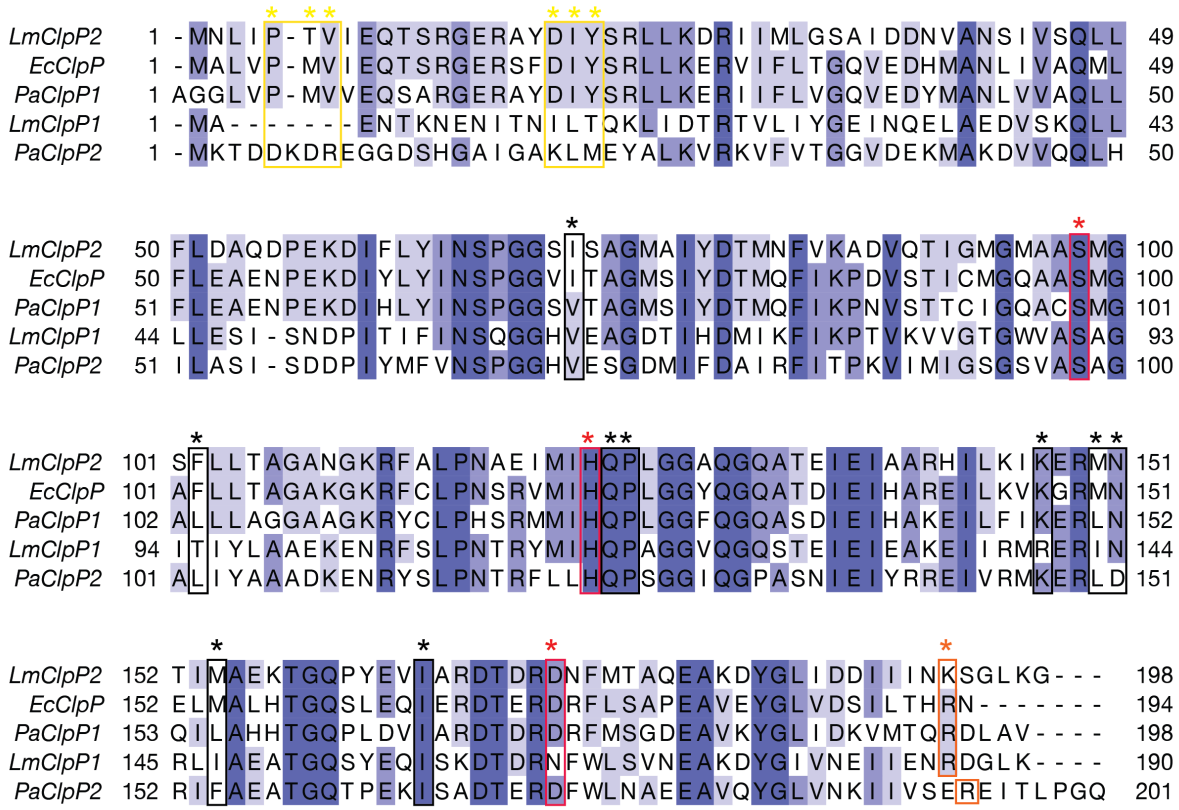

Figure 2-17: Sequence alignment of ClpP from *E. coli*, *L. monocytogenes*, and *P. aeruginosa*. Numbering refers to the mature ClpP sequence. Residues boxed and starred in yellow have been experimentally validated to be important for interaction with AAA+ unfoldases (Bewley et al, 2006). Resides boxed and starred in red make up the catalytic triad. Those boxed and starred in black make up the S1 binding pocket (Szyk and Maurizi, 2006). The residue boxed and starred in orange (R193 in EcClpP) has been shown to be essential for interactions with ClpX in *E. coli* and *P. aeruginosa* (Fei et al., 2020 and this study).

Multiple experiments provide evidence that the active form of PaClpP2 is a 7-mer that stacks with a 7-mer of PaClpP1 to form the active $PaClpP_7ClpP2_7$ hetero-tetradecamer. A remarkable feature of the $ClpP2_7$ side of this complex is that it doesn't productively interact with the AAA+ unfoldases PaClpX or PaClpA. Our high-resolution structure of PaClpP2 reveals that two regions of sequence

that differ from canonical ClpPs, and these regions encode structural elements that explain the lack of unfoldase interaction: (i) the standard ClpP N-terminal flexible loops that interact with the pore 2 loops of the unfoldase to form a contiguous pore between the two functional partners of the protease are absent from PaClp2, being replaced instead by a helix that faces into the PaClpP2 pore; this helix would likely constrict what would otherwise be a substrate entryway. (ii) The ClpP hydrophobic pockets, which are the critical docking sites for IGF/L loops of the ATPases, are at least partially blocked by an "extra" small C-terminal helix in PaClpP2 which likely antagonizes unfoldase docking. We attempted to test this hypothesis by introducing short C-terminal deletions into the *clpP2* gene. However, the extremely poor solubility of these deletion variants prevented further validation of this model. Without a crystal structure of the PaClpP1_7P2_7 heterocomplex, it is unclear whether or how these structural features might toggle into an alternate conformation that might then enable proteolytic complex assembly on both faces of PaClpP1_7P2_7. However, a dramatic conformational switch upon hetero-tetradecamer formation to now allow unfoldase binding appears unlikely, as our biochemical and microscopy studies were carried out with assembled PaClpP1_7P2_7 complexes and revealed predominantly singly-bound ClpX•ClpP complexes in solution (Fig. 2-9). Likewise, introducing a ClpX-binding-defective mutation into the ClpP1_7 side of the PaClpP1_7P2_7 resulted in ClpP assemblies completely defective in PaClpX interaction, as expected if now both the variant ClpP1 and the wild-type ClpP2 faces were unable to bind the unfoldase (Fig 2-10A). Interestingly, like PaClpP1P2, LmClpP1P2, binds asymmetrically to ClpX, with the LmClpX binding to the LmClpP1 side, indicating that failure to bind AAA+ unfoldases is a conserved property of this alternative ClpP subclass. However, LmClpX is reported to bind substantially more tightly to LmClpP1P2 than to LmClpP2_{14}

homocomplexes, a feature that we did not observe for PaClpX and PaClpP1P2 (data not shown, Balogh et al., 2017).

Models for the roles of two ClpP forms in P. aeruginosa: Cleavage specificity as a regulatory function?

Our new insights into *P. aeruginosa* ClpP2, together with previous studies, illuminate possible models for the biological roles of the two distinct active ClpP peptidases during bacterial growth and development. The top panel of Figure 2-18 summarizes the major distinctions between *P. aeruginosa's* two active ClpP complexes. Briefly, $ClpP1_{14}$ is a canonical family member; the *clpP1* gene is in an operon with *clpX* and is constitutively expressed throughout exponential and stationary phase growth. As an enzyme, $PaClpP1_{14}$ is an active peptidase with broad cleavage specificity, and interacts with both PaClpX and PaClpA to form active proteases. These proteases can be singly- or doubly-capped by an unfoldase, and the $ClpP_{14}$ is twice as active when doubly capped. $PaClpP1_{14,}$ therefore is very likely a component of typical ClpXP and ClpAP enzymes that function in growing bacteria to identify and degrade protein quality control substrates, such as ssrA-tagged failed translation products, as well as transcription factors and other regulatory/metabolic proteins. For example, PaClpA, PaClpP1, and the ClpA-adaptor protein (ClpS) are reported to control levels of the transcription factor CdpR, which functions with a group of regulators. CdpR, and its partner regulators control production of pyocyanin, biofilm development, and pathogenicity in model mice (Zhao et al., 2016). Interestingly, *clpP1*-defective strains show increased pyocyanin production, and altered structure (thicker and fluffier) biofilms, likely due to changes in these early regulatory pathways that contribute to pathogenesis (Hall et al., 2017 and this study).

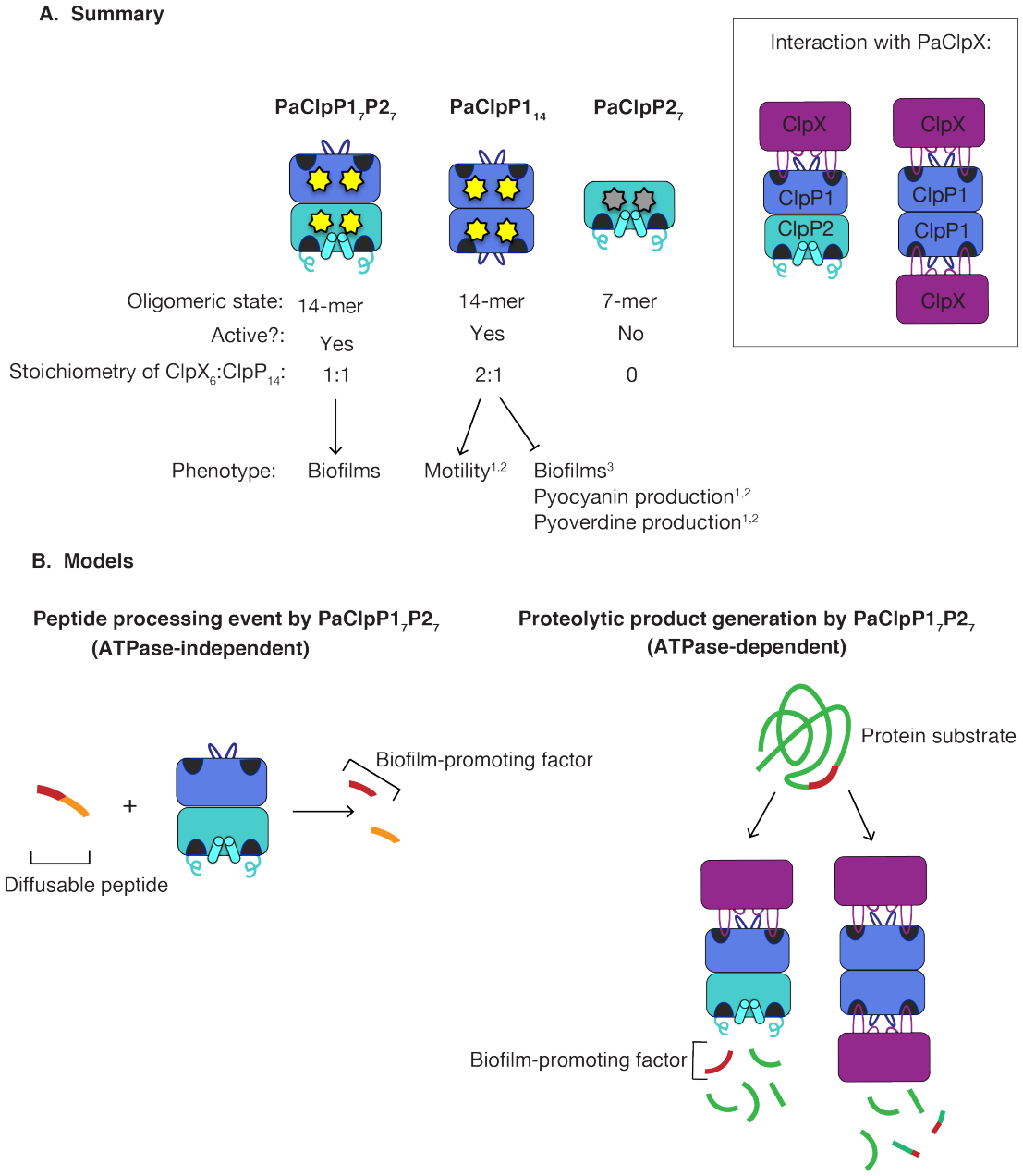

Figure 2-18: Summary and models for PaClpP1 and PaClpP2 assemblies and functions.
A) Graphical summary of the oligomeric states and biochemical activities of PaClpP1$_{14}$, PaClpP2$_7$, and PaClpP1$_7$P2$_7$. Whereas PaClpP2$_7$ is an inactive heptamer (Hall et al., 2017), PaClpP1$_{14}$ and PaClpP1$_7$P2$_7$ are active tetradecamers that can interact with PaClpX and PaClpA to degrade model substrates. PaClpP1$_7$P2$_7$ binds asymmetrically to PaClpX (right panel), whereas one PaClpP1 14-mer can bind two equivalents of PaClpX hexamers (i.e., is doubly-capped). We find in this study that PaClpP1$_7$P2$_7$ promotes biofilm thickness. Previous reports have established functions for PaClpP1 in motility, biofilm formation, and phenazine production *in vivo* (Hall et al, 2017[1]; Zhao et al., 2016[2]; 3: this study, and Shanks et al., 2006[3]). B) Models for *in vivo* functions of PaClpP1$_7$P2$_7$. The left panel shows an ATPase-independent model for PaClpP1$_7$P2$_7$ activity, in which a small propeptide can diffuse

into the axial pore of the heterocomplex and be cleaved by PaClpP1$_7$P2$_7$ to produce an active peptide product. The right panel depicts an ATPase-dependent function of PaClpP1$_7$P2$_7$, where a biologically important proteolytic product is generated by unfoldase-coupled PaClpP1$_7$P2$_7$. Both (A) and (B) represent models in which a currently unknown target with a functional role in biofilm architecture is produced via the cleavage by PaClpP1$_7$P2$_7$.

