Structural Principles of Substrate Recognition and Unfolding by the ClpAP and ClpXP AAA+ Proteases

by

Sora Kim

B.A. Chemistry and Biology, concentration in Biochemistry & Molecular Biology

Williams College, 2013

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

DOCTOR OF PHILOSOPHY

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MAY 2022

Signature of Author: ……....

Department of Biology
April 1, 2022

Certified by: ……

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

Accepted by: ……..

Amy E. Keating
Professor of Biology and Biological Engineering, Associate Department Head
Co-Director, Biology Graduate Committee

Structural Principles of Substrate Recognition and Unfolding by the ClpAP and ClpXP AAA+ Proteases

by

Sora Kim

Submitted to the Department of Biology on April 1, 2022 in partial fulfillment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

Protein degradation is a key regulatory mechanism that controls protein homeostasis in all cells. Found in all domains of life, proteases of the AAA+ (ATPases associated with diverse cellular activities) superfamily perform targeted protein degradation of specific substrates. All AAA+ proteases consist of a hexameric AAA+ unfoldase and a compartmentalized peptidase. AAA+ proteases harness the energy of ATP hydrolysis to mechanically unfold and translocate substrates through the axial channel into the peptidase chamber for degradation. Degradation by AAA+ proteases is carefully regulated by several mechanisms, including selective recognition of specific peptide sequences in the substrate (called degrons), substrate-tethering sequences (called enhancement- or e-tags), and adaptor proteins.

My thesis examines the structural basis of substrate recognition and degradation by the bacterial AAA+ proteases ClpAP and ClpXP, which are composed of either the double-ringed ClpA or the single-ring ClpX unfoldase in complex with the ClpP peptidase. Using covalently crosslinked ClpA–ClpP complexes, I interrogate the symmetry mismatch between the ClpA hexamer and the ClpP heptamer interface to establish that rotation of ClpA relative to ClpP is not required for proteolytic function, contrary to the prediction of structure-based models. Next, I present cryo-EM structures of ClpAP in complex with its adaptor ClpS and an N-degron substrate, revealing the degron-like binding of the ClpS NTE in the ClpA channel and an altered, ‘tucked’ pore-1-loop conformation. I also investigate the function of ClpA D1 pore-2 loops in adaptor-independent and ClpS-assisted degradation. Next, using engineered fusion proteins, I show that a Pro-Pro dipeptide (found in the ClpS junction) near a folded domain is sufficient for protection from ClpAP. Finally, I investigate the redox-sensitive mechanism of ClpXP degradation of the bacterial global transcription regulator FNR and demonstrate that the N-terminal e-tag is a ClpX-tethering motif and the C-terminal recognition signal is a pore-binding tag. Crystal structures of *apo* (oxidized) and [4Fe-4S]-bound *holo*FNR reveal how oxygen-induced conformational changes alter exposure of these two signals to selectively target the *apo* form for degradation. In summary, these studies identify distinct elements within enzymes, adaptors, and substrates that contribute specific functions during the multistep process of degradation by AAA+ proteases.

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

BIOGRAPHICAL NOTE

Sora Kim identifies as gender non-binary and uses they/them pronouns.

PREFACE

I use the first-person plural pronoun, “we”, in my thesis to refer to *(i)* collaborative work performed as noted in the author acknowledgments of each section, or *(ii)* my own ideas, which have been informed by discussions with my thesis advisors and members of the Baker and Sauer labs, as well as research by other investigators.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

OVERVIEW

In all cells, the continual, dynamic processes of protein synthesis and degradation determine the abundance of individual proteins. Cells control the rates of these opposing processes and monitor protein folding to maintain functional proteomes. Once synthesized, proteins exist in a conformational equilibrium, folding from nascent, unfolded states to energetically-favored native conformations, as well as misfolding into energetically competing, aggregate structures. Enzymes and factors in the protein quality control network refold, sequester, and eliminate proteins that misfold, aggregate, or become damaged or unneeded (Chen et al., 2011). For example, AAA+ (ATPases associated with diverse cellular activities) remodeling enzymes/unfoldases can function as chaperones to promote proper folding or as proteases (in a protein complex with a peptidase) to unfold and degrade specific proteins (Mogk et al., 2011). As proteolysis is irreversible and destructive, protein degradation by AAA+ proteases is carefully controlled to avoid indiscriminate activity (Mahmoud and Chien, 2018).

This thesis focuses on the molecular principles regulating protein degradation by AAA+ proteases in bacteria. To contextualize my work on these enzymes, this section first provides background on the biological relevance of regulated proteolysis and the basic architecture of AAA+ proteases in prokaryotes and eukaryotes. The next portion of the introduction describes how substrate proteins are targeted for recognition by bacterial AAA+ proteases, followed by a section describing how additional proteins, called adaptors, modulate substrate recognition.

IMPORTANCE OF REGULATED PROTEIN DEGRADATION

Targeted protein degradation performs protein quality control and regulates numerous cellular processes by altering the proteome in response to stimuli (Varshavsky, 2017; Mahmoud and Chien, 2018). Following exposure to stress, such as oxidative damage, heat shock, and nutrient starvation, or as a result of spontaneous fluctuations in protein conformation, proteases degrade proteins that become damaged, unfolded, misfolded, or aggregation-prone (Hartl and Hayer-Hartl, 2009; Chen et al., 2011). By removing aberrant proteins, proteases also help alleviate chaperone-overload, thereby preventing excess non-native proteins from accumulating and allowing chaperones to effectively buffer against protein misfolding and aggregation.

Although the detailed mechanisms of protein toxicity are poorly understood, misfolded proteins and aggregates can lead to dysfunctional protein homeostasis and even cell death due to *(i)* aberrant interactions with other proteins, membranes, and organelles, *(ii)* the acquisition of detrimental function(s) from altered protein conformations, and *(iii)* by overwhelming the buffering capacity of the protein quality control network (Rao and Bredesen, 2004; Stefani, 2004; Valastyan and Lindquist, 2014; Gonzalez-Garcia et al., 2021). In the absence of properly functioning proteolytic quality control pathways, cells accumulate these often-deleterious species, which contributes to pathological conditions, including cystic fibrosis, lysosomal storage disorders, cancer, and neurodegenerative diseases (Chen et al., 2011). For example, in breast, colon, liver, lung, and ovarian cancer cell lines and tumor tissues, mutations in the tumor-suppressor protein p53 induce aggregation into amyloid-like fibrils, which have been shown to promote oncogenesis and chemoresistance (Yang-Hartwich et al., 2015; Ghosh et al., 2017; de Oliveira et al., 2020; Navalkar et al., 2021). Conversely, cells often activate protein degradation pathways, such as the unfolded protein response and autophagy, to protect against apoptosis (Wang and Kaufman, 2014; Mulcahy Levy and Thorburn, 2020). Rapidly dividing tumor cells in many cancer types also

frequently upregulate proteasomal degradation to meet their increased metabolic demands (Rousseau and Bertolotti, 2018; Morozov and Karpov, 2019).

In addition to their central role in protein homeostasis, proteases help organisms respond to environmental changes often by degrading regulatory proteins. For example, proteases target cell cycle proteins for destruction at specific times, driving unidirectional cell division progression from one phase to another (King et al., 1996). Furthermore, proteolysis is a fast process that directly changes protein levels, allowing for rapid regulation of cellular processes compared to much slower mechanisms that rely on changes in transcription and translation (Mahmoud and Chien, 2018). Thus, protein degradation is a fundamental regulatory mechanism involved in nearly all cellular processes ranging from the molecular to the organismal level, such as *(i)* gene expression, signal transduction, and stress responses, *(ii)* apoptosis, *(iii)* DNA repair, *(iv)* metabolism, *(v)* vesicular trafficking, *(vi)* immunity, and *(vii)* cell division, cell differentiation, and tissue morphogenesis (Brinkmann et al., 2015; Varshavsky, 2017; Mahmoud and Chien, 2018).

AAA+ PROTEASES

AAA+ (ATPases associated with diverse cellular activities) proteases use the energy of ATP hydrolysis to degrade damaged, unneeded, and/or regulatory proteins. AAA+ proteases consist of a AAA+ unfoldase/remodeling enzyme and a self-compartmentalized peptidase, which are either genetically-tethered by being encoded in the same gene or are assembled into stable complexes from separately encoded subunits. AAA+ unfoldases/remodeling enzymes are a subgroup of the AAA+ superfamily (described below, p. 20) that contain at least one AAA+ module (an evolutionarily conserved protein fold containing the ATPase) and a family-specific auxiliary domain. All AAA+ unfoldases/remodeling enzymes assemble into ring-shaped oligomers (usually hexamers), forming an axial channel lined with pore loops that bind a polypeptide substrate. Repeated cycles of ATP binding, hydrolysis, and product release drive conformational changes

in the AAA+ hexamer that transmit mechanical force through these pore loops to the bound substrate, resulting in *(i)* protein remodeling or *(ii)* protein unfolding and subsequent translocation of the unfolded polypeptide into the peptidase chamber (see **Figure 1.1**).

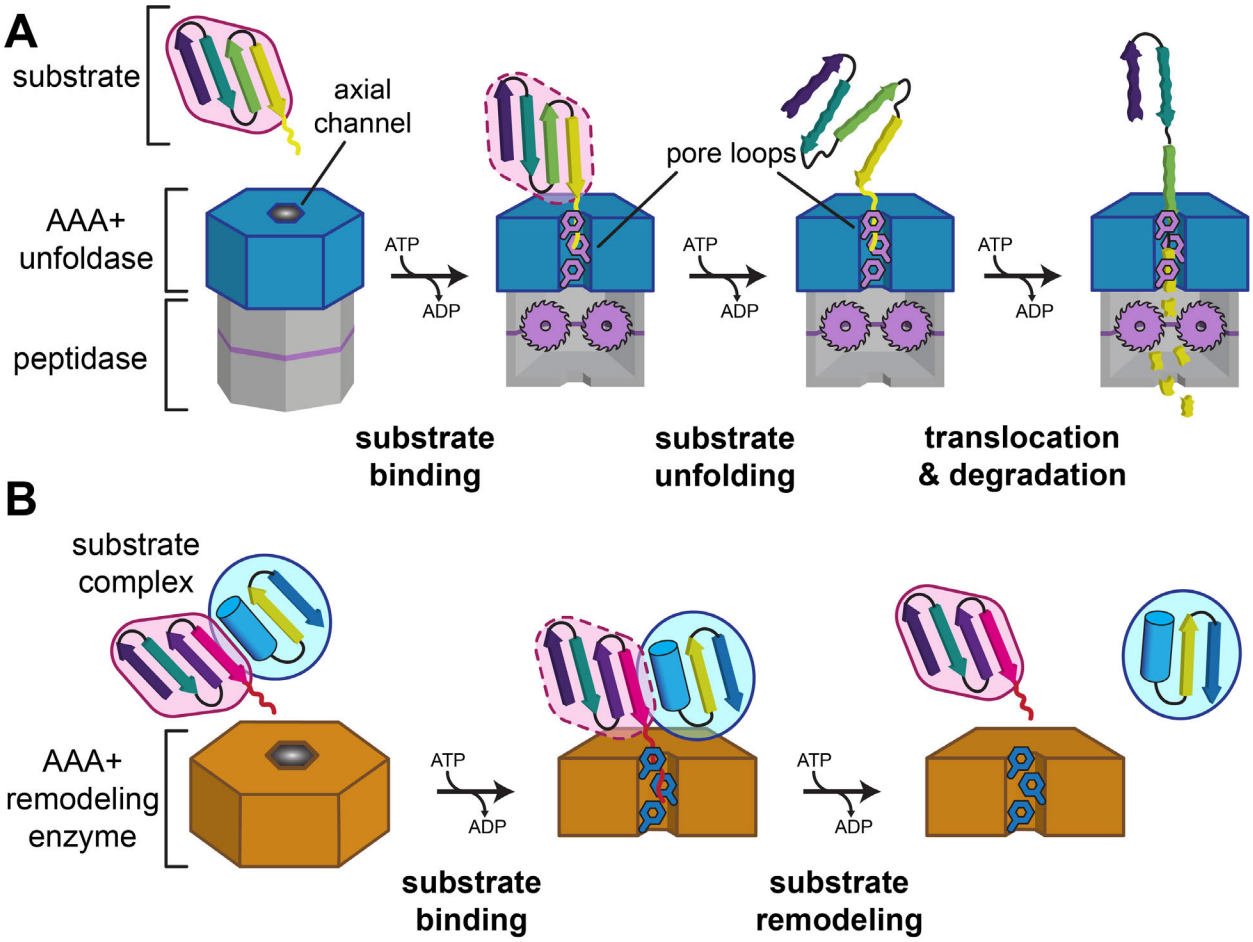


Figure 1.1 Mechanism of AAA+ proteases and remodeling enzymes.
(A) Protein substrate binding, unfolding, and translocation/degradation by AAA+ proteases. The AAA+ protease, which is composed of a hexameric AAA+ unfoldase and a self-compartmentalized peptidase, binds a protein substrate via pore loops located in the axial channel. Multiple cycles of ATP binding, hydrolysis, and product release power conformational changes in the AAA+ unfoldase that lead to substrate unfolding, followed by polypeptide translocation into the peptidase chamber where proteolytic active sites (serrated circles) cleave the polypeptide (yellow fragments).
(B) Disassembly of protein complexes by AAA+ remodeling enzymes. Through axial pore-loop interactions with a protein substrate, AAA+ remodeling enzymes bind and remodel a specific component within the substrate complex, which results in ATP hydrolysis-driven protein complex disassembly.

In the absence of associated peptidases, AAA+ unfoldases/remodeling enzymes can remodel substrate proteins to induce conformational changes in tertiary structure and/or to disassemble protein complexes. For example, the unfoldase ClpX alone disassembles the DNA-bound tetrameric form of the bacteriophage MuA transposase, whereas the protease ClpXP (ClpX in complex with the ClpP peptidase) degrades MuA monomers (Levchenko et al., 1995; Kruklitis et al., 1996; Laachouch et al., 1996; Abdelhakim et al., 2008; Ling et al., 2015). Other AAA+ remodeling enzymes lack partner peptidases and work exclusively as remodeling or disassembly machines, including the major cellular disaggregases ClpB/Hsp104, which collaborate with the refolding chaperones DnaJ/Hsp70 and DnaK/Hsp40; the microtubule-severing enzyme katanin; and the membrane fusion enzyme NSF, which disassembles SNARE (soluble NSF attachment protein receptor) complexes upon completion of vesicle docking to allow SNARE subunit recycling (Sauer and Baker, 2011; Olivares et al., 2015). In contrast to AAA+ proteases, many AAA+ remodeling enzymes do not fully unfold their substrate proteins and instead selectively destabilize individual components of protein complexes to promote a specific conformational change (illustrated in **Figure 1.1B**) (Burton and Baker, 2005).

Self-compartmentalized peptidases form barrel-shaped complexes with their active sites facing the interior of the peptidase chamber. The identity of these active-site residues varies among self-compartmentalized peptidases (*e.g.*, the proteasome and HslV contain N-terminal Thr residues to serve as the catalytic nucleophile, the zinc metalloprotease protease FtsH uses Asp-zincin/HEXXH, Lon uses a Ser-Lys dyad, and ClpP uses a Ser-His-Asp triad), and therefore contributes to different peptide cleavage specificities (Löwe et al., 1995; Wang et al., 1997; Bochtler et al., 2000; Sousa et al., 2002; Botos et al., 2004; Bieniossek et al., 2006; Suno et al., 2006). Most self-compartmentalized peptidases are synthesized as proenzymes in which an N-terminal propeptide is autocatalytically removed to generate the active, mature form, and in some contexts, *e.g.*, for the β subunits of the proteasome, to expose active-site residues hidden

in the unprocessed form (Thompson and Maurizi, 1994; Maupin-Furlow, Julie, 2000). For some peptidases, the propeptide promotes oligomerization by bridging the interface between neighboring subunits (Murata et al., 2009). Whereas for others, the oligomeric assembly helps position the proenzyme subunits for propeptide cleavage (Witt et al., 2006). Thus, formation of a barrel-shaped assembly sequesters peptidolytic active sites, limiting degradation to substrates selectively translocated by their AAA+ unfoldase partner.

In complex with a AAA+ unfoldase, self-compartmentalized peptidases degrade polypeptide substrates that are translocated from the axial channel of the AAA+ enzyme. Narrow central pores that gate access to the proteolytic active sites block folded proteins from entering (Baumeister et al., 1998; Lee et al., 2010b). Furthermore, additional mechanisms prevent proteolysis in the absence of a AAA+ unfoldase partner. Binding of a AAA+ unfoldase partner often controls the width of the peptidase entrance pore by directly modulating the residues that sterically block the opening (Groll et al., 2000; Whitby et al., 2000; Rabl et al., 2008; Lee et al., 2010b; Effantin et al., 2010; Fei et al., 2020a). Alternatively, proteolysis by these barrel-shaped enzymes can be controlled by allosteric mechanisms that induce an active conformation in the active-site residues upon binding of the AAA+ unfoldase (Yoo et al., 1996; Sousa et al., 2002; Kwon et al., 2003).

The two components of AAA+ proteases separate protein unfolding/translocation and peptide cleavage into discrete, sequential activities; however, the two enzymes tightly couple their functions to support regulated protein degradation. Independently, neither the AAA+ unfoldase/remodeling enzyme nor the self-compartmentalized peptidase degrade folded proteins. Stressing the importance of coupling these enzyme activities, small-molecule agonists called acyldepsipeptides (ADEPs) that activate unfoldase-independent ClpP proteolysis kill bacteria as a result of uncontrolled degradation of nascent polypeptides and other unfolded proteins (Brötz-Oesterhelt et al., 2005; Kirstein et al., 2009; Lee et al., 2010a; Li et al., 2010; Schmitz et al., 2014).

Similarly, targeting the AAA+ enzyme with cyclic peptide antibiotics that stimulate ATPase activity also uncouples the unfoldase from the peptidase, demonstrating bacteriocidal activity against both growing and dormant mycobacteria (Schmitt et al., 2011; Vasudevan et al., 2013; Gavrish et al., 2014).

AAA+ Protein Superfamily

The AAA+ superfamily, which includes AAA+ unfoldases/remodeling enzymes, comprises a group of functionally diverse proteins found in all domains of life that contain a structurally-conserved AAA+ ATPase module of ~200–250 amino acids (Iyer et al., 2004; Snider et al., 2008; Wendler et al., 2012). The AAA+ module consists of an αβα/Rossmann-fold large domain, a connecting hinge-linker, and an α-helical small domain. ATP binding and hydrolysis occurs at the interface of the large and small domains, which have structural motifs that interact with the nucleotide and/or mediate catalysis (**Table 1.1** and **Figure 1.2A–C**). These AAA+ motifs can act in *cis* (within the active site of the same subunit) or in *trans* (by contacting the ATPase site in a neighboring subunit). Nucleotide-binding site interactions between the large and small domains of adjacent AAA+ modules contribute to subunit packing, thus stabilizing oligomeric assemblies that are shaped like rings or helical filaments (**Figure 1.2D–E**, respectively). Intersubunit communication also promotes coordinated conformational changes within the AAA+ oligomeric complex, as nucleotide-induced changes in one AAA+ module can be propagated to its neighbors (Erzberger and Berger, 2006; Snider et al., 2008).

By coupling ATP binding/hydrolysis to changes in conformation within an oligomeric assembly, many AAA+ proteins function as versatile molecular machines that perform energy-dependent mechanical work in a variety of cellular processes. In addition to protein quality control and regulated proteolysis, examples of processes using mechanical work by AAA+ proteins include DNA replication, which involves clamp loaders (*e.g.* replication factor C/RFC, DnaX), initiators

(*e.g.* Cdc6, DnaA), and helicases (*e.g.* MCM); vesicle transport by the microtubule motor protein dynein; homologous recombination and DNA repair by RuvB; and cofactor and pigment biosynthesis by cobaltochelatase and the magnesium chelatase BchI, respectively (Snider et al., 2008; Lundqvist et al., 2010). To perform these diverse functions, AAA+ proteins interact with other proteins and/or use specialized auxiliary domains and adaptors (which mediate substrate recognition; see also "Adaptor-mediated Recognition by Bacterial AAA+ Proteases", pp. 41–50), that support the core AAA+ module.

Table 1.1 Conserved motifs in nucleotide-binding sites of AAA+ proteins.
Structural elements in AAA+ modules that mediate nucleotide binding or catalysis, as summarized in (Erzberger and Berger, 2006; Snider et al., 2008).'x' is any residue and 'h' is any hydrophobic residue. A residue in parentheses refers to an optional residue, while those in brackets indicate the most commonly found residues at a given site. '*cis'* or '*trans*' refers to whether the motif contributes to function in the same AAA+ module or is provided from a neighboring subunit.

motif	consensus sequence	function	location in AAA+ module
box-II	none; within α-helix preceding β-sheet	adenine recognition (nucleotide binding)	*cis*
Walker-A	GxxGxGK[T/S]	interacts with β-phosphate of ATP; metal ion coordination (nucleotide binding)	*cis*
Walker-B	hhhhDE	metal ion coordination; catalytic base to activate water (catalysis)	*cis*
sensor-I	polar residue [N/T/H]	interacts with γ-phosphate of ATP or water in active site (catalysis)	*cis* or *trans*
sensor-II	R or positively charged residue	interacts with γ-phosphate of ATP (nucleotide binding and catalysis)	*cis* or *trans*
Arg finger	R(R)	interacts with γ-phosphate of ATP (catalysis)	*trans*

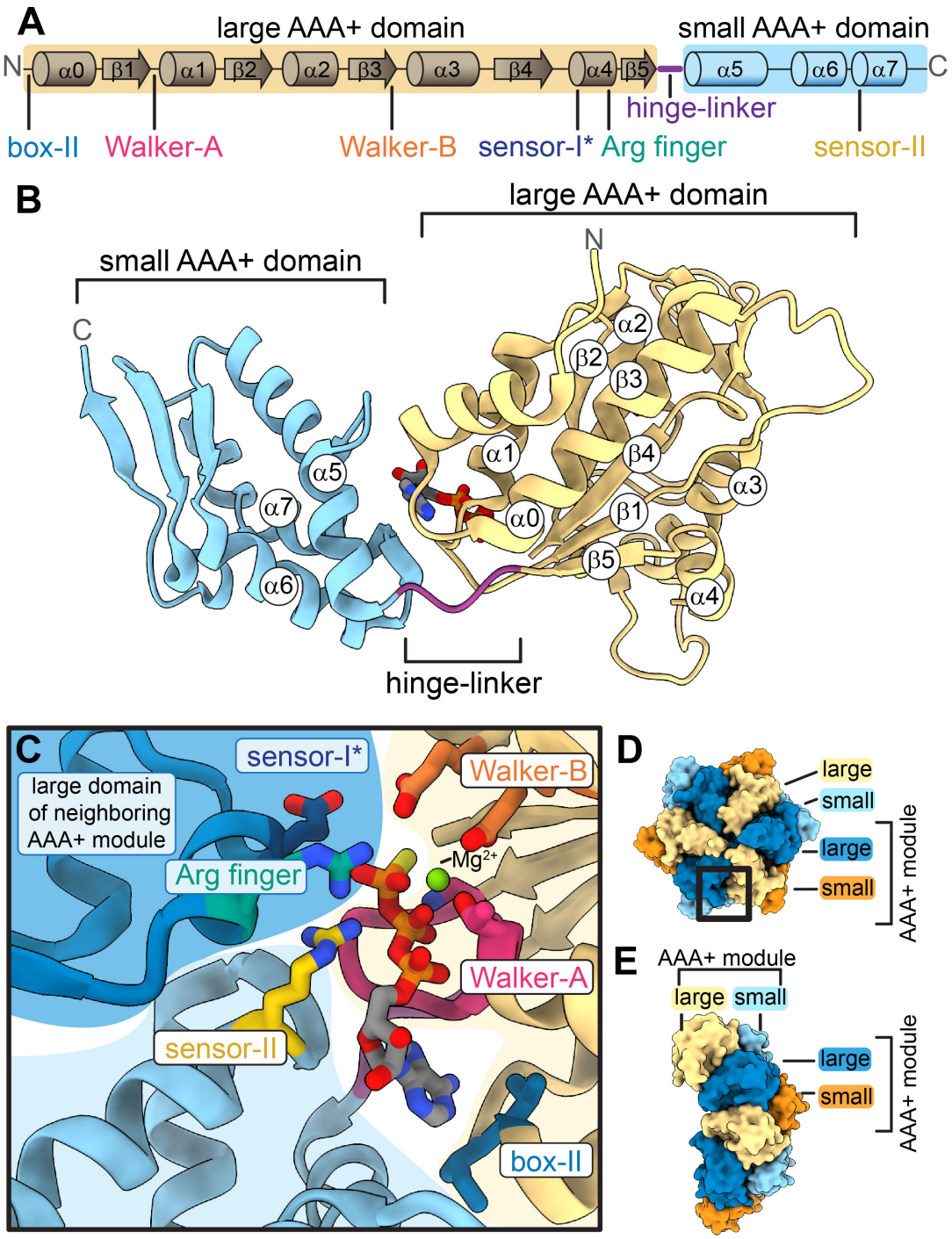


Figure 1.2 Conserved structural elements of AAA+ modules. (A) Secondary structural elements in a AAA+ module (*E. coli* ClpX). 'sensor-I*' refers to the Glu^{303}, which is not a canonical sensor-I residue. **(B)** Atomic model of a AAA+ module (*E. coli* ClpX PDB 6WRF), showing the large and small AAA+ domains. **(C)** Nucleotide-binding site showing structural elements involved in nucleotide binding and/or hydrolysis, colored as shown in (A). The large domain of an adjacent AAA+ module interacts with the small and large domain of the AAA+ module from (B); see boxed inset in (D) for reference. AAA+ modules can assemble into: **(D)** rings (PDB 6WRF) and **(E)** helical filaments (DnaA PDB 2HCB).

AAA+ proteins also contain additional structural features within the AAA+ module used to classify AAA+ clades, comprised of seven distinct evolutionary lineages (clamp loader, initiator, classic, superfamily III helicase, HCLR, H2-insert, and PS-II insert) (Iyer et al., 2004; Erzberger and Berger, 2006). AAA+ unfoldases/remodeling enzymes are members of either the classic or the HCLR clade. The insertion of secondary structural elements, as illustrated in **Figure 1.3**, distinguishes the classic clade (short α-helix between β2 and α2) from the HCLR clade (β-hairpin between α3 and β4). Classic clade members also lack a positively-charged sensor-II residue and have a second Arg finger. In some HCLR AAA+ unfoldases, such as ClpX, a flexible IGF/IGL loop (named after the residues Ile–Gly–Phe/Leu) stabilizes docking of the AAA+ unfoldase with its partner peptidase, ClpP (see next section on prokaryotic AAA+ proteases, p. 27). The β-hairpin insertion of ClpX contains a ClpX-specific RKH loop, named after its residues Arg-Lys-His, that functions similarly to pore-1 and pore-2 loops in substrate binding (Farrell et al., 2007; Martin et al., 2008; Fei et al., 2020a).

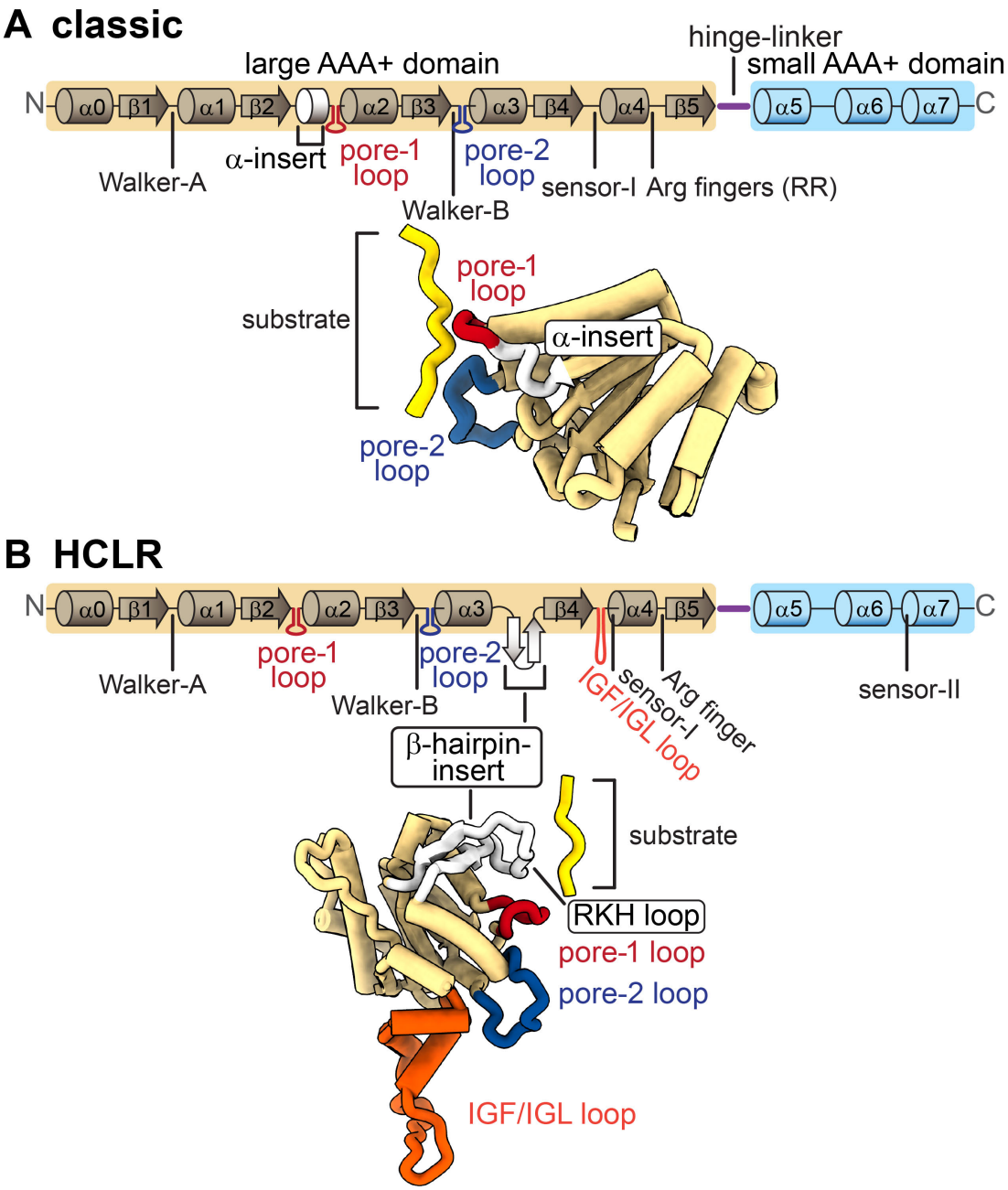


Figure 1.3 Classic and HCLR AAA+ clades. (A) Domain organization of the classic clade. Classic clade members lack a sensor-II arginine and have an additional Arg finger residue. Pore-1 and pore-2 loops contact the substrate and transmit an unfolding or remodeling force. A short α-helical insert precedes the pore-1 loop. Bottom panel shows the large AAA+ domain of a representative classic clade AAA+ unfoldase (Yme1 PDB 6AZ0). **(B)** Domain organization of the HCLR clade. Members of the HCLR clade contain a β-hairpin insertion before the sensor-I motif. The bacterial AAA+ unfoldases ClpA (D2) and ClpX also have an additional flexible IGL or IGF loop that mediates docking with the partner peptidase, ClpP (see next section on prokaryotic AAA+ proteases, p. 27). Bottom panel shows the large AAA+ domain of a representative HCLR clade member (ClpX PDB 6WRF). Adapted from Figure 2, Sauer *et al*., 2021.