The *clpP2* gene and the PaClpP1$_7$P2$_7$ enzyme exhibit several distinctive features compared to PaClpP1. Expression of the *clpP2* gene is activated by quorum sensing, which includes regulation by the transcription factor LasR (Gilbert et al., 2009). As a consequence, PaClpP2 is virtually undetectable during exponential growth phase, but becomes abundant in stationary phase and biofilms. Furthermore, instead of clustering on the genome with other protein quality control enzymes like *clpP1*, *clpP2* is found linked to a cluster of genes called *aze*, which are also controlled by quorum sensing and have recently been shown to synthesize a modified nonribosomal peptide called azetidomonamide B in one study and azabicyclene in another (Hong et al., 2019; Patteson et al., 2019). Genes homologous to *clpP2* and the *aze* cluster have been identified together in a number of bacterial genomes, which suggest a potential functional relationship. However, this intriguing potential interplay between the *aze* gene cluster product and ClpP enzymes is only beginning to be explored.

At the level of enzyme function, two key differences between PaClpP1$_7$P2$_7$ and PaClpP1$_{14}$ stand out. First, PaClpP2 cannot form an active protease on its own because it has evolved structural features that prevent efficient AAA+ unfoldase interaction. Thus, to function as a protease, PaClpP2 must co-assemble with PaClpP1 to form a PaClpP1$_7$P2$_7$ complex; degradation studies reveal this co-complex can bind to one unfoldase ring, on the ClpP1 side, and in this configuration both the PaClpP1 and PaClpP2 peptidase active sites participate in protein degradation. The second significant difference between PaClpP1 and PaClpP2 is that they have distinct amino acid

signatures to their peptide and protein cleavage sites. For example, we found that PaClpP1$_7$P2$_7$ preferentially cleaved substrates with Leu at the P1 position and Ser at P1′, whereas ClpP1$_{14}$ favored the small residues Ala and Gly at P1, and Gln at P2 (Figs. 2-14, 2-15). The differential cleavage specificities of PaClpP1 and PaClpP2 at the P1 position likely reflect the structural properties of the PaClpP1 and PaClpP2 S1 binding pockets (Hedstrom, 2002; Szyk and Maurizi, 2006). Indeed, the S1 binding pocket of PaClpP2 contains a Phe residue (position 154) that may potentially aid in better hydrophobic packing of substrate side-chains during peptide positioning and catalysis (Fig. 2-19).

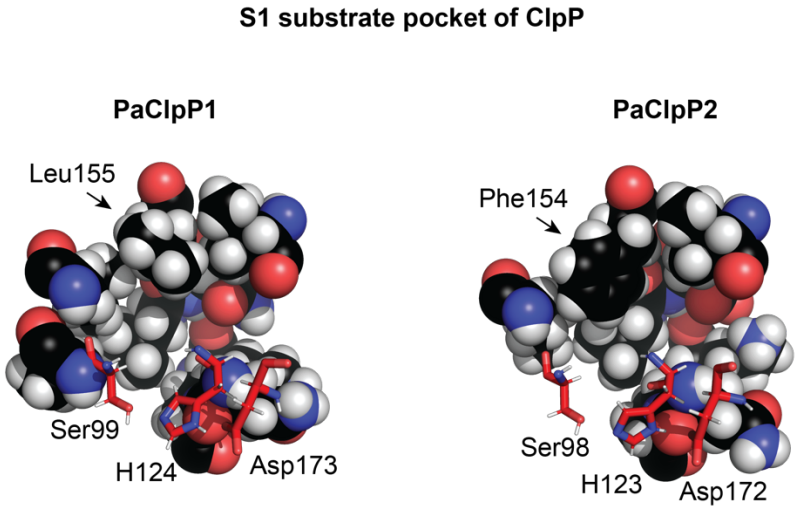

Figure 2-19: S1 binding pockets of PaClpP1 and PaClpP2.
Structures of the S1 binding pockets of PaClpP1 (left) and PaClpP2 (right). Active site resides are represented as red stick figures and residue numbers are labeled. Notable structural differences are pointed out with an arrow (Leu155 in PaClp1, Phe154 in PaClpP2).

Although not comprehensive, these results indicate that ClpP2 in complex with ClpP1 imparts an altered substrate processing ability, both to the ClpP1P2 hetero-multimer itself and to an associated AAA+ protease complex. Our *in vitro* data demonstrate that ClpP2 modulates both ATPase-dependent and -independent processing of substrates by PaClpP1, and our *in vivo* data demonstrate

that growth of normally structured thick biofilm is dependent on ClpP2 enzymatic activity. Taken together, these results suggest that PaClpP1$_7$P2$_7$ likely produces a cleavage product with a downstream positive functional role for biofilm formation (Fig. 7B). The observation that the unique PaClpP2 active-site-cleavage specificity has a positive role in bacterial development emphasizes that the peptidase subunits of AAA+ proteases, not just the ATPase regulatory particles can play a key role in determining biological outcomes of protein/peptide degradation.

Bacteria, (plants,) and animals encode multiple peptidase isoforms.

Although genomes encoding multiple alternative peptidase subunits for AAA+ proteases are continuing to be discovered, the best-characterized to date is the immunoproteasome (i-proteasome) of jawed vertebrates. A biologically important and developmentally regulated example of cells inducing expression of alternative peptidase subunits to help deal with cellular stress is observed with the i-proteasome. In this case, in response to an assault such as viral infection, cells induce expression of three alternative β subunits of the 20S particle. The difference in cleavage preference, especially of the alternative subunit known as β1, results in the pool of peptides generated during proteolysis that is biased away from cleavage after acidic residues and enhanced in production of peptides with N-terminal hydrophobic residues. This latter family of peptides is better suited to be displayed by the peptide-binding receptors of the immune system. This example, together with our early observations indicating that the unique PaClpP2 active-site-cleavage specificity emphasizes that the peptidase subunits of AAA+ proteases, not just the ATPase regulatory particles, can play a key, and perhaps widespread role in determining the biological outcomes of protein/peptide degradation (see Chapter 3 for more discussion on the i-proteasome).

Methods

Strains and plasmids

PCR was carried out on *Pseudomonas aeruginosa* PAO1 genomic DNA provided by the American Type Culture Collection (ATCC). Plasmids were selected using 100 μg/ml ampicillin, 50 μg/ml kanamycin, 10 μg/ml gentamicin (*E. coli*) or 30 μg/ml gentamicin (*P. aeruginosa)* as appropriate. Deletion or allelic-exchange strains of *Pseudomonas aeruginosa* PAO1 were made with homologous recombination as described elsewhere using vector pMQ30, yeast strain *S. cerevisiae* InvSc1 and *E. coli* mating strain ST18 (Hmelo et al., 2016). Briefly, pMQ30 deletion constructs were obtained by co-transforming cut pMQ30 plasmid and PCR-amplified regions flanking *P. aeruginosa* genes of interest into *S. cerevisiae* strain InvSc1. Reconstituted deletion plasmids were isolated, purified at greater concentration and transformed into *E. coli* strain ST18 for mating with *P. aeruginosa* PAO1. Genetic knock outs and allelic-exchange strains were confirmed with colony PCR. PaClpP1_{14}-His_6 and PaClpP2_7-StrepII were expressed with C-terminal polyhistidine or StrepII (IBA) or affinity tags in pET23b (Novagen). PaClpP1_7-His_6 P2_7-StrepII was overexpressed from pETDuet1 vector (Novagen). PaClpX was expressed with an N-terminal thrombin-cleavable polyhistidine tag. Plasmid for C-terminal fusion to MARTX toxin cysteine protease domain (CPD) from *Vibrio cholerae* was a gift from Adrian Olivares. Tagmaster Site-Directed Mutagenesis (GM Biosciences) was used to introduce point mutations.

Protein expression and purification

For protein induction, PaClpP1_{14}, PaClpP2_7 and PaClpP1_7ClpP2_7 variants were expressed with *E. coli* strain JK10 (Ec*clpP*-) grown at 37°C in Luria-Bertani broth supplemented with antibiotic. At OD_{600} of 0.6, temperature was reduced to either 30ºC (for PaClpP1_{14}-His_6 and PaClpP2_7-StrepII)

or 20ºC (for PaClpP1$_7$-His$_6$ P2$_7$-StrepII), 1mM isopropyl β-D-I-thiogalactopyranoside (IPTG) was added to 1 L culture and cells were harvested after 3 hours (PaClpP1$_{14}$-His$_6$ and PaClpP2$_7$-StrepII) or overnight (PaClpP1$_7$-His$_6$ P2$_7$-StrepII). For purification of PaClpP1$_7$-His$_6$ P2$_7$-StrepII, cell pellets were resuspended in Buffer S (50 mM sodium phosphate pH 8.0, 100 mM NaCl, 10% glycerol, 5 mM imidazole), lysed with French press, and cell lysate was mixed with 1% w/v CHAPS detergent for 30 mins at 4ºC. Lysate was then centrifuged 20 minutes at 30,000 rcf and supernatant was applied to Ni-NTA resin equilibrated in Buffer S, washed extensively with Buffer S + 20 mM imidazole and protein was eluted with Buffer S + 300 mM imidazole. Eluate was diluted 1:3 into Buffer W (100 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol) and applied to Strep-Tactin resin (IBA) equilibrated in Buffer W, and washed extensively with Buffer W. Protein was eluted with 2.5 mM Desthiobiotin (IBA) in Buffer W and concentrated by ultracentrifugation with a 30 kDa centrifugal filter (Amicon) to ~500 µL volume. Eluate was further purified with a Superdex200 16/600 HR gel filtration column (GE Healthcare) equilibrated in Buffer X (25 mM Tris pH 7.5, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, 10% glycerol). Fractions containing protein at >95% purity as judged by SDS-PAGE were combined, concentrated, flash frozen in aliquots and stored at -80ºC. For purification of PaClpP1$_{14}$-His$_6$ or PaClpP2$_7$-StrepII variants, the same experimental procedure as above was followed, except the tandem-affinity purification was omitted and either only the Ni-NTA resin was used (for PaClpP1$_{14}$-His$_6$) or only the Strep-Tactin resin was used (for PaClpP2$_7$-StrepII) followed by size-exclusion chromatography (see above). *P. aeruginosa* ClpA was overexpressed and purified with Ni-NTA, as described for PaClpP1$_{14}$-His. Protein was then incubated 2 h at 30°C with His-tagged Ulp1 to cleave between the SUMO domain and ClpA, eluted from Ni-NTA resin, dialyzed into buffer Q50 (50 mM Tris pH 8.0, 1 mM DTT, 10 mM $MgCl_2$, 10% glycerol, 50 mM KCl), and purified on a Mono Q column (GE Healthcare).

Selected fractions were then further purified by size exclusion in Buffer Q200 (50 mM Tris pH 8.0, 1mM DTT, 10 mM $MgCl_2$, 10% glycerol, 200 mM KCl). GFP with a C-terminal SsrA degron tag was expressed and purified as previously described (Dougan et al., 2002).

Enzymatic assays and crosslinking

Protein and peptide degradation assays were carried out in 25 mM HEPES-KOH (pH 7.6), 5 mM KCl, 5 mM $MgCl_2$, 0.032% NP-40, and 10% glycerol (PD buffer) at 30°C. Peptidase assays used 2.85% dimethyl sulfoxide (DMSO), 100 μM fluorogenic peptide (Bachem, Peptides International, or Enzo Life Sciences), and 0.4 μM ClpP tetradecamer ($PaClpP_{14}$), and degradation was monitored by fluorescence with excitation/emission 380/460 nm at 30ºC. Degradation of GFP with a C-terminal SsrA degron tag for generation of a K_M curve was monitored as previously described using 0.54 μM $PaClpP_{14}$, 0.1 μM $ClpX_6$ and a NADH-coupled ATP regeneration system (Nørby 1988, Oliveras et al., 2014). Enzymatic assays were monitored on a SpectraMax M5 microplate reader (Molecular Devices) in a black 384-well flat-bottom plate (Corning). For crosslinking of PaClpP complexes, 0.4 μM $PaClpP_{14}$ was mixed with 0.125% glutaraldehyde at room temperature and reactions were stopped by flash-freezing after indicated time points and analyzed via SDS-PAGE. Casein fluorescein isothiocyanate (FITC-casein) was purchased from Sigma-Aldrich.

Western blotting and co-immunoprecipitation

For co-immunoprecipitation of ClpP2-FLAG from *P. aeruginosa* strains containing a C-terminal Flag tag on *clpP2*, ClpP2-FLAG was isolated from anti-FLAG M2 magnetic beads (Sigma-Aldrich) isolated from cell lysate containing either *clpP2-1xFLAG* (+) or untagged clpP2 (-) alleles at the genomic loci, and eluted with excess 3xFLAG peptide (Sigma-Aldrich). The eluate was

separated by SDS-PAGE and proteins were stained with SYPRO orange (Sigma-Aldrich). The eluate was analyzed by Western blot for PaClpP1 (1:15,000 anti-PaClp1) and PaClpP2 (1:7,500 anti-PaClpP2) as previously described (Hall et al., 2017). Secondary antibody was anti-rabbit IgG-alkaline phosphatase (AP) conjugate (Bio-Rad) used at a 1:10,000 dilution. Blots were incubated with AP ECF substrate (GE Healthcare) and imaged with a Typhoon FLA 9500 scanner (GE Healthcare).