The axial channel of both the classic and HCLR AAA+ unfoldases/remodeling enzymes is lined by pore-1 and pore-2 loops, which contact the protein substrate and transmit an unfolding or remodeling force powered by the energy of ATP hydrolysis. Each set of pore loops forms a spiral-staircase arrangement that wraps around the bound substrate polypeptide, contacting every other residue of the polypeptide backbone (**Figure 1.4A**). The pore-1 loop contains a conserved aromatic residue, usually Tyr, that interacts with the substrate polypeptide and is flanked by additional residues that modulate the characteristics of the pore-1 loop itself or the surrounding environment in the axial channel (Puchades et al., 2020). For example, positively-charged Lys or Arg residues in many classic clade AAA+ motors form cation–π interactions with the pore-1 aromatic residue, which stabilize subunit packing between adjacent AAA+ domains and the spiraling organization of pore loops in the axial channel (**Figure 1.4B**) (Gates et al., 2017; Han et al., 2017; Zehr et al., 2017; de la Peña et al., 2018; Lopez et al., 2020). Alternatively, the conserved aromatic residue in the pore-1 loop may be flanked by hydrophobic residues, such as Val or Trp, and by a Gly residue, which increases flexibility and changes the orientation of the aromatic sidechain (see tyrosines in **Figure 1.4C**) (Puchades et al., 2020).

Pore-2 loops often mediate specialized functions specific to an enzyme's biological activity. In contrast to the pore-1 loop, the substrate-interacting residues of pore-2 loops are also less conserved (Puchades et al., 2020). The pore-2 loops of the microtubule-severing enzymes katanin and spastin contain positively charged residues that neutralize the negative charge within their tubulin substrates (Johjima et al., 2015; Sandate et al., 2019; Zehr et al., 2020; Han et al., 2020). Aromatic residues in the pore-2 loops of Yme1, AFG3L2, and human LONP1 contribute additional hydrophobic interactions with substrate, which are proposed to help these enzymes degrade hydrophobic membrane proteins and unfolded proteins (Puchades et al., 2017, 2019, 2020; Shin et al., 2021). The sequence diversity of pore-1 and pore-2 loops thus reflects the varying functions of these loops in different AAA+ unfoldases/remodeling enzymes.

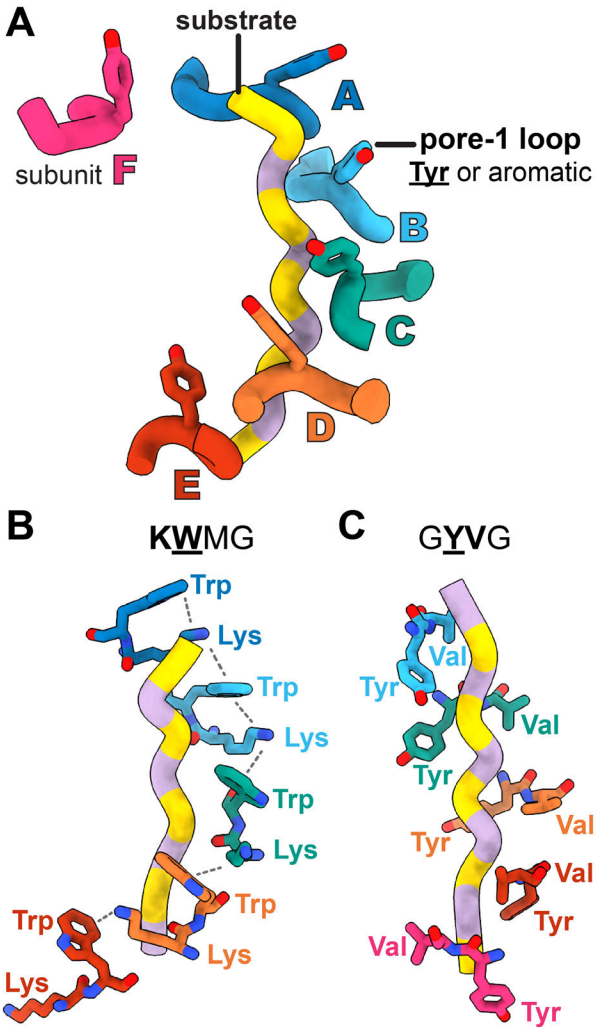


Figure 1.4 Substrate-binding modes of pore-1 loops.
(A) Spiraling organization of pore-1 loops. The conserved aromatic pore-1 loop residues contact two-residue segments of the substrate polypeptide (PDB 6AZ0). Alternating amino acids are colored yellow or light purple to depict the substrate-binding pattern.
For (B) and (C), only the bolded amino acids are shown in sticks, for clarity. Not all pore-1 loops have modeled structures.
(B) Positively-charged residues flanking the pore-1 aromatic residue contribute electrostatic interactions, such as cation-π, in yeast Vps4 (PDB 6BMF).
(C) Hydrophobic amino acids, such as Val, contribute additional hydrophobic interactions to the pore-1 aromatic residue in ClpX (PDB 6WSG). Additionally, the preceding Gly next to the pore-1 Tyr increases flexibility; these pore-1 loop Tyr residues point downward, rather than upward as shown in panel (A).
Adapted from Figure 5, Puchades *et al.*, 2020.

AAA+ unfoldases/remodeling enzymes can be classified as single-ring or double-ring enzymes, based on whether constituent subunits carry a single AAA+ module or tandem AAA+ modules. Each ring forms from the packing of small and large AAA+ domains of adjacent subunits, which usually self-assemble into a hexamer of homotypic AAA+ modules. Therefore, the assembly of two stacked rings, in which each AAA+ module hexamerizes into its own ring, constitutes most double-ring AAA+ unfoldases/remodeling enzymes. Reflecting the vast structural and functional diversity of AAA+ proteins, double-ring enzymes have evolved by multiple mechanisms. For example, the NSF/Cdc48/Pex family arose from internal gene duplication from a single ancestral member, whereas the tandem AAA+ modules of the ClpABC/Hsp104 family, referred to as D1 and D2, appear to evolved from gene fusion of phylogenetically distinct AAA+ modules (Iyer et

al., 2004; Erzberger and Berger, 2006). Specifically, the N-terminal AAA+ module of this latter family, D1, shares sequence and structural features of the classic clade, but the C-terminal module, D2, is part of the HCLR clade (Iyer et al., 2004).

In summary, all AAA+ proteins have conserved structural features within the core AAA+ module. Variations in the organization of secondary structural elements, as well as the insertion of auxiliary domains and/or additional AAA+ modules (in double-ring enzymes), diversify AAA+ proteins, allowing these enzymes to carry out numerous, specialized cellular functions. Within individual families of AAA+ remodeling/enzymes, specific residues in the pore-1 and pore-2 loops can fine-tune substrate interactions to modulate substrate specificity. In the next two sections, I will discuss how these structural features are adapted by AAA+ proteases to support regulated proteolysis in prokaryotes and eukaryotes.

Prokaryotic AAA+ Proteases

Prokaryotes possess several AAA+ proteases: FtsH, HslUV, Lon, ClpXP, ClpAP (usually in gram-negative bacteria) or ClpCP and ClpEP (generally in gram-positive bacteria), and the proteasomal Mpa•20S (in actinobacteria) or PAN•20S (in archaea) (Sauer and Baker, 2011; Gur et al., 2013; Mahmoud and Chien, 2018). The domain organization of these bacterial AAA+ proteolytic systems is provided in **Figure 1.5**. Specialized structural features, such as auxiliary domains and specific active-site residues within the peptidase domains, facilitate subcellular localization (*e.g.*, transmembrane regions in the membrane-bound proteases FtsH and LonB), efficient function under various growth conditions, and recognition of distinct classes of substrates.

The presence of multiple AAA+ proteases in prokaryotes suggests that each enzyme may operate in a defined cellular context. For example, in *Bacillus subtilis*, expression of the ClpP-associated ClpE unfoldase is tightly controlled by elevated temperatures, such that ClpE protein levels are

hardly detectable under standard growth conditions but rapidly accumulate following heat shock at 50°C (Derré et al., 1999; Gerth et al., 2004). Like ClpE, both ClpC and ClpX also partner with ClpP. In contrast to ClpE, protein abundance of ClpC and ClpX remains constant during exponential growth and into the entry of stationary phase. Furthermore, different enzymes have distinct substrate specificities. Compared to ClpXP, which selectively recognizes substrates with specific sequences called degrons (covered in further detail, see "Substrate Recognition by AAA+ Proteases in Bacteria," p. 35), the Lon protease recognizes a broad range of damaged or misfolded proteins that have little sequence similarity via hydrophobic regions that would otherwise be buried in the core of natively folded structures (Gur, 2013; Mahmoud and Chien, 2018). The membrane-bound proteases FtsH (found in most bacteria) and LonB (a Lon isoform usually found in archaea) degrade membrane proteins, as well as some cytoplasmic substrates (Gur, 2013; Bittner et al., 2017). Generally, ClpAP is the only AAA+ protease responsible for degrading N-end-rule substrates in most bacteria (see p. 39), albeit in mycobacteria, which do not encode a ClpA homolog, ClpCP is also capable of this activity (Ziemski et al., 2021).

Multiple AAA+ proteases may exist in a cell as auxiliary backup systems, as exemplified by the partially-overlapping substrate specificity of ClpXP/ClpAP (*e.g.*, ssrA-tagged substrates), Lon/ClpAP (*e.g.*, the replication initiator DnaA in *Caulobacter crescentus*), and Lon/HslUV (*e.g.*, RcsA, a transcription factor that activates the regulation of capsule synthesis operon, and the cell-division inhibitor SulA in *E. coli*) (Gottesman et al., 1998; Wu et al., 1999; Liu et al., 2016). Functional redundancy can be revealed by single-gene deletion mutants, as a single AAA+ protease may be dispensable for degradation of a substrate recognized by a second protease. Redundant AAA+ proteases may help ensure rapid, thorough elimination of shared substrates, especially essential regulatory proteins, such as DnaA, that become toxic when accumulated (Liu et al., 2016).

Beyond their central roles in protein quality control and proteostasis, AAA+ proteases regulate growth, cell development and differentiation, and environmental adaptation in many bacteria (Elsholz et al., 2017; Mahmoud and Chien, 2018). In *E. coli*, *Borrelia burgdorferi*, *Bradyrhizobium japnonicum*, *Helicobacter pylori*, and *Streptococcus mutans*, the membrane-bound FtsH protease is essential for viability (Tomoyasu et al., 1993; Zhongming and Taylor, 1996; Narberhaus et al., 1999; Chu et al., 2016; Wang et al., 2021). By contrast, in *C. crescentus*, *ftsH* is dispensable for growth; instead, the cytoplasmic protease ClpXP is essential and is required for cell cycle progression to degrade several key regulatory proteins, such as the master regulator CtrA and PdeA, a phosphodiesterase that modulates levels of the global bacterial secondary messenger cyclic di-GMP (Jenal and Fuchs, 1998; Christen et al., 2005; Abel et al., 2011). In *B. subtilis*, ClpCP, ClpXP, and FtsH control multiple developmental processes, including competence, sporulation, biofilm formation, and motility (Elsholz et al., 2017). AAA+ proteases also mediate stress responses to DNA damage, nutrient starvation, and the accumulation of reactive oxygen species, promoting cell survival in changing environmental conditions and consequently, contributing also to the development of antimicrobial resistance (Galhardo et al., 2007; Pruteanu and Baker, 2009; Poole, 2012; Händel et al., 2015; Harms et al., 2016).

Furthermore, AAA+ proteases have important roles in virulence regulation in many pathogens, making these enzymes promising drug targets for antimicrobial therapeutic use (Butler et al., 2006; Pressler et al., 2016; Culp and Wright, 2017; Elsholz et al., 2017; Mahmoud and Chien, 2018). Similar to the human 26S proteasomal inhibitors bortezomib, carfilzomib, and ixazomib used to treat multiple myeloma and lymphoma, ClpP can be chemically targeted by covalent active-site inhibitors using β-lactone and phenyl ester scaffolds, as well as by the naturally-produced azabicyclenes (Hong et al., 2019; Patteson et al., 2019; Jayaweera et al., 2021; Culp et al., 2022). However, due to the conserved catalytic mechanism of AAA+ proteases, many of these covalent ‘suicide’ protease inhibitors exhibit off-target inhibition of human orthologs. Hence,

other therapeutic strategies that perturb unfoldase–peptidase interactions in regions that are more divergent between bacterial and human enzymes, such as ADEP ClpP activators and AAA+ unfoldase uncouplers (described previously, pp. 18–19) can achieve specificity against bacterial proteases.

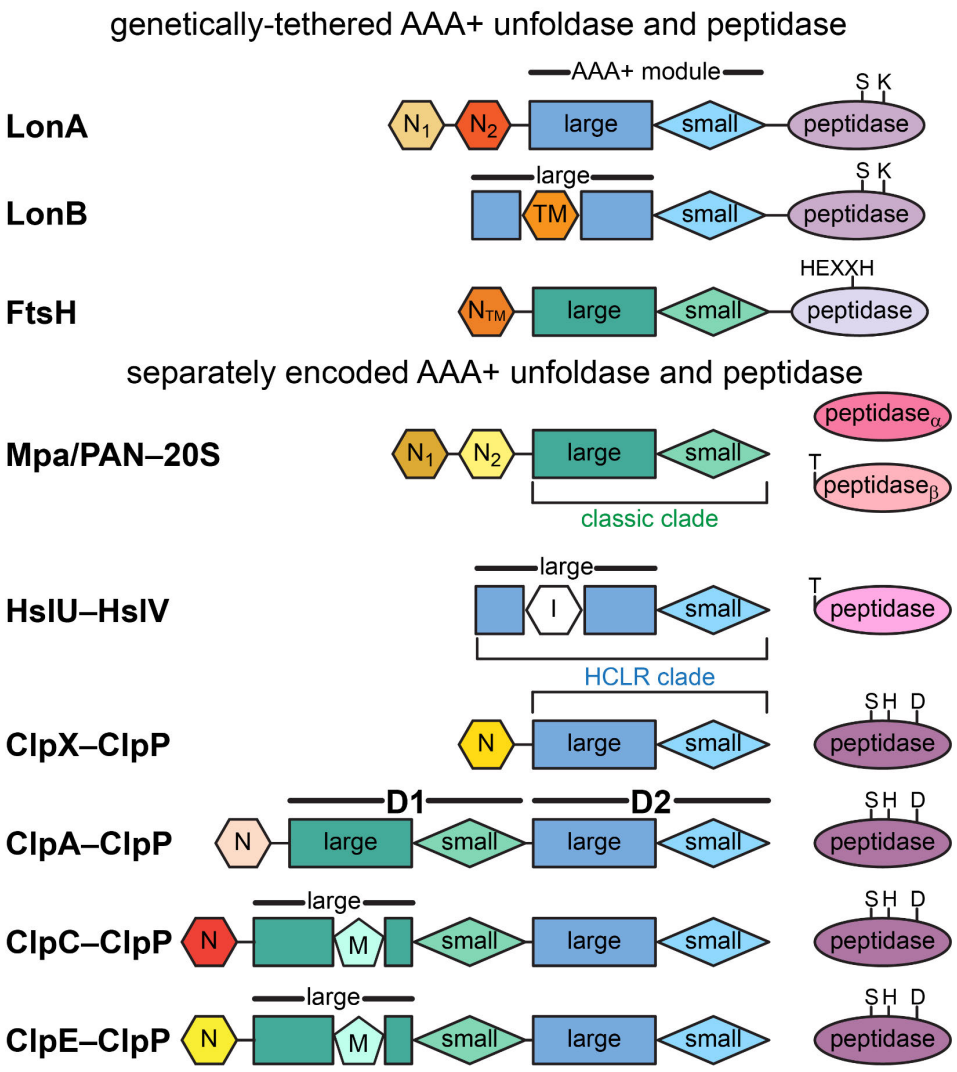


Figure 1.5 Bacterial AAA+ proteases. Domain organization of AAA+ proteases found in prokaryotes showing auxiliary domains, such as the N-domain (N), transmembrane-domain (TM), intermediate-domain (I), and middle-domain (M); AAA+ modules; and peptidase domains (ovals). Insertions in the AAA+ modules are shown by a break in the large AAA+ domain cartoon. For separately encoded subunits that form a functional AAA+ protease complex, the unfoldase is on the left side of the hyphen in the label, and the peptidase is on the right. The AAA+ unfoldase may consist of a single AAA+ module or tandem AAA+ modules called D1 and D2. Classic clade AAA+ modules are indicated in green; HCLR clade AAA+ modules are indicated in blue. The peptidase active-site residues are provided in text above. Adapted from Figure 2, Sauer and Baker, 2011 and Figure 1.1, Gur *et al*., 2013.