Crystallography and negative stain electron microscopy

For crystallization of PaClpP1, hanging drops with 0.5 µL of PaClpP1P2 protein at 6.6 mg/ml in storage buffer (0.05 M Tris HCl pH 7.5, 0.05 M NaCl, 5 mM EDTA, 25 mM DTT, 5% glycerol) were mixed with 0.5 µL well solution (0.1 M sodium malonate, 53.5% MPD) over a reservoir of 500 mL well solution at 20ºC. For crystallization of PaClpP2, hanging drops with 1.5 µL of 32.6 mg/ml ClpP2-StrepII in storage buffer (100 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol) and 1.5 µL of precipitant (1.8 M sodium phosphate monobasic monohydrate, potassium phosphate dibasic pH 6.9) were incubated over a reservoir of 500 ml precipitant at 20ºC. Crystals were soaked in precipitant + 10% ethylene glycol and frozen in liquid N_2. X-ray diffraction data were collected at the Advanced Photon Source (APS) beamline 24-ID-E with X-ray wavelength 0.9792 Å. For the ClpP2 crystal diffraction images were indexed, integrated and scaled using HKL2000. The structure was solved by molecular replacement with Phaser using a portion of *S. aureus* ClpP monomer (PDB 3QWD) as the search model. The asymmetric in the MR solution consisted of seven monomers in a heptameric ring. The ClpP2 amino-acid sequence was inserted with Autobuild in Phenix and the resulting model was refined using Coot and Phenix. For the ClpP1 structure the data were processed using XDS and solved by molecular replacement. The

search model was a heptameric ring derived from the *E. coli* ClpP structure (1YG6). The model was refined by iterative rebuilding in Coot and refinement using Phenix. Coordinates for both structures were deposited in the Protein Data Bank (5BQV and 5BQW). For electron microscopy, 0.5 μM $PaClpX_6$ and 0.2 μM $PaClpP_{14}$ were diluted in 25 mM Hepes pH 7.6, 100 mM KCl, 10 mM MgCl2, 0.5 mM TCEP, 2 mM ATPγS to achieve 18 ng/μL total protein. 5 μL of this mixture was spotted per copper grid (Electron Microscopy Sciences). Stain is 2% uranyl acetate. Images were obtained with the FEI Technai Spirit Transmission Electron microscope.

Mass spectrometry and data analysis

Protein digestion samples were resuspended to 1 mM in 2% formic acid. 1 μL was analyzed for each sample. The sample was loaded with the autosampler directly onto the column, which was made from a 50um ID 360um OD capillary (PicoFrit, New Objectives, Woburn, MA) filled with 25cm of 1.9um Reprosil-Pur C18 AQ. Peptides were eluted at 200 nL/min from the column using an EasyNanoLC (Thermo Scientific, San Jose, CA) with a gradient of 5% buffer B (80% acetonitrile, 0.1% formic acid) over 3 min and followed by 5% buffer B to 25% buffer B over 20 min. The gradient was then switched from 25% to 100% buffer B over 5 min and held constant for 3 min. Finally, the gradient was changed from 100% buffer B to 100% buffer A (100% water, 0.1% formic acid) over 0.5 min and then held constant for 13.5 min. The application of a 2.0 kV distal voltage electrosprayed the eluting peptides directly into the Thermo Orbitrap QEHFX mass spectrometer equipped with a Flex Source (Thermo Scientific). Mass spectrometer-scanning functions and HPLC gradients were controlled by the Xcalibur data system (Thermo Scientific). The mass spectrometer was set to scan MS1 at 60,000 resolution with an AGC target set at 3e6. The scan range was m/z 375-1500. For MS2, spectra were acquired with an AGC target of 1e6 and

a maximum IT at 200ms at 30,000 resolution. The top 5 peaks were analyzed by MS2. Peptides were isolated with an isolation window of m/z 1.2 and fragmented at 28 CE. Only ions with a charge state of 2 and 3 were considered for MS2. Dynamic exclusion was set at 10 sec. Mass spectrometry data was analyzed with ProteomeDiscoverer 2.3 and searched against a database comprising of the four protein sequences (FNR_ECOLI, GFP_ssrA, RARC_BPP22, HTitin_HUMAN) added to an *E. coli* database (released May 14, 2019) containing common contaminant proteins. No enzyme was specified and oxidized methionine as well as N-term acetylation was allowed as protein modification. Only high confidence peptides from the PSM table were analyzed further. Peptides not from the four proteins were eliminated as well as peptides that did not correspond to the protein analyzed in each sample.

Biofilm assays

P. aeruginosa strains were fluorescently tagged at an intergenic neutral chromosomal locus downstream of the *glmS* gene with *egfp* in mini-Tn*7* constructs as described previously (Lambertsen et al., 2004; Choi et al., 2006). Overnight culture of GFP-labeled strains were prepared in LB with shaking at 37°C. Overnight cultures were diluted in ABTGC(Chua et al., 2015) medium to an initial OD600 of 0.01, added to a glass-bottom 96-well plate, and incubated for 48 h at 37°C. Image acquisition was performed using a confocal laser scanning microscope (LSM 800; Zeiss) equipped with a 63×/1.4 NA oil immersion or a 100×/1.4 NA oil immersion objective. Images were analyzed with Zeiss ZEN 2.1 imaging software (Thornwood, NY, USA). The excitation wavelength for GFP was 488 nm. At least five stacks were recorded for each well and at least three independent wells were analyzed for each condition. Images were acquired with

a step size of 0.5 μm. Biofilm quantification was performed using COMSTAT (Heydorn et al., 2000).

References

Akopian, T., Kandror, O., Raju, R. M., Unnikrishnan, M., Rubin, E. J., & Goldberg, A. L. (2012). The active ClpP protease from M. tuberculosis is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO Journal*, *31*(6), 1529–1541. https://doi.org/10.1038/emboj.2012.5

Arevalo-Ferro, C., Hentzer, M., Reil, G., Görg, A., Kjelleberg, S., Givskov, M., … Eberl, L. (2003). Identification of quorum-sensing regulated proteins in the opportunistic pathogen Pseudomonas aeruginosa by proteomics. *Environmental Microbiology*, *5*(12), 1350–1369. https://doi.org/10.1046/j.1462-2920.2003.00532.x

Arnold, I., Wagner-Ecker, M., Ansorge, W., & Langer, T. (2006). Evidence for a novel mitochondria-to-nucleus signalling pathway in respiring cells lacking i-AAA protease and the ABC-transporter Mdl1. *Gene*, *367*(1–2), 74–88. https://doi.org/10.1016/j.gene.2005.09.044

Augustin, S., Nolden, M., Müller, S., Hardt, O., Arnold, L., & Langer, T. (2005). Characterization of peptides released from mitochondria: Evidence for constant proteolysis and peptide efflux. *Journal of Biological Chemistry*, *280*(4), 2691–2699. https://doi.org/10.1074/jbc.M410609200

Baker, T. A., & Sauer, R. T. (2012). ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1823*(1), 15–28. https://doi.org/10.1016/j.bbamcr.2011.06.007

Chua, S. L., Hultqvist, L. D., Yuan, M., Rybtke, M., Nielsen, T. E., Givskov, M., … Yang, L. (2015). In vitro and in vivo generation and characterization of Pseudomonas aeruginosa biofilm-dispersed cells via c-di-GMP manipulation. *Nat Protoc*, *10*(8), 1165–1180. https://doi.org/10.1038/nprot.2015.067

E, V. (2014). *Microarray Analysis of Pseudomonas aeruginosa Quorum-Sensing Regulons : Effects of Growth Phase and Environment Microarray Analysis of Pseudomonas aeruginosa Quorum-Sensing Regulons : Effects of Growth Phase and Environment †*. *185*(7), 2080–2095. https://doi.org/10.1128/JB.185.7.2080

Ferrington, D. A., & Gregerson, D. S. (2012). Immunoproteasomes: Structure, function, and antigen presentation. In *Progress in Molecular Biology and Translational Science* (Vol. 109). https://doi.org/10.1016/B978-0-12-397863-9.00003-1

Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T., & Baker, T. A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Molecular Cell*, *11*(3), 671–683. https://doi.org/10.1016/S1097-2765(03)00060-1

Gersch, M., Stahl, M., Poreba, M., Dahmen, M., Dziedzic, A., Drag, M., & Sieber, S. A. (2016). Barrel-shaped ClpP Proteases Display Attenuated Cleavage Specificities. *ACS Chemical Biology*, *11*(2), 389–399. https://doi.org/10.1021/acschembio.5b00757

Gilbert, K. B., Kim, T. H., Gupta, R., Greenberg, E. P., & Schuster, M. (2009). Global position analysis of the Pseudomonas aeruginosa quorum-sensing transcription factor LasR. *Molecular Microbiology*, *73*(6), 1072–1085. https://doi.org/10.1111/j.1365-2958.2009.06832.x

Hall, B. M., Breidenstein, E. B. M., de la Fuente-Núñez, C., Reffuveille, F., Mawla, G. D., Hancock, R. E. W., & Baker, T. A. (2017). Two isoforms of Clp peptidase in Pseudomonas aeruginosa control distinct aspects of cellular physiology. *Journal of Bacteriology*, *199*(3), 1–15. https://doi.org/10.1128/JB.00568-16

Harris, J. L., Alper, P. B., Li, J., Rechsteiner, M., & Backes, B. J. (2001). Substrate specificity of the human proteasome. *Chemistry and Biology*, *8*(12), 1131–1141. https://doi.org/10.1016/S1074-5521(01)00080-1

Haynes, C. M., Petrova, K., Benedetti, C., Yang, Y., & Ron, D. (2007). ClpP Mediates Activation of a Mitochondrial Unfolded Protein Response in C. elegans. *Developmental Cell*, *13*(4), 467–480. https://doi.org/10.1016/j.devcel.2007.07.016

Haynes, C. M., Yang, Y., Blais, S. P., Neubert, T. A., & Ron, D. (2010). The Matrix Peptide Exporter HAF-1 Signals a Mitochondrial UPR by Activating the Transcription Factor ZC376.7 in C. elegans. *Molecular Cell*, *37*(4), 529–540. https://doi.org/10.1016/j.molcel.2010.01.015

Herget, M., & Tampé, R. (2007). Intracellular peptide transporters in human - Compartmentalization of the "peptidome." *Pflugers Archiv European Journal of Physiology*, *453*(5), 591–600. https://doi.org/10.1007/s00424-006-0083-4

Hong, Z., Bolard, A., Giraud, C., Prévost, S., Genta-Jouve, G., Deregnaucourt, C., … Li, Y. (2019). Azetidine-Containing Alkaloids Produced by a Quorum-Sensing Regulated Nonribosomal Peptide Synthetase Pathway in Pseudomonas aeruginosa. *Angewandte Chemie - International Edition*, *58*(10), 3178–3182. https://doi.org/10.1002/anie.201809981

Kim, Y. I., Burton, R. E., Burton, B. M., Sauer, R. T., & Baker, T. A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. *Molecular Cell*, *5*(4), 639–648. https://doi.org/10.1016/S1097-2765(00)80243-9

Kisselev, A. F., Garcia-Calvo, M., Overkleeft, H. S., Peterson, E., Pennington, M. W., Ploegh, H. L., … Goldberg, A. L. (2003). The caspase-like sites of proteasomes, their substrate specificity, new inhibitors and substrates, and allosteric interactions with the trypsin-like sites. *Journal of Biological Chemistry*, *278*(38), 35869–35877. https://doi.org/10.1074/jbc.M303725200

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., … Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, *47*(W1), W636–W641. https://doi.org/10.1093/nar/gkz268

Melber, A., & Haynes, C. M. (2018). UPR mt regulation and output: A stress response mediated by mitochondrial-nuclear communication. *Cell Research*, *28*(3), 281–295. https://doi.org/10.1038/cr.2018.16

Patteson, J. B., Lescallette, A. R., & Li, B. (2019). Discovery and Biosynthesis of Azabicyclene, a Conserved Nonribosomal Peptide in Pseudomonas aeruginosa. *Organic Letters*, *21*(13), 4955–4959. https://doi.org/10.1021/acs.orglett.9b01383

Rutherford, S. T., & Bassler, B. L. (2012). Bacterial quorum sensing: Its role in virulence and possibilities for its control. *Cold Spring Harbor Perspectives in Medicine*, *2*(11), 1–25. https://doi.org/10.1101/cshperspect.a012427

Schmitza, K. R., Carneyb, D. W., Sellob, J. K., & Sauera, R. T. (2014). Crystal structure of mycobacterium tuberculosis ClpP1p2 suggests a model for peptidase activation by aaa+ partner binding and substrate delivery. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(43), E4587–E4595. https://doi.org/10.1073/pnas.1417120111

Schuster, M., Lostroh, C. P., Ogi, T., & Greenberg, E. P. (2003). (4916) Identification, Timing, and Signal Speci city of. *Society*, *185*(7), 2066–2079. https://doi.org/10.1128/JB.185.7.2066

Sheps, J. A., Ralph, S., Zhao, Z., Baillie, D. L., & Ling, V. (2004). The ABC transporter gene family of Caenorhabditis elegans has implications for the evolutionary dynamics of

multidrug resistance in eukaryotes. *Genome Biology*, *5*(3), 1–17. https://doi.org/10.1186/gb-2004-5-3-r15

St-Pierre, C., Morgand, E., Benhammadi, M., Rouette, A., Hardy, M. P., Gaboury, L., & Perreault, C. (2017). Immunoproteasomes Control the Homeostasis of Medullary Thymic Epithelial Cells by Alleviating Proteotoxic Stress. *Cell Reports*, *21*(9), 2558–2570. https://doi.org/10.1016/j.celrep.2017.10.121

Thompson, M. W., & Maurizi, M. R. (1994). Activity and specificity of Escherichia coli ClpAP protease in cleaving model peptide substrates. *Journal of Biological Chemistry*, *269*(27), 18201–18208.