Prokaryotic AAA+ proteases consist of various assemblies of AAA+ unfoldases and self-compartmentalized peptidases (**Figure 1.5**). As mentioned earlier (p. 16), some AAA+ unfoldase and peptidase partners, such as Lon and FtsH, are genetically-tethered in the same polypeptide, whereas others are encoded by two separate genes and associate to form a two-component complex. All AAA+ proteases share a similar architecture constituted by coaxially-stacked rings, in which the unfoldase and the peptidase each assemble into ring-shaped oligomers. For example, the ClpXP complex is composed of a single $ClpX_6$ ring and two back-to-back heptameric rings of a $ClpP_{14}$ tetradecamer. Additionally, AAA+ proteases formed by the complex of separate unfoldase and peptidase components can adopt a singly-capped (in which only one AAA+ ring unfoldase docks with its peptidase partner) or a doubly-capped structure (in which two AAA+ rings dock on both sides of the peptidase rings) (Grimaud et al., 1998; Ramachandran et al., 2002).

Each unfoldase pairs exclusively with a specific self-compartmentalized peptidase, such that HslU partners with the HslV peptidase, but not ClpP; and ClpX, ClpA, ClpC, and ClpE partner with ClpP—but not HslV. Interestingly, HslU does not contain an IGF/IGL (or more broadly, the consensus tripeptide [L/I/V]-G-[F/L]) loop, which ClpX/A/C/E unfoldases use to dock with ClpP (Kim et al., 2001). HslU instead docks with HslV via its C-terminal tails, with the consensus sequence D-L-[S/A]-[R/K/A]-[Y/F]-I-L (Seong et al., 2002). This difference in docking mechanisms is also reflected in the subunit interface of the unfoldase and peptidase. HslUV contains a symmetrical docking interface between the $HslU_6$ ring hexamer and the two-stacked hexameric rings of $HslV_{12}$ (Bochtler et al., 2000; Sousa et al., 2000). By contrast, the docking interface ClpX and other unfoldases that partner with ClpP features a mismatch between the hexameric AAA+ unfoldase and the seven-fold symmetric ClpP rings (Kessel et al., 1995; Grimaud et al., 1998; Beuron et al., 1998; Bewley et al., 2006).

In addition to possessing several AAA+ proteases, some prokaryotes also encode multiple copies of AAA+ protease components. In actinobacteria, cyanobacteria, and subsets of other phyla, multiple ClpP isoforms further diversify the composition of AAA+ proteases (Yu and Houry, 2007; Nishimura and Van Wijk, 2015). Alternative ClpP isoforms in some bacteria modulate biofilm formation and virulence, as well as physiological adaptation to stress conditions (Gaillot et al., 2001; Qiu et al., 2008; Raju et al., 2012; Hall et al., 2017; Bhandari et al., 2018; Mawla et al., 2021). ClpP isoforms may assemble into homo-heptamers (made up of only one ClpP isoform) or hetero-heptamers (ring formed by a mixture of multiple isoforms) that can be combined to create homocomplexes (*e.g.*, $ClpP_{14}$) or heterocomplexes (*e.g.*, two stacked rings of different homoheptamers or two stacked rings of heteroheptamers). Thus, combinations of different ClpP subunits can generate ClpP assemblies with distinct biochemical activities, such as peptide cleavage specificities and interactions with specific AAA+ unfoldases. Similarly, the 20S peptidase (usually found in mycobacteria and archaea), also known as the core particle, is composed of two different types of peptidase subunits: α, which lacks peptidolytic activity and forms the outer rings of the core particle, and β, which contains peptidase active-site residues and constitutes the inner rings at the center of the core particle (Maupin-Furlow, 2011). The 20S peptidase assembles into four stacked rings with an $\alpha_7\beta_7\beta_7\alpha_7$ arrangement, in which the outer α_7 rings dock with the unfoldase Mpa (from mycobacteria) or PAN (from archaea) to form prokaryotic proteasomes (**Figure 1.6**). Both the prokaryotic and eukaryotic proteasomal unfoldases contain a hydrophobic-Tyr-X (any residue) motif used to dock with the α-rings of their 20S partners (Maupin-Furlow, 2011).

Eukaryotic AAA+ Proteases

Eukaryotes principally utilize a single AAA+ protease, the 26S proteasome, in their main cellular compartments, the nucleus and cytosol, although several homologs of bacterial AAA+ proteases are also found in mitochondria, chloroplasts, and other organelles derived from prokaryotic

endosymbionts (Glynn, 2017; Bard et al., 2018; Bouchnak and van Wijk, 2021). Sharing the stacked-ring architecture of all AAA+ proteases, the 26S proteasome consists of the 20S core particle peptidase and the 19S RP, or regulatory particle (containing the AAA+ unfoldase) (**Figure 1.6**). The 19S RP can be subdivided into the lid (a multiprotein subcomplex functioning in regulation, recognition, and scaffolding) and the base (a hetero-oligomeric AAA+ unfoldase ring containing six different ATPase Rpt1-6 subunits in complex with the non-ATPase Rpn1, Rpn2, Rpn13 subunits) (Bard et al., 2018).

The composition of the 26S proteasome varies depending on the assembly of proteasomal subunits under different conditions or expression in certain tissues, such as the thymus (Murata et al., 2009). The eukaryotic 20S core peptidase differs from its prokaryotic counterparts, as it is composed of seven distinct α and seven different β subunits, of which only three β subunits are active (Maupin-Furlow, 2011). Interestingly, each catalytic β subunit has a unique peptide-bond cleavage preference despite using the same catalytic mechanism, as demonstrated by the activities of caspase-like β1 (acidic residues), trypsin-like β2 (basic residues), and chymotrypsin-like β5 (hydrophobic residues) (Arendt and Hochstrasser, 1997; Rousseau and Bertolotti, 2018). The immunoproteasome, a specialized form of the constitutive 26S proteasome, contains alternative catalytic β subunits, β1i, β2i, and β5i. These immunosubunits are conditionally expressed in response to proinflammatory cytokines and oxidative stress and have altered cleavage specificities compared to their constitutively expressed counterparts (Murata et al., 2018; Rousseau and Bertolotti, 2018). The 26S proteasome's elaborate three-part structure allows multi-level regulation of proteasomal activity: *e.g.*, controlling the assembly of lid, base, and core components, recruitment to specific subcellular sites, and modulation by post-translational modifications (Livneh et al., 2016; Bard et al., 2018).

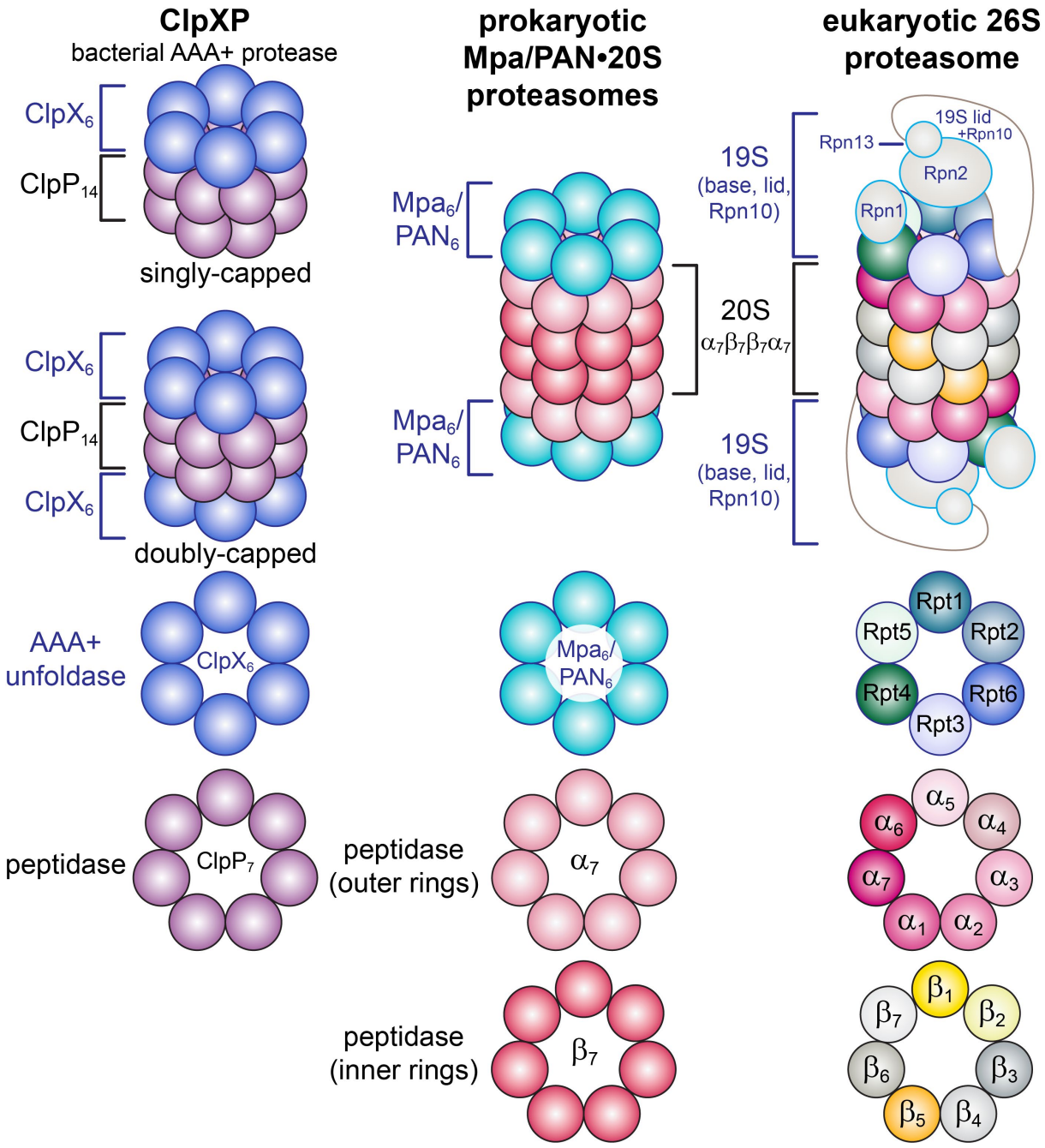


Figure 1.6 Stacked-ring assembly of AAA+ proteases. AAA+ proteases in all domains of life assemble as stacked rings and can be singly-capped (*e.g.*, $ClpX_6$:$ClpP_{14}$) or doubly-capped (*e.g.*, $ClpX_6$:$ClpP_{14}$:$ClpX_6$). Bacterial AAA+ proteases, such as ClpXP, consist of a homohexameric AAA+ unfoldase and usually a homo-oligomeric peptidase, such as $ClpP_{14}$. Prokaryotic proteasomes have identical AAA+ unfoldase subunits and homoheptameric α_7 and β_7 rings. In the eukaryotic 26S proteasome, both the AAA+ unfoldase (Rpt1–6) and peptidase (α_7 and β_7) rings are composed of distinct subunits. The eukaryotic 26S proteasome consists of the 19S regulatory particle—which is formed by the base (containing the Rpt ring unfoldase), lid, and Rpn10—and the 20S peptidase. Both the prokaryotic Mpa/PAN and the eukaryotic 19S regulatory particle dock only to the outer α_7 rings of the 20S peptidase.

SUBSTRATE RECOGNITION BY AAA+ PROTEASES IN BACTERIA

Bacterial AAA+ proteases use multiple strategies to recognize and degrade specific classes of substrate proteins. In the simplest form of recognition, an unstructured peptide sequence (called a degron) in a substrate protein binds specifically to the axial pore loops in the AAA+ unfoldase (**Figure 1.7A**) (Sauer and Baker, 2011). Substrate recognition may also require assistance from additional interactions with the auxiliary domains of the AAA+ unfoldase (**Figure 1.7B**), which occur directly (enzyme•substrate), or are mediated by adaptor proteins (see next section, p. 41) via an enzyme•adaptor•substrate complex (**Figure 1.7C**). Although degrons have been defined in multiple ways, in this thesis, a degron refers to the minimal element sufficient for specific recognition by bacterial AAA+ proteases, or more simply a pore-binding peptide sequence (Varshavsky, 1991; Yang and Bedford, 2012). By contrast, eukaryotic recognition by the 26S proteasome requires both a proteasome-binding tag (a polyubiquitin chain covalently attached to a lysine residue in the substrate) and an initiation region (an unstructured polypeptide tail usually ~20–30 residues long) from which the proteasome engages the substrate polypeptide for processive unfolding and translocation (Inobe et al., 2011; Tomita and Matouschek, 2019).

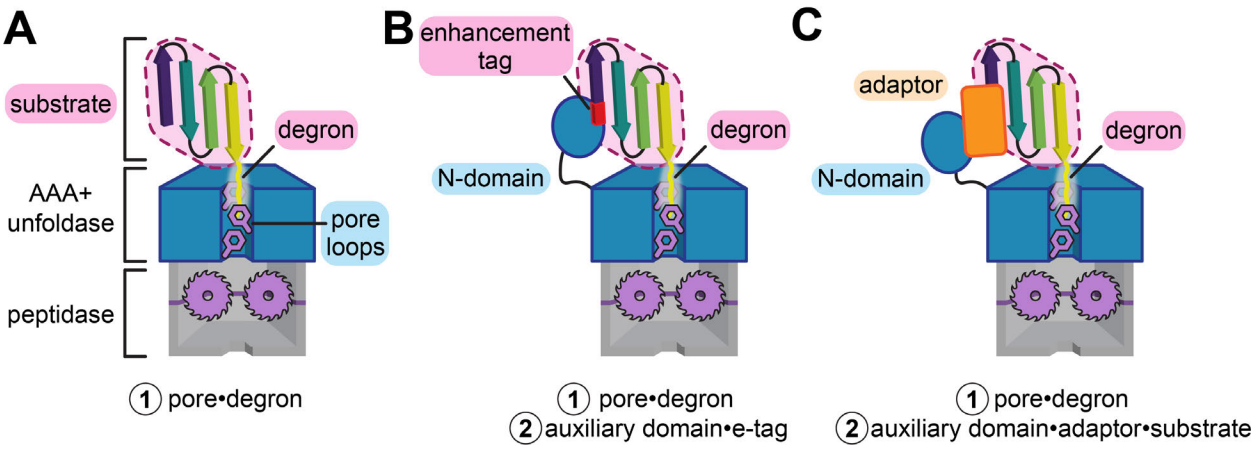


Figure 1.7 Substrate recognition interactions. (A) All substrates require a degron for binding to the pore loops of the AAA+ unfoldase. **(B)** Some substrates also utilize an enhancement tag (or e-tag; colored red) to bind to an auxiliary domain. Only one N-domain (out of six in the AAA+ hexamer) is shown here for clarity. **(C)** Alternatively, an adaptor protein (colored orange) can tether the substrate to the N-domain via an indirect recognition mechanism. Serrated circles in peptidase chamber represent peptidolytic active-site residues.

As mentioned earlier, different AAA+ proteases in bacteria have distinct substrate specificities, often due to the recognition of specific degrons. However, some have overlapping substrate specificities, which recognize a subset of shared substrates either by interacting with the same degron or with different degrons in the same protein. Because binding to a substrate protein requires conformational access to the degron, modulating the exposure of degron sequences is a critical regulatory mechanism for substrate recognition. Degrons are often found at or near the N– or C–termini of substrate proteins, as these areas are frequently solvent-exposed, but can also be located internally (Hoskins et al., 2002; Okuno et al., 2006; Gur and Sauer, 2008a). Degrons can become conditionally exposed by *(i)* proteolytic cleavage or processing to generate fragments with accessible degrons (*e.g.*, LexA, RseA), *(ii)* conformational change, *(iii)* protein unfolding (*e.g.*, following heat shock), and *(iv)* subunit dissociation (Baker and Sauer, 2006).

ssrA-tagging and ssrA-like Degrons

The ssrA tag is a degron that is appended to the C-termini of nascent polypeptides via *trans*-translation, a quality control mechanism used in bacteria to rescue stalled ribosomes. Ribosomes often stall on aberrant mRNAs *i.e.* those that lack a stop codon, which may result from incomplete transcription or nuclease cleavage; mRNAs that contain rare codons; and in response to conditions that promote translational stalling, such as amino acid starvation (Roche and Sauer, 1999, 2001; Karzai et al., 2000; Hayes et al., 2002). The *ssrA* gene encodes transfer-messenger RNA (tmRNA), which has properties of both tRNA and mRNA, containing the ssrA-peptide-coding open reading frame (ORF). In the ssrA-tagging system, this tmRNA molecule is charged with alanine and forms a complex with SmpB (small protein B) and EF-Tu (elongation factor Tu) (Karzai et al., 2000; Janssen and Hayes, 2012). The Ala-tmRNA•SmpB•EF-Tu ternary complex binds to the A site of a stalled ribosome, which then transfers the nascent polypeptide chain to the alanine on the tmRNA. The stalled ribosome resumes translation by switching from the aberrant mRNA to the stop codon-carrying ssrA ORF (in the tmRNA message), thus allowing

translational termination and ribosome recycling (**Figure 1.8**). ssrA-tagging occurs frequently, with an estimated one in 20 to 200 translation events in *E. coli* resulting in ssrA tag addition (Roche and Sauer, 2001; Lies and Maurizi, 2008). These abundant ssrA-tagged polypeptides are recognized and degraded by the AAA+ proteases ClpXP, ClpAP, and FtsH in most bacteria or by Lon in many *Mycoplasma* species, which lack Clp enzymes (Gottesman et al., 1998; Herman et al., 1998; Choy et al., 2007; Gur and Sauer, 2008b; Ge and Karzai, 2009).

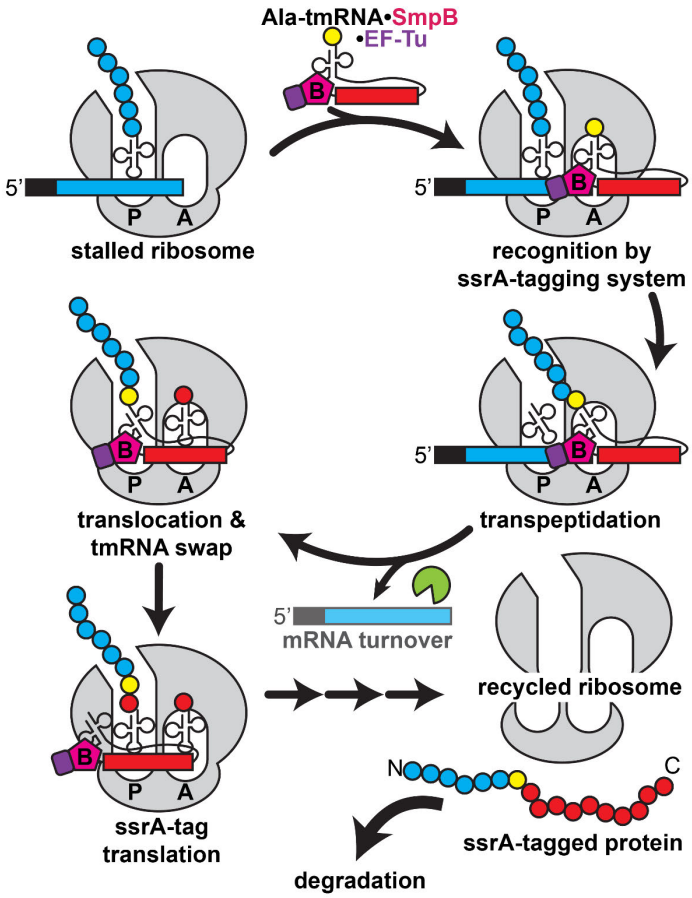


Figure 1.8 ssrA-tagging system.
The Ala-tmRNA•SmpB•EF-Tu complex binds to a stalled ribosome. Transpeptidation proceeds, during which the nascent polypeptide is transferred to the alanine in the tmRNA. Then, translation resumes, switching from the mRNA causing the stall over to the tmRNA message (the ssrA tag ORF). Once the tmRNA encoding the ssrA tag is translated, translation is terminated, freeing the ribosome for recycling and the ssrA-tagged protein. EF-Tu is shown in purple, SmpB is in magenta, the ssrA ORF is in red, and the alanine is in yellow. The green shape represents RNase cleavage of the stall-causing mRNA species.
Adapted from Figure 1, Janssen and Hayes, 2012.

ClpAP and ClpXP recognize partially overlapping regions of the ssrA tag, which has the sequence AANDENYALAA-COO$^-$ in *E. coli* (Keiler et al., 1996). ClpA interacts with AANDENYALAA-COO$^-$ (underlined residues), likely recognizing a pattern of aliphatic residues. By contrast, ClpX, which binds LAA-COO$^-$ in its axial channel, recognizes specific side chains localized near the C-terminus (Flynn et al., 2001; Fei et al., 2020b). Furthermore, this three-residue sequence in the ssrA tag is

the representative member of a class of ssrA-like ClpXP degrons, named C-motif 1 (Flynn et al., 2003). Importantly, the other C-motif 1 degrons do not require ssrA-tagging and instead, are encoded endogenously in the polypeptide sequence. The extent of ClpAP recognition of proteins bearing these C-motif 1 degrons is still poorly characterized, as is the conservation of multiple protease recognition sites within the ssrA tag across bacterial species. For adaptor-assisted recognition of the ssrA tag, see p. 42.

Enhancement Tags

Enhancement tags (e-tags) are peptide sequences in substrate proteins that bind to the auxiliary domains of AAA+ proteases and function in conjunction with a degron, which binds to pore loops in the unfoldase axial channel (**Figure 1.7A**, p. 35). Serving as molecular tethers in the enzyme•substrate complex, e-tags often enhance the affinity of a weakly-recognized degron by increasing the effective concentration of the substrate (**Figure 1.7B**). In the UmuD/UmuD′ heterodimer, the full-length UmuD subunit is tethered to the ClpX N-domain, while UmuD′ (which is generated by DNA damage-induced autocleavage) binds in the ClpX channel (Neher et al., 2003). In the absence of the N-domain or as a result of mutating e-tag residues that interact with the N-domain, degradation of UmuD′ is severely diminished (Gonzalez et al., 2000; Neher et al., 2003). Interestingly, the UmuD/D′ heterodimer forms preferentially, minimizing degradation of full-length $UmuD_2$ homodimers. Because the DNA damage-inducible UmuD′ subunit lacks the e-tag, only the UmuD precursor can dock with the N-domain and thus, specifically targets its *trans*-subunit (UmuD′) for ClpXP degradation.

By coupling multiple recognition signals (at least one degron and at least one e-tag), e-tags enable combinatorial control of substrate unfolding/remodeling, such that accessibility of both the pore-binding degron (or more simply, the pore-tag) and e-tag are necessary for optimal degradation. Furthermore, substrates featuring e-tags also allow specific oligomeric states to be selectively recognized, which is especially advantageous for remodeling and/or disassembly

reactions of multimers or protein complexes. For example, complex disassembly of the MuA tetramer by ClpX is most efficient using a combination of an intrinsically weak pore-tag and two or more e-tags. Although both MuA monomers and multimers contain the MuA pore-tag and the N-domain-binding e-tag, only the MuA tetramer, which supports multivalent ClpX interactions, is efficiently remodeled. Spatial constraints of these recognition signals prevent simultaneous binding to the ClpX N-domain and pore within a single subunit, allowing MuA monomers to bind only one of the two ClpX sites at a given time (Ling et al., 2015).

The Bacterial N-end Rule Pathway

The N-end rule pathway refers to the proteolytic systems in bacteria and eukaryotes that recognize proteins containing N-degrons usually defined by a single amino acid at the N-terminus of the substrate polypeptide. Historically, the 'N-end rule' categorized N-terminal residues as 'destabilizing' (promoting degradation) or 'stabilizing' (not promoting degradation) (Varshavsky, 2011). The identity of N-end (de)/stabilizing residues varies among different organisms, as do the proteolytic mechanisms utilized to recognize N-degrons. In eukaryotes, E3-ubiquitin ligases recognize N-end-rule substrates, resulting in their covalent attachment to ubiquitin (a highly-conserved eukaryotic protein that marks ubiquitin-conjugated proteins for degradation by the 26S proteasome) (Varshavsky, 2011, 2019). In bacteria, the N-end-rule degradation machinery consists of ClpAP, its adaptor ClpS (see p. 44), and the aminoacyl-transferases Aat and Bpt (Tobias et al., 1991; Shrader et al., 1993; Erbse et al., 2006; Graciet et al., 2006; Varshavsky, 2011, 2019). *E. coli* ClpS directly recognizes proteins bearing a Phe, Trp, Tyr, or Leu at the N-terminal position, or a primary destabilizing N-end residue, and delivers these substrates to ClpAP for degradation (**Figure 1.9**). Alternatively, the L/F-transferase Aat appends Leu (or Phe) primary destabilizing residues to proteins bearing N-terminal Lys or Arg (hence referred to as 'secondary' destabilizing residues). Bpt is an L-transferase that conjugates Leu to Asp and Glu secondary destabilizing residues (Varshavsky, 2019).

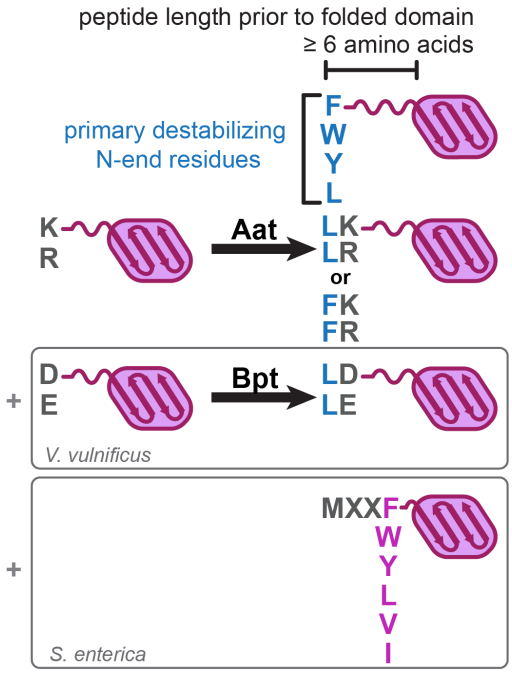


Figure 1.9 The bacterial N-end rule. ClpAPS degrades proteins containing a Phe, Trp, Tyr, or Leu residue at the first N-terminal position (primary destabilizing N-end residues) and an N-terminal peptide tail of six or more residues (includes the N-end residues). The modifying enzyme Aat appends a primary N-end Leu or Phe to the residues indicated, while Bpt (which is present in *Vibrio vulnificus*) appends only a primary N-end Leu. In *Salmonella enterica*, hydrophobic residues at the fourth position from the N-terminus are also recognized by ClpAPS.

Bacterial N-end-rule degrons are generated through various mechanisms and are modulated by additional sequence elements that influence the strength of ClpAPS recognition. In addition to a destabilizing N-end residue, the N-end-rule substrate also requires an unstructured peptide 'tail' of at least six residues prior to its folded domain for productive degradation, presumably to provide a long enough sequence for pore-loop engagement by ClpAP (Wang et al., 2008a). Primary N-end residues are typically revealed by proteolytic processing (Humbard et al., 2013). However, some studies suggest that initiating Met residues may not need to be removed for proteins containing formylated-Met[1] or hydrophobic residues at the fourth position from the N-terminus (Piatkov et al., 2015; Gao et al., 2019). Furthermore, residues adjacent to the N-degron can influence recognition, as shown by the inhibitory effect of negatively charged side chains at the second position and the importance of a short hydrophobic motif within the first 12 residues of a subset of N-end-rule substrates (Wang et al., 2008a; Ninnis et al., 2009).

Post-translational Modifications and Tagging as Degradation Markers

Prokaryotic AAA+ proteases can use post-translational tagging systems to selectively target proteins for degradation, without relying on the proteolytic machine or an adaptor to recognize specific amino acid sequences within substrates. In *B. subtilis*, arginine phosphorylation by the kinase McsB marks proteins for degradation by ClpCP (Kirstein et al., 2007; Elsholz et al., 2012; Trentini et al., 2016). In mycobacteria, the enzymes PafA (proteasome accessory factor A) and Dop (deamidase of Pup) attach specific lysine residue(s) of substrate proteins to a C-terminal diglycine motif in Pup (prokaryotic ubiquitin-like protein). The eukaryotic 26S proteasomal ubiquitination system similarly features conjugation of a small protein (ubiquitin) containing a C-terminal diglycine onto lysine residues in the substrate (Striebel et al., 2014). However, unlike ubiquitin, Pup is an intrinsically disordered protein, except for the formation of a helix (res. 21–64) when docked with Mpa•20S (Wang et al., 2010). Pupylated proteins dock directly with the N-domains of the AAA+ protease and the Pup sequence is engaged by pore loops in the AAA+ channel (Striebel et al., 2014; Kavalchuk et al., 2022). By contrast, ubiquitinated proteins dock to the lid complex in the 26S proteasome, which removes ubiquitin chains from the substrate prior to engagement of its unstructured initiation region by the Rpt-ring AAA+ unfoldase for degradation (Bard et al., 2018). Whether or not Pup is degraded by Mpa•20S following engagement by pore loops or recycled, in a manner analogous to ubiquitin removal and recycling in eukaryotes, is an active area of investigation (Striebel et al., 2010; Zerbib et al., 2021).

ADAPTOR-MEDIATED RECOGNITION BY BACTERIAL AAA+ PROTEASES

Adaptor proteins expand the repertoire of substrate recognition and selection, providing another critical level of conditional regulation that can be induced in response to changing conditions. Adaptors often function as molecular matchmakers, tethering substrates to the auxiliary domains of AAA+ proteases and thus, tighten affinity by increasing the substrate's effective concentration relative to the unfoldase channel (**Figure 1.7C**, p. 35) (Sauer and Baker, 2011). Alternatively,

some adaptors modulate the exposure of degrons in the substrate (*e.g.*, the adaptor RssB and its substrate, the alternative sigma factor RpoS) or tune the substrate specificity of the AAA+ unfoldase itself, as suggested by the ClpXP 'priming' activity of the CpdR adaptor in *C. crescentus* (Mahmoud and Chien, 2018). Additional proteins, called anti-adaptors, inhibit adaptor-mediated recognition by binding to adaptors, preventing adaptor•substrate interactions (Battesti and Gottesman, 2013). Consequently, the combination of multiple adaptor and anti-adaptor proteins can create complex, multi-tiered hierarchies that regulate proteolysis during bacterial growth, development, and adaptation to stress (Joshi et al., 2015; Mahmoud and Chien, 2018). Here, I describe SspB and ClpS in further detail as examples of representative adaptors.

SspB, a specificity-enhancement factor for ClpXP

Although ClpXP degrades ssrA-tagged substrates in the absence of an adaptor protein, SspB further enhances ClpXP recognition, tightening the K_M for degradation of an ssrA-tagged GFP substrate by ~five-fold and increasing the V_{max} by ~25% (Levchenko et al., 2000). Each subunit of the SspB homodimer is composed of a substrate-binding domain and a flexible C-terminal tail containing the XB motif (LRVVK, residues 161–165) that interacts with the ClpX N-domain (Song and Eck, 2003; Levchenko et al., 2003; Wah et al., 2003). Both SspB tails are needed for optimal delivery of ssrA-tagged substrates, but only one tail binds per ClpX N-domain dimer (**Figure 1.10A**). The ClpX N-domain also dimerizes, limiting the total number of available SspB binding sites to three N-domain dimers (per ClpX hexamer). Consequently, only one SspB dimer likely binds to ClpXP, occupying two of the three possible SspB tail binding sites (Bolon et al., 2004).