Vahidi, S., Ripstein, Z. A., Juravsky, J. B., Rennella, E., Goldberg, A. L., Mittermaier, A. K., … Kay, L. E. (2020). An allosteric switch regulates Mycobacterium tuberculosis ClpP1P2 protease function as established by cryo-EM and methyl-TROSY NMR. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(11), 5895–5906. https://doi.org/10.1073/pnas.1921630117

Yewdell, J. W. (2005). Immunoproteasomes: Regulating the regulator. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(26), 9089–9090. https://doi.org/10.1073/pnas.0504018102

Young, L., Leonhard, K., Tatsuta, T., Trowsdale, J., & Langer, T. (2001). Role of the ABC transporter Mdl1 in peptide export from mitochondria. *Science*, *291*(5511), 2135–2138. https://doi.org/10.1126/science.1056957

Yu, A. Y. H., & Houry, W. A. (2007). ClpP: A distinctive family of cylindrical energy-dependent serine proteases. *FEBS Letters*, *581*(19), 3749–3757. https://doi.org/10.1016/j.febslet.2007.04.076

Akopian, T., Kandror, O., Raju, R. M., Unnikrishnan, M., Rubin, E. J., & Goldberg, A. L. (2012). The active ClpP protease from M. tuberculosis is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO Journal*, *31*(6), 1529–1541. https://doi.org/10.1038/emboj.2012.5

Arevalo-Ferro, C., Hentzer, M., Reil, G., Görg, A., Kjelleberg, S., Givskov, M., … Eberl, L. (2003). Identification of quorum-sensing regulated proteins in the opportunistic pathogen Pseudomonas aeruginosa by proteomics. *Environmental Microbiology*, *5*(12), 1350–1369. https://doi.org/10.1046/j.1462-2920.2003.00532.x

Arnold, I., Wagner-Ecker, M., Ansorge, W., & Langer, T. (2006). Evidence for a novel mitochondria-to-nucleus signalling pathway in respiring cells lacking i-AAA protease and the ABC-transporter Mdl1. *Gene*, *367*(1–2), 74–88. https://doi.org/10.1016/j.gene.2005.09.044

Augustin, S., Nolden, M., Müller, S., Hardt, O., Arnold, L., & Langer, T. (2005). Characterization of peptides released from mitochondria: Evidence for constant proteolysis and peptide efflux. *Journal of Biological Chemistry*, *280*(4), 2691–2699. https://doi.org/10.1074/jbc.M410609200

Baker, T. A., & Sauer, R. T. (2012). ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1823*(1), 15–28. https://doi.org/10.1016/j.bbamcr.2011.06.007

Chua, S. L., Hultqvist, L. D., Yuan, M., Rybtke, M., Nielsen, T. E., Givskov, M., … Yang, L. (2015). In vitro and in vivo generation and characterization of Pseudomonas aeruginosa biofilm-dispersed cells via c-di-GMP manipulation. *Nat Protoc*, *10*(8), 1165–1180. https://doi.org/10.1038/nprot.2015.067

E, V. (2014). *Microarray Analysis of Pseudomonas aeruginosa Quorum-Sensing Regulons : Effects of Growth Phase and Environment Microarray Analysis of Pseudomonas aeruginosa Quorum-Sensing Regulons : Effects of Growth Phase and Environment †*. *185*(7), 2080–2095. https://doi.org/10.1128/JB.185.7.2080

Ferrington, D. A., & Gregerson, D. S. (2012). Immunoproteasomes: Structure, function, and antigen presentation. In *Progress in Molecular Biology and Translational Science* (Vol. 109). https://doi.org/10.1016/B978-0-12-397863-9.00003-1

Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T., & Baker, T. A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Molecular Cell*, *11*(3), 671–683. https://doi.org/10.1016/S1097-2765(03)00060-1

Gersch, M., Stahl, M., Poreba, M., Dahmen, M., Dziedzic, A., Drag, M., & Sieber, S. A. (2016). Barrel-shaped ClpP Proteases Display Attenuated Cleavage Specificities. *ACS Chemical Biology*, *11*(2), 389–399. https://doi.org/10.1021/acschembio.5b00757

Gilbert, K. B., Kim, T. H., Gupta, R., Greenberg, E. P., & Schuster, M. (2009). Global position analysis of the Pseudomonas aeruginosa quorum-sensing transcription factor LasR. *Molecular Microbiology*, *73*(6), 1072–1085. https://doi.org/10.1111/j.1365-2958.2009.06832.x

Hall, B. M., Breidenstein, E. B. M., de la Fuente-Núñez, C., Reffuveille, F., Mawla, G. D., Hancock, R. E. W., & Baker, T. A. (2017). Two isoforms of Clp peptidase in Pseudomonas aeruginosa control distinct aspects of cellular physiology. *Journal of Bacteriology*, *199*(3), 1–15. https://doi.org/10.1128/JB.00568-16

Harris, J. L., Alper, P. B., Li, J., Rechsteiner, M., & Backes, B. J. (2001). Substrate specificity of the human proteasome. *Chemistry and Biology*, *8*(12), 1131–1141. https://doi.org/10.1016/S1074-5521(01)00080-1

Haynes, C. M., Petrova, K., Benedetti, C., Yang, Y., & Ron, D. (2007). ClpP Mediates Activation of a Mitochondrial Unfolded Protein Response in C. elegans. *Developmental Cell*, *13*(4), 467–480. https://doi.org/10.1016/j.devcel.2007.07.016

Haynes, C. M., Yang, Y., Blais, S. P., Neubert, T. A., & Ron, D. (2010). The Matrix Peptide Exporter HAF-1 Signals a Mitochondrial UPR by Activating the Transcription Factor ZC376.7 in C. elegans. *Molecular Cell*, *37*(4), 529–540. https://doi.org/10.1016/j.molcel.2010.01.015

Herget, M., & Tampé, R. (2007). Intracellular peptide transporters in human - Compartmentalization of the "peptidome." *Pflugers Archiv European Journal of Physiology*, *453*(5), 591–600. https://doi.org/10.1007/s00424-006-0083-4

Hong, Z., Bolard, A., Giraud, C., Prévost, S., Genta-Jouve, G., Deregnaucourt, C., … Li, Y. (2019). Azetidine-Containing Alkaloids Produced by a Quorum-Sensing Regulated Nonribosomal Peptide Synthetase Pathway in Pseudomonas aeruginosa. *Angewandte Chemie - International Edition*, *58*(10), 3178–3182. https://doi.org/10.1002/anie.201809981

Kim, Y. I., Burton, R. E., Burton, B. M., Sauer, R. T., & Baker, T. A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. *Molecular Cell*, *5*(4), 639–648. https://doi.org/10.1016/S1097-2765(00)80243-9

Kisselev, A. F., Garcia-Calvo, M., Overkleeft, H. S., Peterson, E., Pennington, M. W., Ploegh, H. L., … Goldberg, A. L. (2003). The caspase-like sites of proteasomes, their substrate specificity, new inhibitors and substrates, and allosteric interactions with the trypsin-like sites. *Journal of Biological Chemistry*, *278*(38), 35869–35877. https://doi.org/10.1074/jbc.M303725200

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., … Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, *47*(W1), W636–W641. https://doi.org/10.1093/nar/gkz268

Melber, A., & Haynes, C. M. (2018). UPR mt regulation and output: A stress response mediated by mitochondrial-nuclear communication. *Cell Research*, *28*(3), 281–295. https://doi.org/10.1038/cr.2018.16

Patteson, J. B., Lescallette, A. R., & Li, B. (2019). Discovery and Biosynthesis of Azabicyclene, a Conserved Nonribosomal Peptide in Pseudomonas aeruginosa. *Organic Letters*, *21*(13), 4955–4959. https://doi.org/10.1021/acs.orglett.9b01383

Rutherford, S. T., & Bassler, B. L. (2012). Bacterial quorum sensing: Its role in virulence and possibilities for its control. *Cold Spring Harbor Perspectives in Medicine*, *2*(11), 1–25. https://doi.org/10.1101/cshperspect.a012427

Schmitza, K. R., Carneyb, D. W., Sellob, J. K., & Sauera, R. T. (2014). Crystal structure of mycobacterium tuberculosis ClpP1p2 suggests a model for peptidase activation by aaa+ partner binding and substrate delivery. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(43), E4587–E4595. https://doi.org/10.1073/pnas.1417120111

Schuster, M., Lostroh, C. P., Ogi, T., & Greenberg, E. P. (2003). (4916) Identification, Timing, and Signal Speci city of. *Society*, *185*(7), 2066–2079. https://doi.org/10.1128/JB.185.7.2066

Sheps, J. A., Ralph, S., Zhao, Z., Baillie, D. L., & Ling, V. (2004). The ABC transporter gene family of Caenorhabditis elegans has implications for the evolutionary dynamics of multidrug resistance in eukaryotes. *Genome Biology*, *5*(3), 1–17. https://doi.org/10.1186/gb-2004-5-3-r15

St-Pierre, C., Morgand, E., Benhammadi, M., Rouette, A., Hardy, M. P., Gaboury, L., & Perreault, C. (2017). Immunoproteasomes Control the Homeostasis of Medullary Thymic Epithelial Cells by Alleviating Proteotoxic Stress. *Cell Reports*, *21*(9), 2558–2570. https://doi.org/10.1016/j.celrep.2017.10.121

Thompson, M. W., & Maurizi, M. R. (1994). Activity and specificity of Escherichia coli ClpAP protease in cleaving model peptide substrates. *Journal of Biological Chemistry*, *269*(27), 18201–18208.

Vahidi, S., Ripstein, Z. A., Juravsky, J. B., Rennella, E., Goldberg, A. L., Mittermaier, A. K., … Kay, L. E. (2020). An allosteric switch regulates Mycobacterium tuberculosis ClpP1P2 protease function as established by cryo-EM and methyl-TROSY NMR. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(11), 5895–5906. https://doi.org/10.1073/pnas.1921630117

Yewdell, J. W. (2005). Immunoproteasomes: Regulating the regulator. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(26), 9089–9090. https://doi.org/10.1073/pnas.0504018102

Young, L., Leonhard, K., Tatsuta, T., Trowsdale, J., & Langer, T. (2001). Role of the ABC transporter Mdl1 in peptide export from mitochondria. *Science*, *291*(5511), 2135–2138. https://doi.org/10.1126/science.1056957

Yu, A. Y. H., & Houry, W. A. (2007). ClpP: A distinctive family of cylindrical energy-dependent serine proteases. *FEBS Letters*, *581*(19), 3749–3757. https://doi.org/10.1016/j.febslet.2007.04.076

Chapter 3:

Discussion and Future Directions

Overview

In Chapter 2, I presented evidence that *P. aeruginosa* makes an alternative hetero-ClpP complex during stationary-phase growth ($PaClpP1_7P2_7$), and that the catalytic serine residues of PaClpP2 are specifically important for the development of biofilms. In this chapter, I will discuss the biological implications of this work, focusing on how alternative proteolytic components and products may regulate and inform cellular responses. Finally, I will suggest some experimental approaches that could extend these studies in future work.

Why might cells produce alternative proteolytic subunits?

Why might cells benefit from making alternative peptidase subunits? One reason could be for the generation of diverse peptide products with downstream cellular functions. In this scenario, what we normally would consider to be "waste products" of proteolysis (i.e. products from the breakdown of polypeptides), would instead be functional molecules. An example of this phenomenon occurs with the immunoproteasome (i-proteasome), where in jawed vertebrates, immune cells produce alternative active β subunits of the proteasome 20S particle to generate peptides for antigen presentation (refer to Chapter 1 for a brief introduction on the structure of the eukaryotic proteasome). Effectively, i-proteasomes generate a different spectrum of peptides from standard proteasomes, and this diversity is benefited by adaptive immune system responses. The subunit organization, proteolytic activities, and peptide product generation of the two proteasome variants will be discussed below.

The standard proteasome catalytic subunits, β1, β2, and β5 are constitutively expressed in eukaryotic cells, with each subunit employing a different catalytic activity: caspase-like, trypsin-

like, and chymotrypsin-like, for cleavage after acidic, basic, and hydrophobic amino acids, respectively. Conversely, in the i-proteasome, these catalytic subunits are replaced by inducible subunits LMP2, MECL-1, and LMP7. Whereas the MECL-1 and LMP7 have the same type of catalytic activities as the β2 (basic preferred) and β5 subunits (hydrophobic preferred), the LMP2 subunit does not parallel β1 (cleaving after acidic residues), but rather cleaves after hydrophobic amino acids. Thus, cleavage after acidic residues is practically abolished in the i-proteasome. The direct outcome of replacing a caspase-like with a chymotrypsin-like subunit in the i-proteasome is a shift towards the more abundant production of peptides with hydrophobic C-termini, which act as major histocompatibility complex (MHC) class I-compatible peptides. Interestingly, a proteasome does not have to be composed of "all constitutive" or "all inducible" subunits. Instead, given that intermediate proteasomes can contain any mixture of constitutive and inducible subunits, it is theoretically possible to generate up to 36 different 20S core particles.

It is important to note that the generation of MHC class I presenting peptides is achieved in all nucleated vertebrate cells by the standard proteasome. As discussed in Chapter 1, the β subunits of the eukaryotic proteasome are not identical but homologous. The range of cleavage specificities of the standard β subunits generate wide diversity in the resulting peptidome. This diversity is advantageous because it sets up a "surveillance scheme" by which the immune system can monitor what proteins individual cells are making, enabling host CD8 T cells to rapidly elicit an immune response in the event of viral infection (since viruses use cellular ribosomes to propagate). Infection leads to the synthesis of new and foreign proteins, and thus formation of corresponding new peptides. In response to cytokines (such as IFN-γ or IFN-α) at early viral infection, cells rapidly upregulate the expression of i-proteasome subunits, and as an effect, diversify their

peptidome as to be more suitable for display by the peptide-recognition proteins key to the immune system.