SspB differentially regulates ClpXP and ClpAP recognition of ssrA-tagged substrates due to its interactions with specific residues in the ssrA tag. The SspB substrate-binding domain binds to a portion of the ssrA tag, AANDENYALAA-COO^- (underlined residues), that does not overlap with the critical C-terminal segment needed for ClpX pore-loop engagement (LAA-COO^-) (Flynn et al., 2001). Both the ClpX pore loops and SspB can concurrently bind an ssrA-tagged protein in the

adaptor•substrate•ClpXP ternary complex, albeit with relatively minor steric clashes due to the close spacing between the two binding sites in the ssrA tag (Hersch et al., 2004). Conversely, SspB and ClpA compete for binding to overlapping residues (**AA**NDENYALAA-COO$^-$, ClpA-binding residues underlined, SspB/ClpA overlap in bold), allowing SspB to function both as a competitive inhibitor for ssrA-tagged protein degradation by ClpAP and an enhancer for ClpXP degradation of these substrates (Flynn et al., 2001). Interestingly, the N-terminal fragment of RseA (residues 1–108), a ClpXP substrate, competes with ssrA-tagged substrates for binding to overlapping sites in SspB, demonstrating the interplay of substrates, adaptors, and AAA+ proteases in recognition (Levchenko et al., 2005).

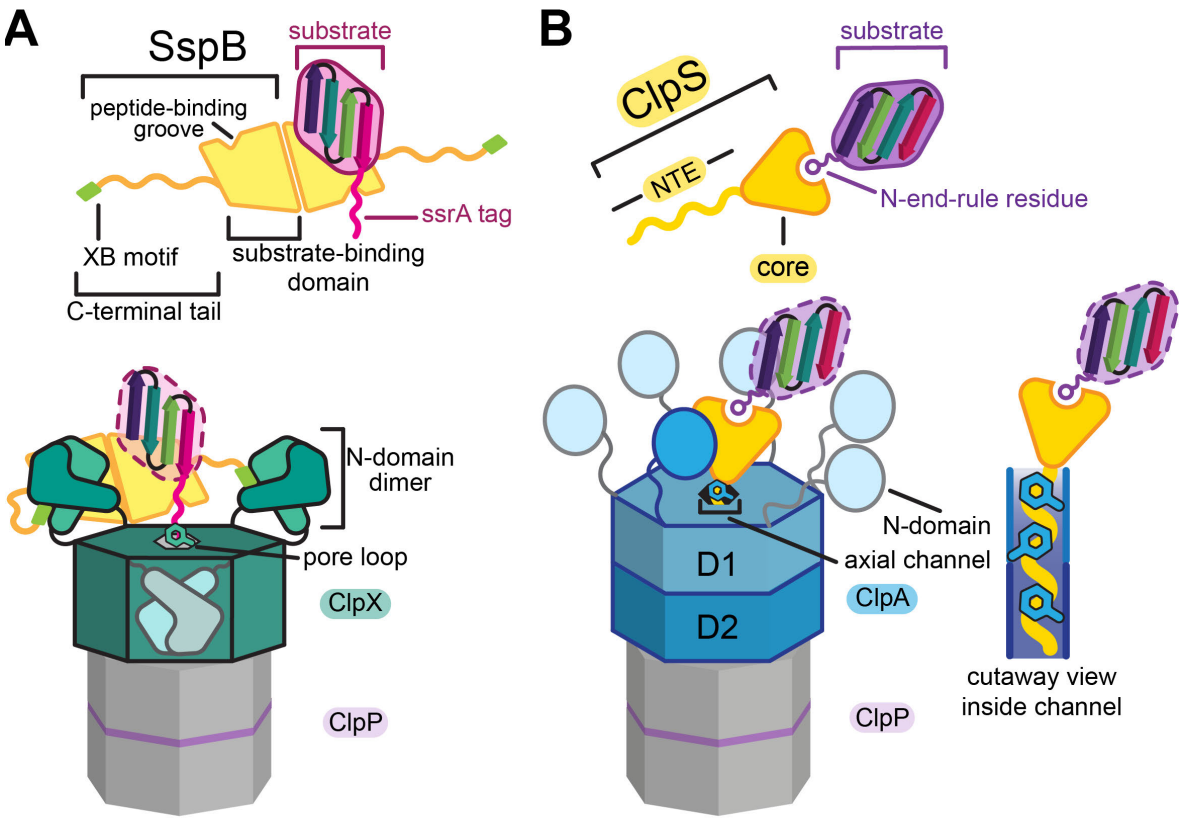


Figure 1.10 Adaptor–AAA+ protease delivery complexes. N-domains not bound by adaptor proteins are in gray outlines. **(A)** The SspB dimer tethers an ssrA-tagged substrate to the ClpX N-domain dimers. The C-terminal tail containing the XB (ClpX-binding) motif of each SspB subunit binds to a distinct ClpX N-domain dimer. An ssrA-tagged substrate can simultaneously bind to SspB and pore loops in the ClpX axial channel. **(B)** ClpAPS•N-end-rule substrate high-affinity delivery complex. The ClpS core domain binds the ClpA N-domain, and the ClpS NTE is engaged by ClpA pore loops.

ClpS, an adaptor protein that reprograms ClpAP

ClpS is a bifunctional regulator, promoting ClpAP degradation of N-end-rule substrates and inhibiting degradation of ssrA-tagged substrates and other classes of substrates recognized directly by ClpAP (Dougan et al., 2002; Erbse et al., 2006; Wang et al., 2008b; Torres-Delgado et al., 2020). ClpS (~10 kDa) contains a normally intrinsically disordered N-terminal extension (NTE, residues 1–25) and a folded core domain (residues 26–106) that has a hydrophobic N-end residue binding pocket (**Figure 1.11A**). The ClpS substrate-binding pocket, in which Asn^{47}, Asp^{49}, and His^{79} (in *C. crescentus*) participate in hydrogen bonding with the protonated α-amino group of the N-end residue, is preformed and accessible in the free adaptor (Wang et al., 2008b; Román-Hernández et al., 2009, 2011). Structures of the ClpS pocket bound to substrates carrying a primary destabilizing N-end residue are nearly identical to the *apo* structure (**Figure 1.11B**) (Román-Hernández et al., 2009). Consequently, binding of the N-end-rule substrate to ClpS has a relatively low entropic cost, as this step only requires proper conformational selection of a small number of primary destabilizing N-end rotamers. Conversely, ClpS binding to poorly recognized N-terminal residues, such as Met, is disfavored due to having a much larger range of possible rotamers (carries a larger entropic penalty), or as a result of steric clashes with other residues in the ClpS pocket (Wang et al., 2008b; Román-Hernández et al., 2009). For example, ClpS Met^{53} excludes β-branched amino acids, such as Ile and Val. An M53A ClpS variant has expanded N-end-rule specificity, as this mutant has a larger binding pocket and can thus deliver substrates carrying these bulkier amino acids for efficient degradation by ClpAP (**Figure 1.11C**) (Wang et al., 2008b).

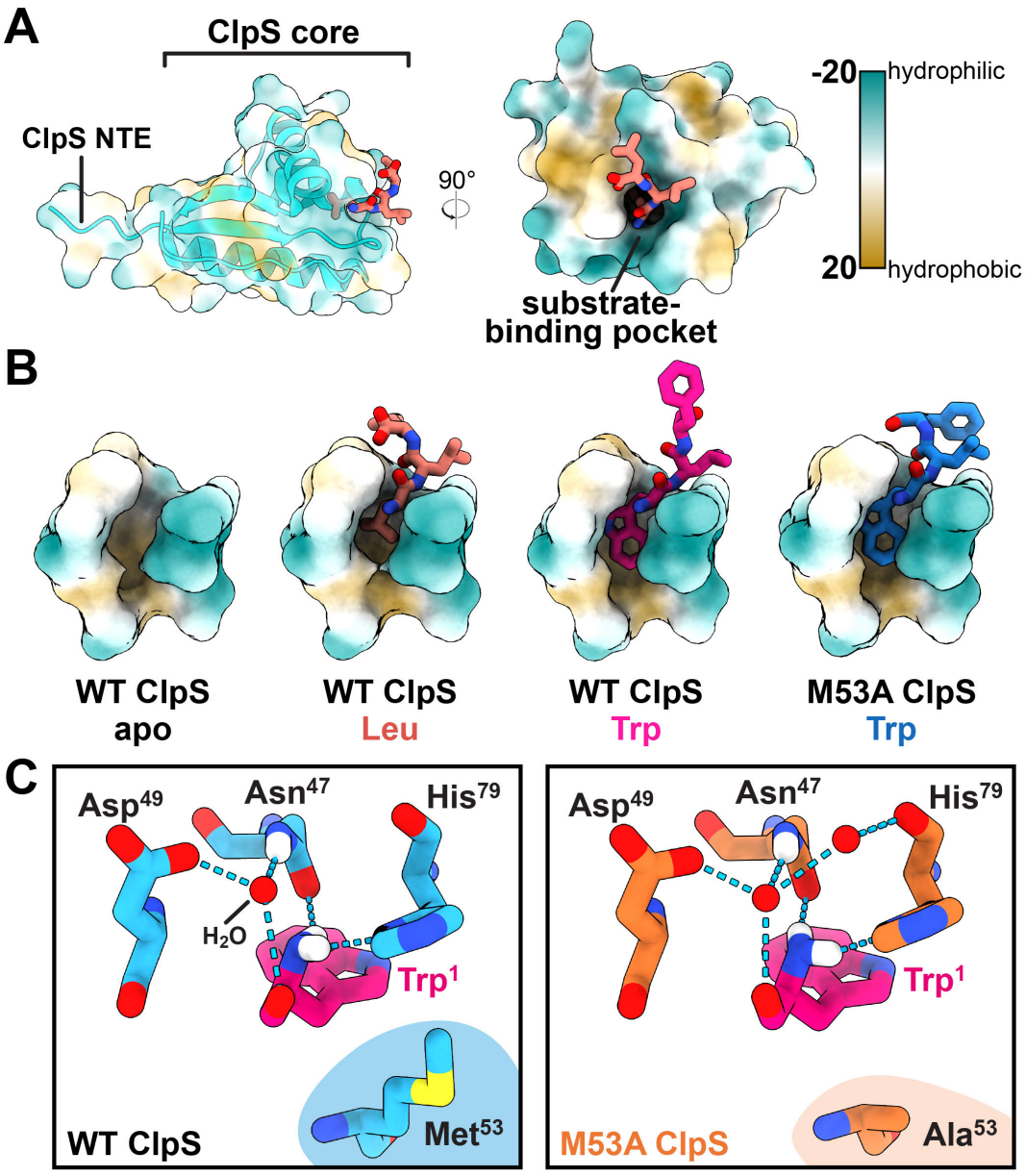


Figure 1.11 ClpS preformed substrate-binding pocket.
(A) The ClpS core (*C. crescentus* PDB 3G19) contains a hydrophobic binding pocket. The ClpS core is colored as teal (hydrophilic) and gold (hydrophobic) surfaces. **(B)** Side views of the ClpS binding pocket in the absence (PDB 3GQ0) or presence of primary destabilizing N-end peptides, Leu (PDB 3G19) and Trp (PDB 3GQ1 and 3G1B with M53A ClpS), indicated in the second row of labels. ClpS residues 45, 48, 49, 51, 53, 56, 75, 78, 79, and 112 are shown as surfaces.
(C) Hydrogen bonding with Asn^{47}, Asp^{49}, and His^{79} stabilizes the N-end Trp residue. The M53A ClpS mutant has a larger binding pocket than WT (lower right). Adapted from Figure 2 and 4, Wang *et al.*, 2008b.

The ClpS NTE interacts with pore loops in the ClpA axial channel, while the ClpS core binds to the ClpA N-domain (Zeth et al., 2002; Wang et al., 2008b; Román-Hernández et al., 2009, 2011; Rivera-Rivera et al., 2014). Both these binding modes are required to form high-affinity ClpAPS•N-degron delivery complexes (**Figure 1.10B**). The affinity of $ClpA_6$•ClpS•N-degron peptide assembly is < 50 nM. When full-length ClpS is replaced with just the ClpS core domain, the K_D for the ternary complex assembled without the ClpS NTE is ~1.5 μM (30-fold weaker) (Román-Hernández et al., 2011). Similarly, mutation of residues in the ClpS core domain that contact the ClpA N-domains severely impairs binary assembly of $ClpA_6$•ClpS complexes (Zeth et al., 2002). Furthermore, protease•adaptor•substrate ternary complex assembly is also stabilized by the presence of substrate in the binding pocket within the ClpS core domain (Román-Hernández et al., 2011). Mutation of *E. coli* ClpS His^{66} (analogous to *C. crescentus* His^{79}, which forms hydrogen bonds with the α-amino group of the substrate's N-terminal residue) to alanine weakens $ClpA_6$•ClpS•N-degron peptide assembly ~9 fold, supporting the importance of this residue in adaptor-mediated substrate binding enhancement (Román-Hernández et al., 2011). Thus, all interfaces involved in ClpAPS•N-degron complexes are important in assembling high-affinity N-end-rule substrate delivery complexes.

Many bacteria and organelles derived from prokaryotic endosymbionts possess ClpS orthologs, and some α-proteobacteria, as well as cyanobacteria, also encode a second paralog, ClpS2 (Lupas and Koretke, 2003; Tryggvesson et al., 2015; Stein et al., 2016). The *clpS* gene is usually located in the same operon as *clpA*. By contrast, *clpS2* genes are not found in the same operon as *clpA* and *clpS(1)*. Furthermore, *clpS2* genetic loci are scattered throughout the genomes of many different bacterial groups, supporting the possibility that this second class of ClpS proteins may have been acquired through horizontal gene transfer (Lupas and Koretke, 2003; Tryggvesson et al., 2015). Reflecting the evolutionary divergence of ClpS1 (the 'parental' ClpS well-conserved across α-proteobacteria) and ClpS2, the two ClpS isoforms have distinct substrate

preferences, which result from differences in the substrate-binding pocket, as demonstrated for *Agrobacterium tumefaciens* ClpS2 (Tryggvesson et al., 2015; Stein et al., 2016).

Examining sequence conservation of the ClpS protein reveals additional insights regarding the key functions of the ClpS NTE and the highly-conserved substrate-binding core domain (Dougan et al., 2002; Guo et al., 2002). Supporting the hypothesis that bacterial and eukaryotic recognition of N-end-rule substrates have a common evolutionary origin, the ClpS core shares significant sequence homology to the substrate-binding pocket of Ubr1 *N*-recognins, a family of eukaryotic E3 ubiquitin ligases that function like ClpS to recognize hydrophobic N-end residues in the eukaryotic N-end rule pathway (Lupas and Koretke, 2003; Varshavsky, 2011). Conversely, the ClpS NTE sequence is poorly conserved among ClpS orthologs, with the exception of ~4-5 residues adjacent to the ClpS core domain, called the junction sequence (Hou et al., 2008; Román-Hernández et al., 2011). For proper NTE function, length but not sequence appears critical. The ClpS NTE sequence can be substituted with a Gly-Ser-Lys linker with no observable difference in ClpS activity from the wild-type sequence, while truncating the NTE by 13 or more residues severely impairs N-end-rule substrate degradation by ClpAPS (Hou et al., 2008; Román-Hernández et al., 2011).

Despite the lack of significant NTE sequence conservation that otherwise suggests this region is dispensable, the ClpS NTE is a critical component of adaptor function. Beyond its immediate role in ternary complex assembly with ClpAP and N-degron substrates, the NTE participates in multiple reaction steps of the N-end-rule pathway. The intrinsically disordered NTE is a ‘degron-mimic’ that lacks tertiary structure and is engaged by ClpA pore loops in the same fashion as a substrate polypeptide (see Chapter 3, p. 86) (Rivera-Rivera et al., 2014). Consequently, ClpAP attempts to unfold and degrade the substrate-like ClpS adaptor. ClpS however resists proteolysis by ClpAP. Structural features, such as the stable ClpS core domain and proline

residues in the junction sequence, protect this 'pseudo-substrate' from successful degradation (see Chapter 4, p. 134) (Dougan et al., 2002; Román-Hernández et al., 2011; Rivera-Rivera et al., 2014; Rivera-Rivera, 2015).