A function for *clpP2* in *P. aeruginosa*

What is the function of *clpP2* in *P. aeruginosa*? Our previous study established that whereas PaClpP1 purifies as a fully-assembled 14-mer with peptidase and protease activity when partnered to PaClpX or PaClpA, PaClpP2 on its own assembles as a 7-mer with no detectable enzymatic activity (Hall et al., 2017). In Chapter 2, I show for the first time that PaClpP2 is a functional peptidase when co-complexed with PaClpP1, and that its catalytic sites are specifically important for the development of biofilms. Based on this work, I propose that there are two functional ClpP peptidase complexes in *P. aeruginosa*: $ClpP1_{14}$ and $ClpP1_7P2_7$. Genome-wide expression data show that whereas *clpP1* gene product levels are stable throughout the various stages of growth and development (refer to Fig. 1-12), PaClpP2 mRNA and protein levels are highly upregulated upon quorum sensing (Wagner et al., 2003; Schuster et al., 2003; Arevalo-Ferro et al., 2003). This induction of *clpP2* is directly dependent on the quorum-sensing master transcription factor, LasR (Arevalo-Ferro et al., 2003; Gilbert et al., 2009). Therefore, based on expression and biochemical data, it seems likely that $PaClpP1_{14}$ is the dominant peptidase in the absence of quorum-sensing molecules (during log-phase growth and in planktonic cells) and interacts symmetrically with PaClpX and PaClpA to degrade protein substrates (Hall et al., 2017 and this study). However, upon entry into stationary-phase growth and in response to quorum-sensing, *clpP2* is upregulated, and $PaClpP1_7P2_7$ complexes accumulate.

Why might have *P. aeruginosa* evolved to make a specialized, quorum-sensing-specific ClpP? By analogy to the eukaryotic proteasome during antigen presentation, one possibility is that PaClpP1P2 contributes to the diversity of the peptidome by generating alternatively-cleaved peptides. This change in the profile of peptide cleavage products could then serve in an as-of-yet undiscovered pathway with downstream effects on biofilm development. Such a positive role for peptides specifically generated by cleavage by PaClpP2 is supported by the observation that an active site knockout allele of *clpP2* (*clpP2*^{SA}) results in thinner biofilms (Fig. 2-11).

One formal, but unlikely, possibility that cannot be completely excluded in this model is that PaClpP1$_7$P2$_7$ may interact non-canonically with a currently uncharacterized AAA+ unfoldase with alternative substrates for proteolysis. Through rigorous biochemical and mutational analysis, I have established that a functional interaction with the canonical AAA+ unfoldases PaClpX (and likely PaClpA) is not observed with the PaClpP2 ring of PaClpP1$_7$P2$_7$ (Fig. 2-9, 2-10). This is supported by two structural features: (i) the divergent N-terminal sequence of PaClpP2 (which is important for interacting with the pore-2 loops of AAA+ unfoldases), and (ii) the C-terminal extension on PaClpP2 that sterically antagonizes access to the hydrophobic binding pockets of PaClpP2 (the region where the IGL/IGF loops on ClpX/A dock to) (Fig. 2-5, 2-7). Even if there was a C-terminal processing event *in vivo* that relieved access to the docking pockets, it likely would not be sufficient for allowing productive interaction with AAA+ unfoldases, since the identity of N-terminal residues would still not be compatible with interacting with the pore-2 loops of unfoldases. Even so, a mechanism by which PaClpP1P2 binds to a currently unknown AAA+ unfoldase with unique substrate specificity would still not explain the *clpP2*^{SA} phenotype, since the proteolytic sites of PaClpP1 would most likely be able to compensate for inactive PaClpP2 in

"mystery unfoldase•PaClpP1$_7$P2$^{SA}{}_7$" complexes if general degradation activity was all that was needed (Fig. 2-12, 2-13, 2-14). Therefore, instead we favor models where PaClpP2, in complex with PaClpP1, produces a peptide or peptides with positive downstream effects on biofilm formation (Fig. 2-18).

Proteolysis-derived peptide products as cellular sensors

In Chapter 2, I provide evidence that the active sites of PaClpP2 perform a function that promotes biofilm thickness. This result, together with my analysis that reveals that PaClpP1 and PaClpP2 have distinguishable features in their cleavage-site specificity, lead me to favor models that posit the peptide products of PaClpP1$_7$P2$_7$ are functional molecules. These functional peptides may perhaps be generated by the hetero-peptidase acting alone, or as part of a PaClpX/PaClpA-containing protease (see Fig. 2-18). Typically, ClpP partnered with an AAA+ unfoldase degrades proteins to small oligopeptides that are then processed to amino acids by cytoplasmic peptidases (Yu and Houry, 2007). However, a functional role for protease-generated peptides in biological systems is not unprecedented, as already discussed with the example of the i-proteosome. In the section below, I will draw upon examples of functional proteolytic products from select model organisms, to support my ideas that the PaClpP1$_7$P2$_7$ cleavage product(s) have a biological/regulatory function. Furthermore, I explain how some genome mining analysis encourages me to also suggest that there may be a ClpP-dependent peptide export pathway in Gram-negative bacteria.

In the mitochondrial matrix of *Saccharomyces cerevisiae* (*S. cerevisiae*), peptides derived from protein degradation by the m-AAA protease (which is composed of Yta10 and Yta12 subunits; m stands for matrix) are actively transported into the intermembrane space; once out of the mitochondria some of these peptides function to regulate nuclear expression of genes involved in oxidative phosphorylation, and thus may be key communication molecules between mitochondria and other cellular components (Young et al., 2001; Augustin et al., 2005; Arnold et al., 2006). Worth noting in this context is that Mdl1p, the yeast mitochondrial peptide transporter that functions in the above pathway, is homologous to the Transporters associated with Antigen Presentation (TAP transporters), which, in mammals, TAP transports peptides from the cytosol into the endoplasmic reticulum for assembly onto MHC class I complexes in mammals (Herget and Tampé, 2007). Both Mdlp1 and TAP are ATP-binding cassette (ABC) transporters with a transmembrane region and a single ATP-binding cassette (Sheps et al., 2004). The existence and importance of these peptide transporters stronger supports the idea that purposeful peptide trafficking events are conserved from yeast to mammals.

Another related example in which protein degradation products are useful to the cell occurs during the mitochondrial unfolded protein response (UPR^{mt}) in the nematode, *Caenorhabditis elegans* (*C. elegans*). When the protein-folding environment in the mitochondria is perturbed, nuclear genes encoding mitochondrial matrix chaperones are induced in what is called UPR^{mt} (Melber and Haynes, 2018). A protease homologous to bacterial ClpP (*ZK970.2*) was first identified in a genome-wide screen for genes involved in UPR^{mt} (Haynes et al., 2007). It was later discovered that the mitochondria-localized peptide transporter, HAF-1, is responsible for the efflux of stress-induced, ClpP-derived peptides out of the mitochondria and into the cytosol, where they are then

transported into the nucleus and signal the UPR^{mt} response (Haynes et al., 2010). Notably, HAF-1 is also homologous to the yeast mitochondrial peptide transporter, Mdl1p, and to human TAP.

Do bacteria possess similar peptide-based signaling mechanisms, whereby proteolysis-derived peptides take center stage? Analogous to acylated homoserine lactones as quorum sensing molecules in Gram-negative bacteria (refer to Chapter 1), Gram-positive bacteria use processed oligopeptides to communicate (Autoinducer Peptides; AIP). AIPs are 5-17 amino acids in length and are processed by the same ABC transporters that transport them out of the cell (Rutherford and Bassler, 2012). Yet, to date, there have been no reports of a ClpP-dependent peptide production pathway for cellular signaling. My work on the PaClpP1$_7$P2$_7$ heterocomplex suggests the existence of such a pathway in *P. aeruginosa*. Interestingly, BLAST search queries for homologs of the human TAP transporter in *P. aeruginosa* strain PAO1 reveal a promising candidate. PA1113 is annotated as a probable "ATP-binding/permease fusion ABC transporter" with greater than 30% sequence identity to human TAP2, nematode HAF-1, and yeast Mdl1p (pseudomonas.com, Table 3-1).

	PA1113 (*P. aeruginosa*)	TAP2 (*H. sapiens*)	HAF-1 (*C. elegans*)	Mdl1p (*S. cerevisiae*)
PA1113	100	33.6	31.2	31.6
TAP2	33.6	100	28.1	28.7
HAF-1	31.2	28.1	100	32.5
Mdl1p	31.6	28.7	32.5	100

Table 3-1: Percent identity matrix of select homologs of the human TAP2 transporter.
PA1113, TAP2, HAF-1, and Mdl1p are homologous proteins, sharing 28.1-33.6% sequence identity. Percent identity matrix was calculated using alignments made in Clustal 2.1 (Madeira et al., 2019).

Given the functional links between the mammalian proteasome and the TAP transporter, yeast mitochondrial m-AAA protease and Mdl1p transporter, and nematode ClpP and the HAF-1

transporter, it is tempting (although still only speculation) to draw a similar parallel in the Gram-negative bacteria *P. aeruginosa* between PaClpP1$_7$P2$_7$ and an ABC transporter (perhaps PA1113). To establish such a link, further experiments are warranted. A pressing biological question that will motivate related future studies is: are PaClpP2-dependent peptides exported from cells? Tandem mass-spectrometry on supernatants of *P. aeruginosa* cultures (e.g. comparing wild-type with a *clpP2* null) to identify specific excreted peptides would be an important means to start to probe this line of research. As an initial approach to determine whether PA1113 plays a role in this pathway, biofilms cultured from genetic knockouts of PA1113 will need to be assessed for thickness and compared to those prepared from wild-type and *clpP2*^{SA} cells (recall from Chapter 2 that *clpP2*^{SA} biofilms are thinner than wildtype). If biofilms from PA1113 knockouts phenocopy those of *clpP2*^{SA}, then that would be preliminary evidence consistent with a ClpP-dependent peptide export pathway important in *P. aeruginosa* development.

On the peptidase activity of PaClpP1$_7$P2$_7$

Although the PaClpP1$_7$P2$_7$ heterocomplex hydrolyzes the small peptide substrates carboxybenzyl-Leu-Leu-Leu-7-amino-4-methylcoumarin (Z-LLL-AMC) and carboxybenzyl-Gly-Gly-Leu-AMC (Z-GGL-AMC) convincingly faster than does PaClpP1$_{14}$ (Fig. 2-12), this sequence specificity is not clearly reflected in the products generated in large-scale protein degradation experiments (refer to WebLogos in Fig. 2-14). This apparent dampening in cleavage specificity when comparing peptide cleavage to proteolysis of proteins with the assistance of a AAA+ unfoldase is thought to be due to the high local concentration of active sites in ClpP [~350 mM] during proteolysis. Under these proteolysis conditions, ClpX or ClpA is actively pumping protein into the chamber, whereas

during peptide hydrolysis, the concentration of substrate in the ClpP chamber depends on its ability to diffuse into ClpP (Baker and Sauer, 2012; Gersch et al., 2015). However, AAA+ ATPase-independent ClpP activity is still a useful means for determining the relative activities of different types of active sites within proteolytic complexes (Thompson and Maurizi, 1994; Harris et al., 2001; Kisselev et al., 2003).

Thus, it is useful to consider the underlying mechanisms explaining why PaClpP1$_7$P2$_7$ is more active against Z-LLL-AMC and Z-GGL-AMC compared to PaClpP1$_{14}$. A reasonable possibility is that due to structural differences between the S1 binding pocket of PaClpP1 and PaClpP2, binding of different peptides, especially those with different residues at the peptide's P1 site would be favored (Fig. 2-19). An alternative, less obvious, possibility is that the peptides themselves are acting as allosteric activators of the heterocomplex PaClpP1$_7$P2$_7$. This type of model is worth considering since we know that in the actinobacterium *M. tuberculosis*, MtClpP1P2 heterocomplex activity requires the binding of activator peptides, such as benzoyl-leucyl-leucine (Bz-LL). In this enzyme, peptide agonists directly bind all 14 serine active sites of MtClpP1P2; the outcome of an agonist-bound MtClpP1P2 is a more stable and more catalytically active ClpP heterocomplex (Akopian et al., 2012; Schmitz et al., 2014; Gersch et al., 2015; Vahidi et al., 2020). It is curious that out of the 18 peptide substrates assayed for activity against PaClpP1$_{14}$ and PaClpP1$_7$P2$_7$, the two that were cleaved faster by the heterocomplex, Z-LLL-AMC and Z-GGL-AMC, are also the most similar to the dipeptide that stabilizes and activates MtClpP1P2. Future experiments will need to be done to address whether peptide agonists bind and activate PaClpP1$_7$P2$_7$ in a manner reminiscent to that observed for MtClpP1P2.

Future directions

Work presented in this thesis establishes PaClpP1$_7$P2$_7$ as a *bona fide* ClpP peptidase in *P. aeruginosa* with a positive role in biofilm formation. The co-complex deviates from canonical ClpPs in that, first, it is composed of two different types of subunits (PaClpP1 and PaClpP2), second, it contains different peptidase activity/specificity than the canonical form, and third, AAA+ unfoldases only bind to one side of the co-complex. PaClpP1$_7$P2$_7$ has a specialized role in the cell, as it is induced by quorum sensing and in stationary phase cells and plays a positive role in the regulation of biofilm architecture. Here, I propose future directions I find interesting regarding this project. Topics that will be addressed in this section are focused around the structure, stability, and substrates of PaClpP1$_7$P2$_7$.

Although we were able to solve the crystal structures of PaClpP1 and PaClpP2 individually, the structure of PaClpP1$_7$P2$_7$ remains elusive despite significant attempts by the author. The few, rare PaClpP1$_7$P2$_7$ crystals that formed diffracted poorly, or in one case, turned out to be purely PaClpP1$_{14}$, indicating that PaClpP1$_7$P2$_7$ may not exist in a unique conformation and that inter-subunit rearrangements can occur over the time course of crystallization. Importantly, however, given this observation, we established using multiple affinity-capture approaches that PaClpP1$_7$P2$_7$ does not detectably dissociate and rearrange to form homo-complexes at least over the time of our enzymatic experiments (Fig. 3-1).

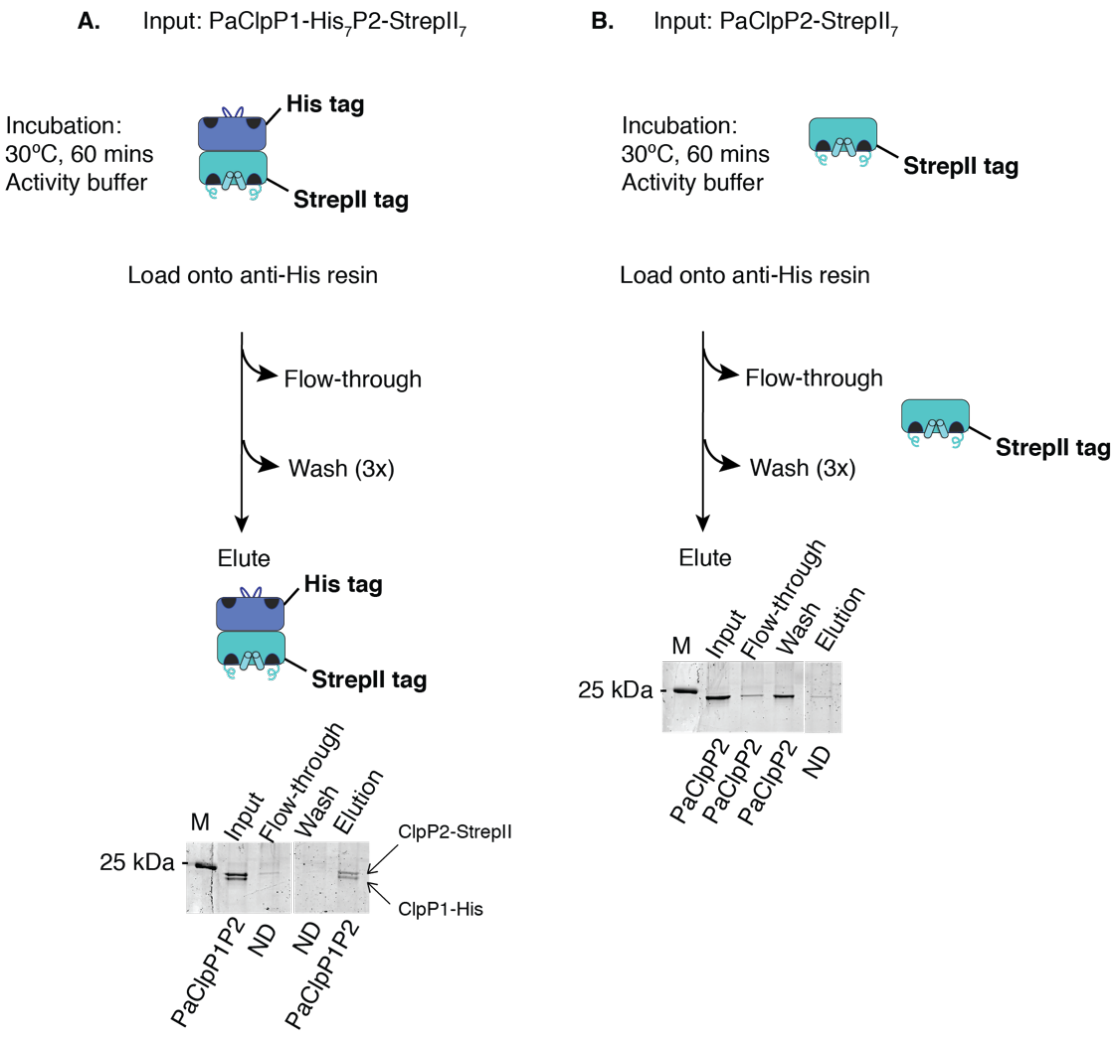

Figure 3-1: PaClpP2$_7$ remains associated with PaClpP1$_7$ under experimental reaction conditions.
A) PaClpP1$_7$P2$_7$ was incubated in PD reaction buffer (see Materials and Methods) at 30ºC for one hour to mimic a biochemical reaction. After incubation, the mixture was added to His-tag Dynabeads (Life Technologies) to select for complexes containing PaClpP1-6xHis$_7$. Samples corresponding to input, flow-through, wash, and elution were analyzed for the presence of the two ClpP isoforms. Bands corresponding to an approximately 1:1 ratio of ClpP1 and ClpP2 were detected by SDS-PAGE in the final elutes, but no ClpP1 was detected in the His-column flow-through or wash, suggesting that subunit dissociation between ClpP1 and ClpP2 and assembly of ClpP1 14-mers does not occur to a detectable extent under reaction conditions. B) Control showing Clp2-StrepII$_7$ alone does not stably-bind His-tag beads. Procedure is same as in (A), except purified Clp2-StrepII$_7$ was incubated with the resin. In (A) and (B), "ND" stands for "not detected."

The inability to form high-resolution PaClpP1$_7$P2$_7$ crystals could indicate that the co-complex is present in a structurally-heterogenous population in solution which is incompatible with the formation of crystalline lattices. To get around this barrier, one could take a cryogenics approach coupled to electron microcopy (cryo-EM). By this approach, it would be interesting to see whether, and to what extent, PaClpP1$_7$P2$_7$ adopts an "extended" form (a form associated with active ClpPs; see Chapter 2) in solution. Particular attention should be paid to the conformation(s) of the C-terminal ~8 residues of PaClpP2, which could inform future models of PaClpP1$_7$P2$_7$ function. For example, if the PaClpP2 C-termini of heterocomplexes become disordered or are no longer structured loops blocking the hydrophobic pockets (as seen in our current PaClpP2 structure), this conformation change could point to a mechanism in which an alternative unfoldase docks to the PaClpP2 side of the heterocomplex.

How does PaClpP1$_7$P2$_7$ co-assemble in *P. aeruginosa*? PaClpP1$_7$P2$_7$ formation from *E. coli* depends on the co-overexpression of PaClpP1 and PaClpP2 from a common vector (Chapter 2), suggesting that PaClpP1 and PaClpP2 monomers need to co-assemble during or right after translation to form in *E. coli*. Multiple types of mixing experiments using the two purified forms never yielded active PaClpP1$_7$P2$_7$. Although this feature could be explained by the fact that PaClpP1$_{14}$, once assembled, is very stable/long-lived, and therefore there is never sufficient PaClpP1 to assemble with PaClpP2 *in vitro*. However, assembly of PaClpP1$_7$P2$_7$ may be more complicated *in vivo*. Is there a small molecule or factor that stabilizes PaClpP1$_7$P2$_7$ formation in *P. aeruginosa*? Intriguingly, located in the genome of *P. aeruginosa* immediately adjacent of *clpP2* (PA3326), operating under the same promoter, is a nonribosomal peptide synthetase (NRPS) gene cluster of 10 genes (PA3327-PA3336). The product of this NRPS biosynthetic operon is a

modified peptide called azetidomonamide B, or azabicyclene (Fig. 3-2; Hong et al., 2019; Patteson et al., 2019). Like *clpP2*, the NRPS genes are upregulated during quorum sensing (Schuster et al., 2003; Arevalo-Ferro et al., 2003). Genes homologous to *clpP2* and the NRPS gene cluster co-localize in a number of bacterial genomes, suggestive of a potential functional relationship, for example perhaps to preferentially promote assembly of the PaClpP1$_7$P2$_7$ heterocomplex. However, it is currently unknown how, or even if, the NRPS product functionally interacts with the ClpP machinery. Experiments to test this relationship may uncover insight into potential assembly factors and/or regulators of ClpP in *P. aeruginosa*.

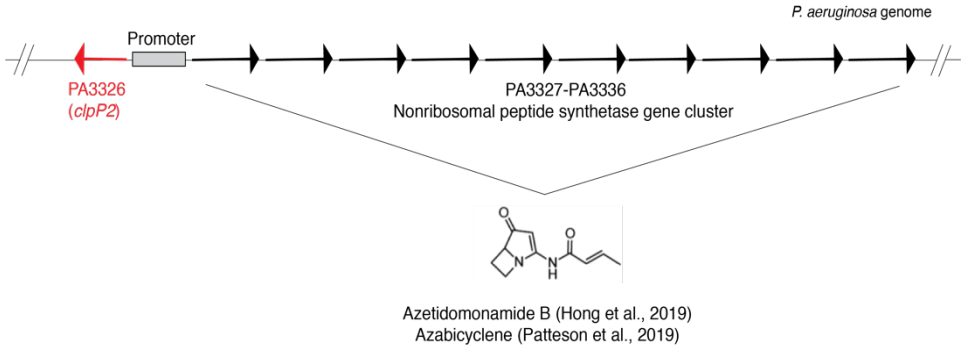

Figure 3-2: A Nonribosomal peptide synthetase gene cluster is located upstream of *clppP2*.
Genetic organization of the *clpP2* locus in *P. aeruginosa* PAO1. The red arrow represents the *clpP2* gene, and the black arrows represent ten nonribosomal peptide synthetase (NRPS) genes in a cluster located directly upstream of *clpP2*. Shown below the cluster is the proposed modified peptide that is produced by enzymes of the NRPS gene cluster.

To determine whether azetidomonamide B/azabicyclene physically binds to PaClpP2, one could purify PaClpP2-Flag from PAO1 stationary phase cultures and perform liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS), specifically looking for the azabicyclene, either associated or covalently bound to one of the isoforms of ClpP. LC-MS/MS was performed on both the PaClpP1 and PaClpP2-Flag recovered from *P. aeruginosa* cultures at

the beginning of Chapter 2, but coverage was only moderate, sufficient to identify both proteins but not to identify modifications, specifically as for neither isoform was a peptide containing the active-site Ser recovered (Fig. 3-3). Therefore, optimization of LC-HRMS or LC-MS/MS approach would be crucial for identifying covalent adducts on target residues in PaClpP1 and PaClpP2.

>PaClpP1 (1-198)
MSRNSFIPHVPDIQAAGGLVPMVVEQSARGERAYDIYSRLLKERIIFLVGQVEDYMANLV
VAQLLFLEAENPEKDIHLY**INSPGGSVTAGMSIY**DTMQFIKPNVSTTCIGQACSMGALL**L**
AGGAAGKRYCLPHSRMMIHQPLGGFQGQASDIEIHAK**EILFIKERLNQILAHHTGQPLDV**
IARDTDRDR**FMSGDEAVKYGLIDKVMTQR**DLAV

>PaClpP2 (1-201)
MKTDDKDR**EGGDSHGAIGAKLMEYALKVRKVFVTGGVDEK**MAKDVVQQLHILASISDDPI
YMFVNSPGGHVESGDMIFDAIRFITPKVIMIGSGSVASAGALIY**AAADKENRYSLPNTR**F
LL**HQPSGGIQGPASNIEIY**RREIVRMKER**LDRIFAEATGQTPEKISADTERDFW**LNAEEA
VQYGLVNKIIVSEREITLPGQ

Figure 3-3: Sequence coverage of PaClpP1 and PaClpP2 from LC-MS/MS done on PaClpP2-Flag pull-downs in *P. aeruginosa*.
Co-IP was done on PaClpP2 from PAO1 as in Fig. 2-1. Bands corresponding to PaClpP1 and PaClpP2 were excised and each was subjected to digestion by trypsin and chymotrypsin in separate reactions. Sequences in bold were identified by LC-MS/MS. Residues in red are the catalytic triad.

What substrates are turned over by PaClpP1$_{14}$ and PaClpP1$_7$P2$_7$ *in vivo*? To identify differential substrates between the complexes, I propose that future investigators should perform a substrate trapping approach in *P. aeruginosa* stationary-phase cultures. This technique relies on the ability of inactivated ClpP to "trap" proteins that have translocated into its chamber by AAA+ unfoldases (Kim et al., 2000; Flynn et al., 2003). ClpP variants can be made by mutating the ClpP active site (Ser->Ala substitution) and appending a tandem-affinity tag onto the C-terminus, referred to as

ClpP^{trap} variants. This method should allow for the highly-enriched purification of and subsequent detection of ClpP^{trap} substrates.