Furthermore, these degron-like interactions of the NTE within the ClpA channel are essential for ClpS modulation of ClpAP activity. ClpS tightens the K_M for N-end-rule substrate degradation, while weakening the K_M for degradation of ssrA-tagged substrates via a mixed inhibition mechanism (Dougan et al., 2002; Erbse et al., 2006; Wang et al., 2008b; Torres-Delgado et al., 2020). Interestingly, ClpA contains an intrinsic degron and is itself a ClpAP substrate (Maglica et al., 2008). The presence of ClpS inhibits this autodegradation, perhaps through a similar mechanism used to inhibit degradation of ssrA-tagged substrates (Dougan et al., 2002). ClpS also reduces the ATP hydrolysis rate of ClpAP ~two-fold and therefore slows down unfolding and translocation by ClpAP, indicating allosteric control of the ATPase sites (especially those in the D2 ring) and global enzyme activity by ClpS (Hou et al., 2008; Torres-Delgado et al., 2020). Disrupting the interactions between the ClpA pore loops and the ClpS NTE diminishes *(i)* formation of ClpAPS complex assembly, *(ii)* ClpS-assisted enhancement of N-end-rule substrate degradation, and *(iii)* ClpS inhibition of ssrA-tagged substrate degradation (Torres-Delgado et al., 2020; Zuromski et al., 2021). Most importantly, the ClpS NTE is the key structural feature used to transmit force from the ClpA unfoldase to the substrate-bound ClpS core domain, as proposed by our current model of ClpS-assisted N-end-rule substrate delivery (see below).

In contrast to the SspB dimer, which uses its C-terminal tails to tether ssrA-tagged substrates to promote ClpXP recognition, ClpS is a monomeric adaptor that delivers N-end-rule substrates to ClpAP via an 'active' delivery mechanism. That is, ClpS delivery requires ATP-fueled mechanical work to induce conformational changes in the ClpS•substrate complex. The substrate is bound to the ClpS core domain distal to the ClpA axial pore loops, and the N-degron (bearing the primary

destabilizing N-end residue) is hidden in the ClpS substrate-binding pocket. Consequently, the ClpA channel cannot bind/engage the N-end-rule substrate without remodeling ClpS, which is bound to ClpA pore loops via its NTE. Therefore, substrate delivery by ClpS appears to require ATP-hydrolysis-driven remodeling of the ClpS core domain to promote substrate transfer from ClpS to ClpA (**Figure 1.12**) (Román-Hernández et al., 2011; Rivera-Rivera et al., 2014; Zuromski et al., 2021).

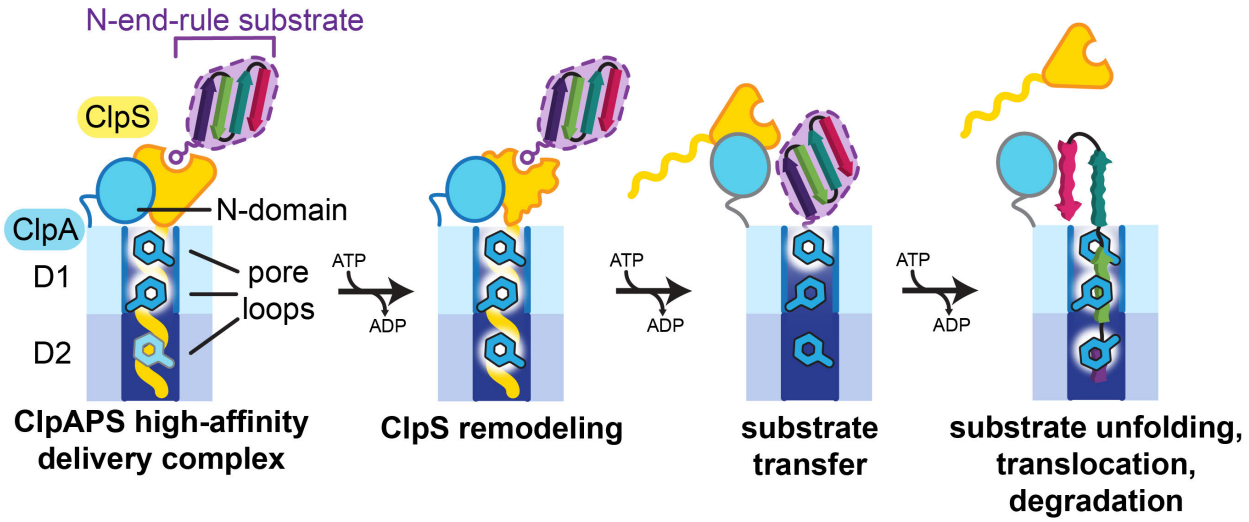


Figure 1.12 ClpS-assisted N-end-rule substrate delivery mechanism. Pore loops in the D1 ring are required for ClpAP•ClpS•N-end-rule substrate complex assembly. The grayed-out loop in the D2 ring indicates these pore-1 loops are not required. Powered by ATP hydrolysis, pore-1 loops in both rings transmit a remodeling force to the ClpS core that results in substrate transfer from ClpS to ClpA. The ClpS adaptor is likely released from the ClpA axial channel, and the N-end-rule substrate is processively unfolded and translocated into ClpP for degradation.

Briefly, in the proposed mechanism of ClpS-assisted substrate delivery, upon formation of the high-affinity ClpAPS•N-end-rule substrate ternary complex, pore loops from both D1 and D2 rings in the ClpA axial channel 'pull' on the NTE and remodel the ClpS core. This ClpA-dependent conformational change in the ClpS adaptor likely destabilizes the high-affinity ternary complex, which may lead to transfer of the substrate from the ClpS core to the ClpA axial channel. Subsequently, ClpA pore loops engage the N-end-rule substrate, instead of the ClpS NTE,

allowing processive substrate unfolding and subsequent translocation into the ClpP chamber for degradation.

The precise reaction sequence involved in ClpS-assisted degradation of N-end-rule substrates, especially the substrate handoff step, remains to be fully characterized. Furthermore, little is known about the progression of structural interactions between substrate or adaptor•substrate complexes and many of the AAA+ unfoldases during degradation, and their functional implications to AAA+ protease mechanisms. For example, many AAA+ proteases feature a symmetry mismatch between the AAA+ ring hexamer and self-compartmentalized peptidase heptamer. The next chapter of my thesis describes the subunit interface mismatch involved in a separately-encoded unfoldase (ClpA) and peptidase (ClpP) complex and whether rotation of these two components of a AAA+ protease is required for unfolding, translocation, and degradation activity. Work in my thesis also focuses on the structural basis of ClpS NTE engagement by the pore-1 and pore-2 loops of the ClpA D1 and D2 rings (Chapter 3, p. 86) and the contribution of a Pro^{24}-Pro^{25} sequence in the junction region of the NTE towards ClpS protection from ClpAP proteolysis (Chapter 4, p. 134). In Chapter 5 (p. 155), I describe how another AAA+ protease, ClpXP, selectively recognizes the inactive, *apo* form the transcriptional regulator FNR under aerobic conditions in *E. coli*, showing how the structural properties of a substrate control degradation by a AAA+ protease in response to oxygen exposure. Then I explore the biochemical and biological implications of my work, as well as prospective studies that could be performed to address remaining questions and to test models I propose in Chapter 6 (p. 195). Finally, in the Appendix (p. 213), I present *in vitro* and *in vivo* characterizations of ClpXP-dependent degradation of an extracytoplasmic function sigma (σ) factor, σ^{AntA}, that regulates morphological differentiation and antimycin biosynthesis in streptomycetes, a genus of Gram-positive bacteria known for their ability to produce bioactive secondary metabolites.

REFERENCES

Abdelhakim, A.H., Oakes, E.C., Sauer, R.T., and Baker, T.A. (2008). Unique contacts direct high-priority recognition of the tetrameric Mu transposase-DNA complex by the AAA+ unfoldase ClpX. Mol. Cell *30*, 1, 39–50.

Abel, S., Chien, P., Wassmann, P., Schirmer, T., Kaever, V., Laub, M.T., Baker, T.A., and Jenal, U. (2011). Regulatory cohesion of cell cycle and cell differentiation through interlinked phosphorylation and second messenger networks. Mol. Cell *43*, 4, 550–560.

Arendt, C.S., and Hochstrasser, M. (1997). Identification of the yeast 20S proteasome catalytic centers and subunit interactions required for active-site formation. Proc. Natl. Acad. Sci. U. S. A. *94*, 14, 7156–7161.

Baker, T.A., and Sauer, R.T. (2006). ATP-dependent proteases of bacteria: recognition logic and operating principles. Trends Biochem. Sci. *31*, 12, 647–653.

Bard, J.A.M., Goodall, E.A., Greene, E.R., Jonsson, E., Dong, K.C., and Martin, A. (2018). Structure and function of the 26S proteasome. Annu. Rev. Biochem. *87*, 697–724.

Battesti, A., and Gottesman, S. (2013). Roles of adaptor proteins in regulation of bacterial proteolysis. Curr. Opin. Microbiol. *16*, 2, 140–147.

Baumeister, W., Walz, J., Zühl, F., and Seemüller, E. (1998). The proteasome: Paradigm of a self-compartmentalizing protease. Cell *92*, 3, 367–380.

Beuron, F., Maurizi, M.R., Belnap, D.M., Kocsis, E., Booy, F.P., Kessel, M., and Steven, A.C. (1998). At sixes and sevens: Characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease. J. Struct. Biol. *123*, 3, 248–259.

Bewley, M.C., Graziano, V., Griffin, K., and Flanagan, J.M. (2006). The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. J. Struct. Biol. *153*, 2, 113–128.

Bhandari, V., Wong, K.S., Zhou, J.L., Mabanglo, M.F., Batey, R.A., and Houry, W.A. (2018). The role of ClpP protease in bacterial pathogenesis and human diseases. ACS Chem. Biol. *13*, 6, 1413–1425.

Bieniossek, C., Schalch, T., Bumann, M., Meister, M., Meier, R., and Baumann, U. (2006). The molecular architecture of the metalloprotease FtsH. Proc. Natl. Acad. Sci. U. S. A. *103*, 9, 3066.

Bittner, L.M., Arends, J., and Narberhaus, F. (2017). When, how and why? Regulated proteolysis by the essential FtsH protease in *Escherichia coli*. Biol. Chem. *398*, 5–6, 625–635.

Bochtler, M., Hartmann, C., Song, H.K., Bourenkov, G.P., Bartunik, H.D., and Huber, R. (2000). The structures of HsIU and the ATP-dependent protease HsIU-HsIV. Nature *403*, 6771, 800–805.

Bolon, D.N., Wah, D.A., Hersch, G.L., Baker, T.A., and Sauer, R.T. (2004). Bivalent tethering of SspB to ClpXP is required for efficient substrate delivery: a protein-design study. Mol. Cell *13*, 3, 443–449.

Botos, I., Melnikov, E.E., Cherry, S., Tropea, J.E., Khalatova, A.G., Rasulova, F., Dauter, Z., Maurizi, M.R., Rotanova, T. V., Wlodawer, A., et al. (2004). The catalytic domain of *Escherichia coli* Lon protease has a unique fold and a Ser-Lys dyad in the active site. J. Biol. Chem. *279*, 9, 8140–8148.

Bouchnak, I., and van Wijk, K.J. (2021). Structure, function, and substrates of Clp AAA+ protease systems in cyanobacteria, plastids, and apicoplasts: A comparative analysis. J. Biol. Chem. *296*, 100338.

Brinkmann, K., Schell, M., Hoppe, T., and Kashkar, H. (2015). Regulation of the DNA damage response by ubiquitin conjugation. Front. Genet. *6*, 98.

Brötz-Oesterhelt, H., Beyer, D., Kroll, H.P., Endermann, R., Ladel, C., Schroeder, W., Hinzen, B., Raddatz, S., Paulsen, H., Henninger, K., et al. (2005). Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. Nat. Med. *11*, 10, 1082–1087.

Butler, S.M., Festa, R.A., Pearce, M.J., and Darwin, K.H. (2006). Self-compartmentalized bacterial proteases and pathogenesis. Mol. Microbiol. *60*, 3, 553–562.

Burton, B.M., and Baker, T.A. (2005). Remodeling protein complexes: Insights from the AAA+ unfoldase ClpX and Mu transposase. Protein Sci. *14*, 8, 1945–1954.

Chen, B., Retzlaff, M., Roos, T., and Frydman, J. (2011). Cellular strategies of protein quality control. Cold Spring Harb. Perspect. Biol. *3*, 8, 1–14.

Choy, J.S., Aung, L.L., and Karzai, A.W. (2007). Lon Protease Degrades Transfer-Messenger RNA-Tagged Proteins. J. Bacteriol. *189*, 18, 6564.

Christen, M., Christen, B., Folcher, M., Schauerte, A., and Jenal, U. (2005). Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. J. Biol. Chem. *280*, 35, 30829–30837.

Chu, C.Y., Stewart, P.E., Bestor, A., Hansen, B., Lin, T., Gao, L., Norris, S.J., and Rosa, P.A. (2016). Function of the *Borrelia burgdorferi* FtsH homolog is essential for viability both *in vitro* and *in vivo* and independent of HflK/C. mBio *7*, 2, e00404-16.

Culp, E., and Wright, G.D. (2017). Bacterial proteases, untapped antimicrobial drug targets. J. Antibiot. (Tokyo). *70*, 4, 366–377.

Culp, E.J., Sychantha, D., Hobson, C., Pawlowski, A.C., Prehna, G., and Wright, G.D. (2022). ClpP inhibitors are produced by a widespread family of bacterial gene clusters. Nat. Microbiol. *7*, 3, 451–462.

Derré, I., Rapoport, G., Devine, K., Rose, M., and Msadek, T. (1999). ClpE, a novel type of Hsp100 ATPase, is part of the CtsR heat shock regulon of *Bacillus subtilis*. Mol. Microbiol. *32*, 3, 581–593.

Dougan, D.A., Reid, B.G., Horwich, A.L., and Bukau, B. (2002). ClpS, a substrate modulator of the ClpAP machine. Mol. Cell *9*, 3, 673–683.

Effantin, G., Maurizi, M.R., and Steven, A.C. (2010). Binding of the ClpA unfoldase opens the axial gate of ClpP peptidase. J. Biol. Chem. *285*, 19, 14834–14840.

Elsholz, A.K.W., Turgay, K., Michalik, S., Hessling, B., Gronau, K., Oertel, D., Mäder, U., Bernhardt, J., Becher, D., Hecker, M., et al. (2012). Global impact of protein arginine phosphorylation on the physiology of *Bacillus subtilis*. Proc. Natl. Acad. Sci. U. S. A. *109*, 19, 7451–7456.

Elsholz, A.K.W., Birk, M.S., Charpentier, E., and Turgay, K. (2017). Functional diversity of AAA+ protease complexes in *Bacillus subtilis*. Front. Mol. Biosci. *4*, 44.

Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D.A., and Bukau, B. (2006). ClpS is an essential component of the N-end rule pathway in *Escherichia coli*. Nature *439*, 7077, 753–756.

Erzberger, J.P., and Berger, J.M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu. Rev. Biophys. Biomol. Struct. *35*, 93–114.

Farrell, C.M., Baker, T.A., and Sauer, R.T. (2007). Altered specificity of a AAA+ protease. Mol. Cell *25*, 1, 161–166.

Fei, X., Bell, T.A., Jenni, S., Stinson, B.M., Baker, T.A., Harrison, S.C., and Sauer, R.T. (2020a). Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. eLife *9*, e52774.

Fei, X., Bell, T.A., Barkow, S.R., Baker, T.A., and Sauer, R.T. (2020b). Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. eLife *9*, e61496, 1–39.

Flynn, J.M., Levchenko, I., Seidel, M., Wickner, S.H., Sauer, R.T., and Baker, T.A. (2001). Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. Proc. Natl. Acad. Sci. U. S. A. *98*, 19, 10584–10589.

Flynn, J.M., Neher, S.B., Kim, Y.I., Sauer, R.T., and Baker, T.A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. Mol. Cell *11*, 3, 671–683.

Gaillot, O., Bregenholt, S., Jaubert, F., Di Santo, J.P., and Berche, P. (2001). Stress-induced ClpP serine protease of *Listeria monocytogenes* is essential for induction of listeriolysin O-dependent protective immunity. Infect. Immun. *69*, 8, 4938–4943.

Galhardo, R.S., Hastings, P.J., and Rosenberg, S.M. (2007). Mutation as a stress response and the regulation of evolvability. Crit. Rev. Biochem. Mol. Biol. *42*, 5, 399–435.