To identify cellular proteins captured by PaClpP1$_{14}$^{trap}, complexes should be isolated from Δ*clpP2* strains and subjected to tandem mass-spectrometry (Fig. 3-4A). For the identification of proteins captured by PaClpP1$_7$P2$_7$^{trap}, heterocomplexes should be purified from doubly-inactivated ClpP strains, with the tandem-affinity tag appended on *clpP2* (*clpP1*^{SA} *clpP2*^{SA}*-Myc-TEV-His*) (Fig 3-4B). Certainly, in both PaClpP1$_{14}$^{trap} and PaClpP1$_7$P2$_7$^{trap} systems, a large portion of the proteins trapped inside PaClpP will be substrates of PaClpX or PaClpA (see Table 3-1 for likely targets of PaClpXP based on my probing of the PAO1 annotated genome for proteins terminating in the sequence AA-coo$^-$). To determine ClpA/X-independent substrates of PaClpP1$_{14}$ and PaClpP1$_7$P2$_7$, trapping experiments followed by mass spectrometry should be performed from strains deleted for PaClpA and PaClpX (Δ*clpA* Δ*clpX*) (Fig. 3-5). Substrates found in this trapping setup would represent proteins internalized into ClpP either by an alternative AAA+ unfoldase, or peptides that passively diffuse into the ClpP chamber without the aid of unfoldases.

Work presented in this thesis establishes PaClpP1$_7$P2$_7$ as a new ClpP peptidase in *P. aeruginosa* with positive functions in biofilm formation. Because biofilm formation is central to the bacterium's mode of virulence, PaClpP1$_7$P2$_7$ should be considered as a target of therapeutic importance. Specifically, the design and manufacture of drugs to target and dampen the activity of PaClpP1$_7$P2$_7$ may be a useful strategy to prevent the colonization and development of *P. aeruginosa* infections. It will be worthwhile for future studies to investigate the potential druggability of PaClpP1$_7$P2$_7$ in preventing the spread of *P. aeruginosa* in clinical settings.

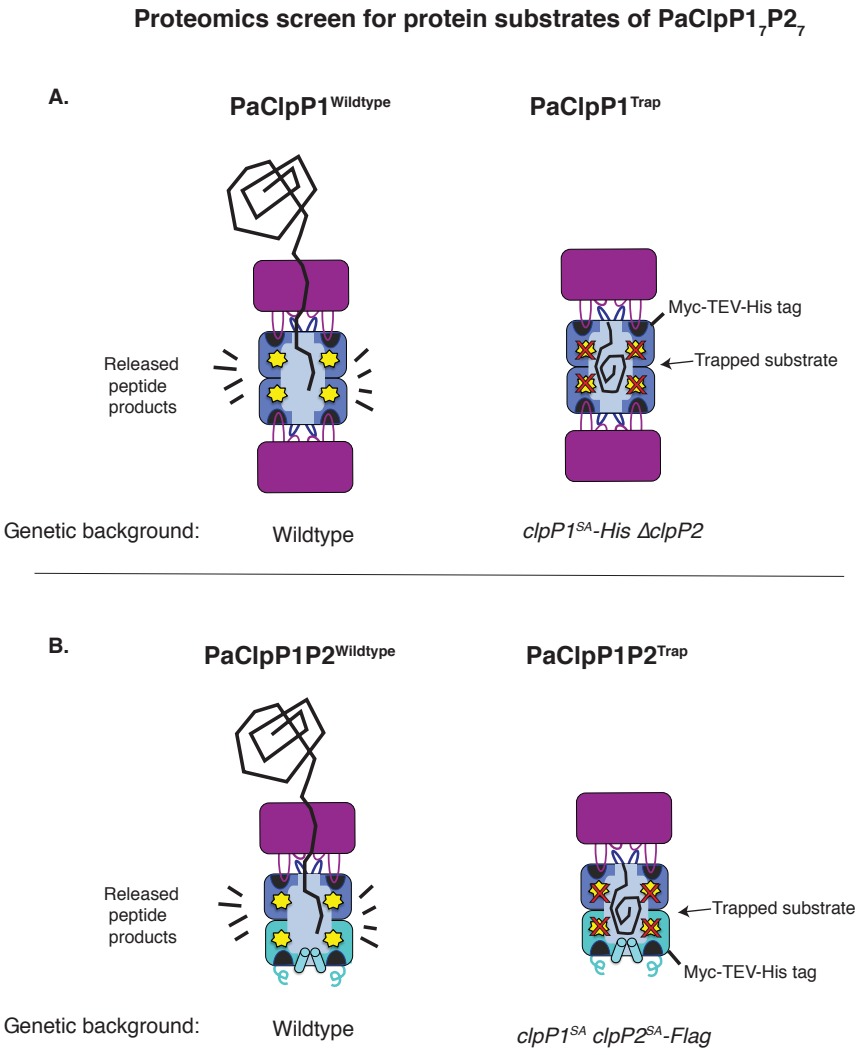

Figure 3-4: ClpP-trapping approach to identify protein substrates of PaClpP complexes.
A) Targets of PaClpX•P1$_{14}$ and PaClpA•P1$_{14}$ can be identified by this method. Briefly, *P. aeruginosa* strains expressing an affinity-tagged, catalytically-null PaClpP1 (*clpP1^{SA} Myc-TEV-His*) would be cultured to stationary-phase, and pull-downs for PaClpP1^{SA}-Myc-TEV-His would be performed. Proteins co-immunoprecipitating with PaClpP1^{SA}-Myc-TEV-His would be identified using mass spectrometry. Importantly, to isolate pure PaClpP1$^{SA}{}_{14}$-His complexes, PaClpP1-trapping experiments should be done in a Δ*clpP2* genetic background. B) Targets of PaClpX•P1$_7$P2$_7$ and PaClpA•P1$_7$P2$_7$ can be identified by this method. Workflow is same as in (A), except both PaClpP1 and PaClpP2 should harbor inactivating mutations, with a tandem affinity tag appended onto PaClpP2 for pull-downs. Targets of the co-complex would be identified using mass-spectrometry.

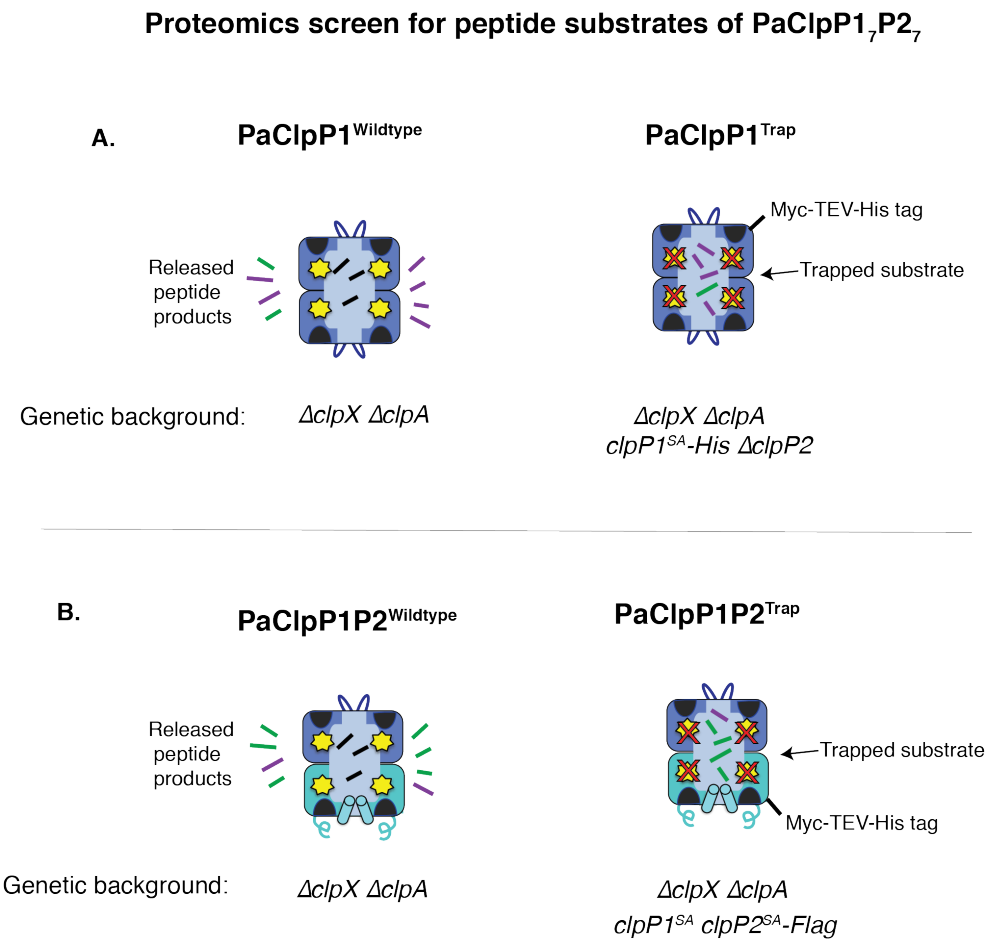

Figure 3-5: ClpP-trapping approach to identify peptide substrates of PaClpP complexes.
A) Targets of PaClpP1$_{14}$ can be identified by this method. Δ*clpX* Δ*clpA P. aeruginosa* strains expressing an affinity-tagged, catalytically-null PaClpP1 (*clpP1^{SA}-Myc-TEV-His*) would be cultured to stationary-phase, and pull-downs for PaClpP1^{SA}-Myc-TEV-His would be performed. Peptides co-immunoprecipitating with PaClpP1^{SA}-His would be identified using mass spectrometry. Importantly, to isolate pure PaClpP1$^{SA}{}_{14}$-Myc-TEV-His complexes, PaClpP1-trapping experiments should be done in a Δ*clpP2* genetic background. B) Peptide targets of PaClpP1$_7$P2$_7$ and PaClpP1$_7$P2$_7$ can be identified by this method. Workflow is same as in (A), except both PaClpP1 and PaClpP2 should harbor inactivating mutations, with a tandem affinity tag appended onto PaClpP2 for pull-downs. Peptide targets of the co-complex would be identified using mass-spectrometry.

Table 3-1: "-AA" ending proteins in *P. aeruginosa* PAO1

Gene	Description	C-terminal residues
PA0064	Uncharacterized protein	-QLAA
PA0067	Peptidase_M3 domain-containing protein	-GEAA
PA0098	Uncharacterized protein	-VRAA
PA0112	Uncharacterized protein	-LKAA
PA0167	HTH tetR-type domain-containing protein	-PPAA
PA0177	CheW-like domain-containing protein	-EAAA
PA0191	HTH lysR-type domain-containing protein	-HAAA
PA0245	3-dehydroquinate dehydratase 2	-EAAA
PA0276	Uncharacterized protein	-KRAA
PA0341	Prolipoprotein diacylglyceryl transferase	-PKAA
PA0440	Probable oxidoreductase	-VEAA
PA0444	M20_dimer domain-containing protein	-RLAA
PA0480	AB hydrolase-1 domain-containing protein	-LGAA
PA0524	Nitric oxide reductase subunit B	-KAAA
PA0532	Uncharacterized protein	-KHAA
PA0805	Uncharacterized protein	-TLAA
PA0826.2	Proteolysis tag peptide encoded by tmRNA	-ALAA
PA0830	Uncharacterized protein	-PKAA
PA0872	Phenylalanine-4-hydroxylase	-KQAA
PA0887	Acetyl-coenzyme A synthetase 1	-MQAA
PA1056	Proton_antipo_M domain-containing protein	-GGAA
PA1221	Carrier domain-containing protein	-RSAA
PA1245	Uncharacterized protein	-ALAA
PA1286	MFS domain-containing protein	-GGAA
PA1291	AB hydrolase-1 domain-containing protein	-PDAA
PA1297	Probable metal transporter	-PLAA
PA1309	HTH lysR-type domain-containing protein	-LAAA
PA1418	Probable sodium:solute symport protein	-ASAA
PA1491	Probable transporter	-LHAA
PA1520	HTH gntR-type domain-containing protein	-DSAA
PA1578.1	UPF0153 protein PA1578.1	-ITAA
PA1593	4HBT domain-containing protein	-GGAA
PA1654	Aminotran_1_2 domain-containing protein	-AEAA
PA1674	GTP cyclohydrolase FolE2	-RGAA
PA1700	Uncharacterized protein	-VLAA
PA1731	Alpha-E domain-containing protein	-LEAA
PA1749	N-acetyltransferase domain-containing protein	-DPAA
PA1800	Trigger factor	-PQAA
PA1822	Uncharacterized protein	-PVAA
PA1836	HTH tetR-type domain-containing protein	-VRAA
PA1958	Probable transporter	-GVAA
PA1962	FMN-dependent NADH-azoreductase 2	-FAAA
PA1971	Branched-chain amino acid transport system 3 carrier protein	-QVAA
PA2007	Maleylacetoacetate isomerase	-TPAA
PA2022	UDP-glucose 6-dehydrogenase	-ASAA
PA2042	Serine/threonine transporter SstT	-ASAA
PA2125	Aldedh domain-containing protein	-TGAA
PA2157	Uncharacterized protein	-ESAA
PA2171	Hemerythrin domain-containing protein	-KRAA

PA2207	Uncharacterized protein	-RHAA
PA2212	5-oxoprolinase subunit A 2	-LEAA
PA2288	Uncharacterized protein	-QRAA
PA2298	Probable oxidoreductase	-NDAA
PA2304	TauD domain-containing protein	-GGAA
PA2321	Gluconokinase	-HAAA
PA2341	ABC transporter domain-containing protein	-QQAA
PA2366	Uricase PuuD	-PVAA
PA2381	Uncharacterized protein	-SSAA
PA2467	Anti-sigma factor FoxR	-VVAA
PA2484	HTH tetR-type domain-containing protein	-VEAA
PA2516	Toluate 1,2-dioxygenase electron transfer component	-ASAA
PA2524	Sensor protein	-TGAA
PA2687	Sensor protein PfeS	-PAAA
PA2897	HTH gntR-type domain-containing protein	-LIAA
PA2910	Putative manganese efflux pump MntP	-LSAA
PA2930	HTH lysR-type domain-containing protein	-SAAA
PA2987	Lipoprotein-releasing system ATP-binding protein LolD	-LIAA
PA3122	HTH lysR-type domain-containing protein	-TPAA
PA3129	tRNA-dihydrouridine (16) synthase	-SQAA
PA3147	Glyco_trans_4-like_N domain-containing protein	-GRAA
PA3219	Metallophos_2 domain-containing protein	-EPAA
PA3258	Uncharacterized protein	-IEAA
PA3323	Peptidase_M24 domain-containing protein	-REAA
PA3334	Carrier domain-containing protein	-HRAA
PA3335	Uncharacterized protein	-SSAA
PA3407	Heme acquisition protein	-ALAA
PA3490	Ion-translocating oxidoreductase complex subunit B	-ERAA
PA3530	Fer2_BFD domain-containing protein	-FVAA
PA3534	4Fe-4S Mo/W bis-MGD-type domain-containing protein	-VEAA
PA3573	Bcr/CflA family efflux transporter	-RQAA
PA3584	Glycerol-3-phosphate dehydrogenase	-VHAA
PA3602	Glu_synthase domain-containing protein	-DAAA
PA3607	Spermidine/putrescine import ATP-binding protein PotA	-LPAA
PA3657	Methionine aminopeptidase	-TSAA
PA3741	HotDog ACOT-type domain-containing protein	-SQAA
PA3775	Probable membrane transporter protein	-YLAA
PA3792	2-isopropylmalate synthase	-AKAA
PA3816	O-acetylserine synthase	-NPAA
PA3954	Bac_luciferase domain-containing protein	-REAA
PA3962	Uncharacterized protein	-KPAA
PA3981	PhoH domain-containing protein	-HDAA
PA3989	DNA polymerase III, delta subunit	-LPAA
PA3990	Uncharacterized protein	-VMAA
PA4038	Uncharacterized protein	-KKAA
PA4092	4-hydroxyphenylacetate 3-monooxygenase reductase component	-RRAA
PA4115	Uncharacterized protein	-TEAA
PA4151	Transket_pyr domain-containing protein	-RDAA
PA4171	DJ-1_PfpI domain-containing protein	-QAAA
PA4223	Probable ATP-binding component of ABC transporter	-EVAA
PA4230	Isochorismate pyruvate lyase	-RGAA
PA4272	50S ribosomal protein L10	-AAAA
PA4312	Uncharacterized protein	-DHAA

PA4344	M20_dimer domain-containing protein	-NGAA
PA4351	PlsC domain-containing protein	-RRAA
PA4366	Superoxide dismutase [Fe]	-NFAA
PA4504	ABC transmembrane type-1 domain-containing protein	-KNAA
PA4563	30S ribosomal protein	-TAAA
PA4567	50S ribosomal protein L27	-IVAA
PA4783	Uncharacterized protein	-RAAA
PA4784	HTH asnC-type domain-containing protein	-GCAA
PA4852	tRNA-dihydrouridine synthase	-GVAA
PA4893	Urease accessory protein UreG	-LTAA
PA4903	MFS domain-containing protein	-GDAA
PA4908	Uncharacterized protein	-DNAA
PA5147	Adenine DNA glycosylase	-GTAA
PA5199	AmgS	-RGAA
PA5222	Uncharacterized protein	-RNAA
PA5224	AMP_N domain-containing protein	-SEAA
PA5259	Uroporphyrinogen-III synthase	-TSAA
PA5347	Uncharacterized protein	-KKAA
PA5400	ETF domain-containing protein	-RDAA
PA5419	Sarcosine oxidase gamma subunit	-QVAA
PA5421	Glutathione-independent formaldehyde dehydrogenase	-FRAA
PA5486	UPF0056 inner membrane protein	-HIAA
PA5535	CobW C-terminal domain-containing protein	-GHAA

References

Akopian, T., Kandror, O., Raju, R. M., Unnikrishnan, M., Rubin, E. J., & Goldberg, A. L. (2012). The active ClpP protease from M. tuberculosis is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO Journal*, *31*(6), 1529–1541. https://doi.org/10.1038/emboj.2012.5

Arevalo-Ferro, C., Hentzer, M., Reil, G., Görg, A., Kjelleberg, S., Givskov, M., … Eberl, L. (2003). Identification of quorum-sensing regulated proteins in the opportunistic pathogen Pseudomonas aeruginosa by proteomics. *Environmental Microbiology*, *5*(12), 1350–1369. https://doi.org/10.1046/j.1462-2920.2003.00532.x

Arnold, I., Wagner-Ecker, M., Ansorge, W., & Langer, T. (2006). Evidence for a novel mitochondria-to-nucleus signalling pathway in respiring cells lacking i-AAA protease and the ABC-transporter Mdl1. *Gene*, *367*(1–2), 74–88. https://doi.org/10.1016/j.gene.2005.09.044

Augustin, S., Nolden, M., Müller, S., Hardt, O., Arnold, L., & Langer, T. (2005). Characterization of peptides released from mitochondria: Evidence for constant proteolysis and peptide efflux. *Journal of Biological Chemistry*, *280*(4), 2691–2699. https://doi.org/10.1074/jbc.M410609200

Baker, T. A., & Sauer, R. T. (2012). ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1823*(1), 15–28. https://doi.org/10.1016/j.bbamcr.2011.06.007

Chua, S. L., Hultqvist, L. D., Yuan, M., Rybtke, M., Nielsen, T. E., Givskov, M., … Yang, L. (2015). In vitro and in vivo generation and characterization of Pseudomonas aeruginosa biofilm-dispersed cells via c-di-GMP manipulation. *Nat Protoc*, *10*(8), 1165–1180. https://doi.org/10.1038/nprot.2015.067

E, V. (2014). *Microarray Analysis of Pseudomonas aeruginosa Quorum-Sensing Regulons : Effects of Growth Phase and Environment Microarray Analysis of Pseudomonas aeruginosa Quorum-Sensing Regulons : Effects of Growth Phase and Environment †*. *185*(7), 2080–2095. https://doi.org/10.1128/JB.185.7.2080

Ferrington, D. A., & Gregerson, D. S. (2012). Immunoproteasomes: Structure, function, and antigen presentation. In *Progress in Molecular Biology and Translational Science* (Vol. 109). https://doi.org/10.1016/B978-0-12-397863-9.00003-1

Flynn, J. M., Neher, S. B., Kim, Y. I., Sauer, R. T., & Baker, T. A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Molecular Cell*, *11*(3), 671–683. https://doi.org/10.1016/S1097-2765(03)00060-1

Gersch, M., Stahl, M., Poreba, M., Dahmen, M., Dziedzic, A., Drag, M., & Sieber, S. A. (2016). Barrel-shaped ClpP Proteases Display Attenuated Cleavage Specificities. *ACS Chemical Biology*, *11*(2), 389–399. https://doi.org/10.1021/acschembio.5b00757

Gilbert, K. B., Kim, T. H., Gupta, R., Greenberg, E. P., & Schuster, M. (2009). Global position analysis of the Pseudomonas aeruginosa quorum-sensing transcription factor LasR. *Molecular Microbiology*, *73*(6), 1072–1085. https://doi.org/10.1111/j.1365-2958.2009.06832.x

Hall, B. M., Breidenstein, E. B. M., de la Fuente-Núñez, C., Reffuveille, F., Mawla, G. D., Hancock, R. E. W., & Baker, T. A. (2017). Two isoforms of Clp peptidase in Pseudomonas aeruginosa control distinct aspects of cellular physiology. *Journal of Bacteriology*, *199*(3),

1–15. https://doi.org/10.1128/JB.00568-16
Harris, J. L., Alper, P. B., Li, J., Rechsteiner, M., & Backes, B. J. (2001). Substrate specificity of the human proteasome. *Chemistry and Biology*, *8*(12), 1131–1141. https://doi.org/10.1016/S1074-5521(01)00080-1
Haynes, C. M., Petrova, K., Benedetti, C., Yang, Y., & Ron, D. (2007). ClpP Mediates Activation of a Mitochondrial Unfolded Protein Response in C. elegans. *Developmental Cell*, *13*(4), 467–480. https://doi.org/10.1016/j.devcel.2007.07.016
Haynes, C. M., Yang, Y., Blais, S. P., Neubert, T. A., & Ron, D. (2010). The Matrix Peptide Exporter HAF-1 Signals a Mitochondrial UPR by Activating the Transcription Factor ZC376.7 in C. elegans. *Molecular Cell*, *37*(4), 529–540. https://doi.org/10.1016/j.molcel.2010.01.015
Herget, M., & Tampé, R. (2007). Intracellular peptide transporters in human - Compartmentalization of the "peptidome." *Pflugers Archiv European Journal of Physiology*, *453*(5), 591–600. https://doi.org/10.1007/s00424-006-0083-4
Hong, Z., Bolard, A., Giraud, C., Prévost, S., Genta-Jouve, G., Deregnaucourt, C., … Li, Y. (2019). Azetidine-Containing Alkaloids Produced by a Quorum-Sensing Regulated Nonribosomal Peptide Synthetase Pathway in Pseudomonas aeruginosa. *Angewandte Chemie - International Edition*, *58*(10), 3178–3182. https://doi.org/10.1002/anie.201809981
Kim, Y. I., Burton, R. E., Burton, B. M., Sauer, R. T., & Baker, T. A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. *Molecular Cell*, *5*(4), 639–648. https://doi.org/10.1016/S1097-2765(00)80243-9
Kisselev, A. F., Garcia-Calvo, M., Overkleeft, H. S., Peterson, E., Pennington, M. W., Ploegh, H. L., … Goldberg, A. L. (2003). The caspase-like sites of proteasomes, their substrate specificity, new inhibitors and substrates, and allosteric interactions with the trypsin-like sites. *Journal of Biological Chemistry*, *278*(38), 35869–35877. https://doi.org/10.1074/jbc.M303725200
Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., … Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, *47*(W1), W636–W641. https://doi.org/10.1093/nar/gkz268
Melber, A., & Haynes, C. M. (2018). UPR mt regulation and output: A stress response mediated by mitochondrial-nuclear communication. *Cell Research*, *28*(3), 281–295. https://doi.org/10.1038/cr.2018.16
Patteson, J. B., Lescallette, A. R., & Li, B. (2019). Discovery and Biosynthesis of Azabicyclene, a Conserved Nonribosomal Peptide in Pseudomonas aeruginosa. *Organic Letters*, *21*(13), 4955–4959. https://doi.org/10.1021/acs.orglett.9b01383
Rutherford, S. T., & Bassler, B. L. (2012). Bacterial quorum sensing: Its role in virulence and possibilities for its control. *Cold Spring Harbor Perspectives in Medicine*, *2*(11), 1–25. https://doi.org/10.1101/cshperspect.a012427
Schmitza, K. R., Carneyb, D. W., Sellob, J. K., & Sauera, R. T. (2014). Crystal structure of mycobacterium tuberculosis ClpP1p2 suggests a model for peptidase activation by aaa+ partner binding and substrate delivery. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(43), E4587–E4595. https://doi.org/10.1073/pnas.1417120111
Schuster, M., Lostroh, C. P., Ogi, T., & Greenberg, E. P. (2003). (4916) Identification, Timing, and Signal Speci city of. *Society*, *185*(7), 2066–2079. https://doi.org/10.1128/JB.185.7.2066
Sheps, J. A., Ralph, S., Zhao, Z., Baillie, D. L., & Ling, V. (2004). The ABC transporter gene

family of Caenorhabditis elegans has implications for the evolutionary dynamics of multidrug resistance in eukaryotes. *Genome Biology*, *5*(3), 1–17. https://doi.org/10.1186/gb-2004-5-3-r15

St-Pierre, C., Morgand, E., Benhammadi, M., Rouette, A., Hardy, M. P., Gaboury, L., & Perreault, C. (2017). Immunoproteasomes Control the Homeostasis of Medullary Thymic Epithelial Cells by Alleviating Proteotoxic Stress. *Cell Reports*, *21*(9), 2558–2570. https://doi.org/10.1016/j.celrep.2017.10.121

Thompson, M. W., & Maurizi, M. R. (1994). Activity and specificity of Escherichia coli ClpAP protease in cleaving model peptide substrates. *Journal of Biological Chemistry*, *269*(27), 18201–18208.

Vahidi, S., Ripstein, Z. A., Juravsky, J. B., Rennella, E., Goldberg, A. L., Mittermaier, A. K., … Kay, L. E. (2020). An allosteric switch regulates Mycobacterium tuberculosis ClpP1P2 protease function as established by cryo-EM and methyl-TROSY NMR. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(11), 5895–5906. https://doi.org/10.1073/pnas.1921630117

Yewdell, J. W. (2005). Immunoproteasomes: Regulating the regulator. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(26), 9089–9090. https://doi.org/10.1073/pnas.0504018102

Young, L., Leonhard, K., Tatsuta, T., Trowsdale, J., & Langer, T. (2001). Role of the ABC transporter Mdl1 in peptide export from mitochondria. *Science*, *291*(5511), 2135–2138. https://doi.org/10.1126/science.1056957

Yu, A. Y. H., & Houry, W. A. (2007). ClpP: A distinctive family of cylindrical energy-dependent serine proteases. *FEBS Letters*, *581*(19), 3749–3757. https://doi.org/10.1016/j.febslet.2007.04.076