Gao, X., Yeom, J., and Groisman, E.A. (2019). The expanded specificity and physiological role of a widespread N-degron recognin. Proc. Natl. Acad. Sci. U. S. A. *116*, 37, 18629–18637.

Gates, S.N., Yokom, A.L., Lin, J., Jackrel, M.E., Rizo, A.N., Kendsersky, N.M., Buell, C.E., Sweeny, E.A., Mack, K.L., Chuang, E., et al. (2017). Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. Science. *357*, 6348, 273–279.

Gavrish, E., Sit, C.S., Cao, S., Kandror, O., Spoering, A., Peoples, A., Ling, L., Fetterman, A., Hughes, D., Bissell, A., et al. (2014). Lassomycin, a ribosomally synthesized cyclic peptide, kills *Mycobacterium tuberculosis* by targeting the ATP-dependent protease ClpC1P1P2. Chem. Biol. *21*, 4, 509–518.

Ge, Z., and Karzai, A.W. (2009). Co-evolution of multipartite interactions between an extended tmRNA tag and a robust Lon protease in Mycoplasma. Mol. Microbiol. *74*, 5, 1083–1099.

Gerth, U., Kirstein, J., Mostertz, J., Waldminghaus, T., Miethke, M., Kock, H., and Hecker, M. (2004). Fine-tuning in regulation of Clp protein content in *Bacillus subtilis*. J. Bacteriol. *186*, 1, 179–191.

Ghosh, S., Salot, S., Sengupta, S., Navalkar, A., Ghosh, D., Jacob, R., Das, S., Kumar, R., Jha, N.N., Sahay, S., et al. (2017). P53 amyloid formation leading to its loss of function: Implications in cancer pathogenesis. Cell Death Differ. *24*, 10, 1784–1798.

Glynn, S.E. (2017). Multifunctional mitochondrial AAA proteases. Front. Mol. Biosci. *4*, 34.

Gonzalez-Garcia, M., Fusco, G., and De Simone, A. (2021). Membrane interactions and toxicity by misfolded protein oligomers. Front. Cell Dev. Biol. *9*, 395.

Gonzalez, M., Rasulova, F., Maurizi, M.R., and Woodgate, R. (2000). Subunit-specific degradation of the UmuD/D' heterodimer by the ClpXP protease: The role of *trans* recognition in UmuD' stability. EMBO J. *19*, 19, 5251–5258.

Gottesman, S., Roche, E., Zhou, Y., and Sauer, R.T. (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the ssrA-tagging system. Genes Dev. *12*, 9, 1338–1347.

Graciet, E., Hu, R.G., Piatkov, K., Rhee, J.H., Schwarz, E.M., and Varshavsky, A. (2006). Aminoacyl-transferases and the N-end rule pathway of prokaryotic/eukaryotic specificity in a human pathogen. Proc. Natl. Acad. Sci. *103*, 9, 3078–3083.

Grimaud, R., Kessel, M., Beuron, F., Steven, A.C., and Maurizi, M.R. (1998). Enzymatic and structural similarities between the *Escherichia coli* ATP-dependent proteases, ClpXP and ClpAP. J. Biol. Chem. *273*, 20, 12476–12481.

Groll, M., Bajorek, M., Köhler, A., Moroder, L., Rubin, D.M., Huber, R., Glickman, M.H., and Finley, D. (2000). A gated channel into the proteasome core particle. Nat. Struct. Biol. *7*, 11, 1062–1067.

Guo, F., Esser, L., Singh, S.K., Maurizi, M.R., and Xia, D. (2002). Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA. J. Biol. Chem. *277*, 48, 46753–46762.

Gur, E. (2013). The Lon AAA+ protease. Subcell. Biochem. *66*, 35–51.

Gur, E., and Sauer, R.T. (2008a). Recognition of misfolded proteins by Lon, a AAA+ protease. Genes Dev. *22*, 16, 2267–2277.

Gur, E., and Sauer, R.T. (2008b). Evolution of the ssrA degradation tag in Mycoplasma: Specificity switch to a different protease. Proc. Natl. Acad. Sci. U. S. A. *105*, 42, 16113–16118.

Gur, E., Ottofueling, R., and Dougan, D.A. (2013). Machines of destruction – AAA+ proteases and the adaptors that control them. Subcell. Biochem. *66*, 3–33.

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

Han, H., Monroe, N., Sundquist, W.I., Shen, P.S., and Hill, C.P. (2017). The AAA ATPase Vps4 binds ESCRT-III substrates through a repeating array of dipeptide-binding pockets. eLife *6*, e31324.

Han, H., Schubert, H.L., McCullough, J., Monroe, N., Purdy, M.D., Yeager, M., Sundquist, W.I., and Hill, C.P. (2020). Structure of spastin bound to a glutamate-rich peptide implies a hand-over-hand mechanism of substrate translocation. J. Biol. Chem. *295*, 2, 435–443.

Händel, N., Hoeksema, M., Freijo Mata, M., Brul, S., and ter Kuile, B.H. (2015). Effects of stress, reactive oxygen species, and the SOS response on *de novo* acquisition of antibiotic resistance in *Escherichia coli*. Antimicrob. Agents Chemother. *60*, 3, 1319–1327.

Harms, A., Maisonneuve, E., and Gerdes, K. (2016). Mechanisms of bacterial persistence during stress and antibiotic exposure. Science *354*, 6318.

Hartl, F.U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding *in vitro* and *in vivo*. Nat. Struct. Mol. Biol. *16*, 6, 574–581.

Hayes, C.S., Bose, B., and Sauer, R.T. (2002). Proline residues at the C terminus of nascent chains induce ssrA tagging during translation termination. J. Biol. Chem. *277*, 37, 33825–33832.

Herman, C., Thévenet, D., Bouloc, P., Walker, G.C., and D'Ari, R. (1998). Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH). Genes Dev. *12*, 9, 1348.

Hersch, G.L., Baker, T.A., and Sauer, R.T. (2004). SspB delivery of substrates for ClpXP proteolysis probed by the design of improved degradation tags. Proc. Natl. Acad. Sci. *101*, 33, 12136–12141.

Hong, Z., Bolard, A., Giraud, C., Prévost, S., Genta-Jouve, G., Deregnaucourt, C., Häussler, S., Jeannot, K., and Li, Y. (2019). Azetidine-containing alkaloids produced by a quorum-sensing regulated nonribosomal peptide synthetase pathway in *Pseudomonas aeruginosa*. Angew. Chemie *131*, 10, 3210–3214.

Hoskins, J.R., Yanagihara, K., Mizuuchi, K., and Wickner, S. (2002). ClpAP and ClpXP degrade proteins with tags located in the interior of the primary sequence. Proc. Natl. Acad. Sci. U. S. A. *99*, 17, 11037–11042.

Hou, J.Y., Sauer, R.T., and Baker, T.A. (2008). Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. Nat. Struct. Mol. Biol. *15*, 3, 288–294.

Humbard, M.A., Surkov, S., De Donatis, G.M., Jenkins, L.M., and Maurizi, M.R. (2013). The N-degradome of *Escherichia coli*: limited proteolysis *in vivo* generates a large pool of proteins bearing N-degrons. J. Biol. Chem. *288*, 40, 28913–28924.

Inobe, T., Fishbain, S., Prakash, S., and Matouschek, A. (2011). Defining the geometry of the two-component proteasome degron. Nat. Chem. Biol. *7*, 3, 161–167.

Iyer, L.M., Leipe, D.D., Koonin, E. V., and Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. J. Struct. Biol. *146*, 1–2, 11–31.

Janssen, B.D., and Hayes, C.S. (2012). The tmRNA ribosome-rescue system. Advances in Protein Chemistry and Structural Biology. *86*,151-91

Jayaweera, S.P.E., Wanigasinghe Kanakanamge, S.P., Rajalingam, D., and Silva, G.N. (2021). Carfilzomib: a promising proteasome inhibitor for the treatment of relapsed and refractory multiple myeloma. Front. Oncol. *11*, 740796.

Jenal, U., and Fuchs, T. (1998). An essential protease involved in bacterial cell-cycle control. EMBO J. *17*, 19, 5658–5669.

Johjima, A., Noi, K., Nishikori, S., Ogi, H., Esaki, M., and Ogura, T. (2015). Microtubule severing by katanin p60 AAA+ATPase requires the C-terminal acidic tails of both α-and β-tubulins and basic amino acid residues in the AAA+ring pore. J. Biol. Chem. *290*, 18, 11762–11770.

Joshi, K.K., Bergé, M., Radhakrishnan, S.K., Viollier, P.H., and Chien, P. (2015). An adaptor hierarchy regulates proteolysis during a bacterial cell cycle. Cell *163*, 2, 419–431.

Karzai, A.W., Roche, E.D., and Sauer, R.T. (2000). The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. Nat. Struct. Biol. *7*, 6, 449–455.

Kavalchuk, M., Jomaa, A., Müller, A.U., and Weber-Ban, E. (2022). Structural basis of prokaryotic ubiquitin-like protein engagement and translocation by the mycobacterial Mpa-proteasome complex. Nat. Commun. *13*, 1, 1–13.

Keiler, K.C., Waller, P.R.H., and Sauer, R.T. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Science. *271*, 5251, 990–993.

Kessel, M., Maurizi, M.R., Kim, B., Kocsis, E., Trus, B.L., Singh, S.K., and Steven, A.C. (1995). Homology in structural organization between *E. coli* ClpAP protease and the eukaryotic 26S proteasome. J. Mol. Biol. *250*, 5, 587–594.

Kim, Y.I., Levchenko, I., Fraczkowska, K., Woodruff, R. V., Sauer, R.T., and Baker, T.A. (2001). Molecular determinants of complex formation between Clp/Hsp 100 ATPases and the ClpP peptidase. Nat. Struct. Biol. *8*, 3, 230–233.

King, R.W., Deshaies, R.J., Peters, J.M., and Kirschner, M.W. (1996). How proteolysis drives the cell cycle. Science. *274*, 5293, 1652–1659.

Kirstein, J., Dougan, D.A., Gerth, U., Hecker, M., and Turgay, K. (2007). The tyrosine kinase McsB is a regulated adaptor protein for ClpCP. EMBO J. *26*, 8, 2061–2070.

Kirstein, J., Hoffmann, A., Lilie, H., Schmidt, R., Helga, R.W., Heike, B.O., Mogk, A., and Turgay, K. (2009). The antibiotic ADEP reprogrammes CIP, switching it from a regulated to an uncontrolled protease. EMBO Mol. Med. *1*, 1, 37–49.

Kruklitis, R., Welty, D.J., and Nakai, H. (1996). ClpX protein of *Escherichia coli* activates bacteriophage Mu transposase in the strand transfer complex for initiation of Mu DNA synthesis. EMBO J. *15*, 4, 935–944.

Kwon, A.R., Kessler, B.M., Overkleeft, H.S., and McKay, D.B. (2003). Structure and reactivity of an asymmetric complex between HslV and I-domain deleted HslU, a prokaryotic homolog of the eukaryotic proteasome. J. Mol. Biol. *330*, 2, 185–195.

de la Peña, A.H., Goodall, E.A., Gates, S.N., Lander, G.C., and Martin, A. (2018). Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-driven translocation. Science. *362*, 6418.

Laachouch, J.E., Desmet, L., Geuskens, V., Grimaud, R., and Toussaint, A. (1996). Bacteriophage Mu repressor as a target for the *Escherichia coli* ATP-dependent Clp protease. EMBO J. *15*, 2, 437–444.

Lee, B.G., Park, E.Y., Lee, K.E., Jeon, H., Sung, K.H., Paulsen, H., Rübsamen-Schaeff, H., Brötz-Oesterhelt, H., and Song, H.K. (2010a). Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. Nat. Struct. Mol. Biol. *17*, 4, 471–478.

Lee, M.E., Baker, T.A., and Sauer, R.T. (2010b). Control of substrate gating and translocation into ClpP by channel residues and ClpX binding. J. Mol. Biol. *399*, 5, 707–718.

Levchenko, I., Luo, L., and Baker, T.A. (1995). Disassembly of the Mu transposase tetramer by the ClpX chaperone. Genes Dev. *9*, 19, 2399–2408.

Levchenko, I., Seidel, M., Sauer, R.T., and Baker, T.A. (2000). A specificity-enhancing factor for the ClpXP degradation machine. Science. *289*, 5488, 2354–2356.

Levchenko, I., Grant, R.A., Wah, D.A., Sauer, R.T., and Baker, T.A. (2003). Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag. Mol. Cell *12*, 2, 365–372.

Levchenko, I., Grant, R.A., Flynn, J.M., Sauer, R.T., and Baker, T.A. (2005). Versatile modes of peptide recognition by the AAA+ adaptor protein SspB. Nat. Struct. Mol. Biol. *12*, 6, 520–525.

Li, D.H.S., Chung, Y.S., Gloyd, M., Joseph, E., Ghirlando, R., Wright, G.D., Cheng, Y.Q., Maurizi, M.R., Guarné, A., and Ortega, J. (2010). Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: A model for the ClpX/ClpA-bound state of ClpP. Chem. Biol. *17*, 9, 959–969.

Lies, M., and Maurizi, M.R. (2008). Turnover of endogenous ssrA-tagged proteins mediated by ATP-dependent proteases in *Escherichia coli*. J. Biol. Chem. *283*, 34, 22918–22929.

Ling, L., Montaño, S.P., Sauer, R.T., Rice, P.A., and Baker, T.A. (2015). Deciphering the roles of multicomponent recognition signals by the AAA+ unfoldase ClpX. J. Mol. Biol. *427*, 18, 2966–2982.

Liu, J., Francis, L.I., Jonas, K., Laub, M.T., and Chien, P. (2016). ClpAP is an auxiliary protease for DnaA degradation in *Caulobacter crescentus*. Mol. Microbiol. *102*, 6, 1075–1085.

Livneh, I., Cohen-Kaplan, V., Cohen-Rosenzweig, C., Avni, N., and Ciechanover, A. (2016). The life cycle of the 26S proteasome: From birth, through regulation and function, and onto its death. Cell Res. *26*, 8, 869–885.

Lopez, K.E., Rizo, A.N., Tse, E., Lin, J.B., Scull, N.W., Thwin, A.C., Lucius, A.L., Shorter, J., and Southworth, D.R. (2020). Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis. Nat. Struct. Mol. Biol. *27*, 5, 406–416.

Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995). Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. Science. *268*, 5210, 533–539.

Lundqvist, J., Elmlund, H., Wulff, R.P., Berglund, L., Elmlund, D., Emanuelsson, C., Hebert, H., Willows, R.D., Hansson, M., Lindahl, M., et al. (2010). ATP-induced conformational dynamics in the AAA+ motor unit of magnesium chelatase. Structure *18*, 3, 354–365.

Lupas, A.N., and Koretke, K.K. (2003). Bioinformatic analysis of ClpS, a protein module involved in prokaryotic and eukaryotic protein degradation. J. Struct. Biol. *141*, 1, 77–83.

Maglica, Ž., Striebel, F., and Weber-Ban, E. (2008). An intrinsic degradation tag on the ClpA C-terminus regulates the balance of ClpAP complexes with different substrate specificity. J. Mol. Biol. *384*, 2, 503–511.

Mahmoud, S.A., and Chien, P. (2018). Regulated proteolysis in bacteria. Annu. Rev. Biochem. *87*, 677–696.

Martin, A., Baker, T.A., and Sauer, R.T. (2008). Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. Mol. Cell *29*, 4, 441–450.

Maupin-Furlow, Julie, A. (2000). Proteasomes in the archaea: from structure to function. Front. Biosci. *5*, 1, d837.

Maupin-Furlow, J. (2011). Proteasomes and protein conjugation across domains of life. Nat. Rev. Microbiol. 2012 102 *10*, 2, 100–111.

Mawla, G.D., Hall, B.M., Cárcamo-Oyarce, G., Grant, R.A., Zhang, J.J., Kardon, J.R., Ribbeck, K., Sauer, R.T., and Baker, T.A. (2021). ClpP1P2 peptidase activity promotes biofilm formation in *Pseudomonas aeruginosa*. Mol. Microbiol. *115*, 6, 1094–1109.

Mogk, A., Huber, D., and Bukau, B. (2011). Integrating protein homeostasis strategies in prokaryotes. Cold Spring Harb. Perspect. Biol. *3*, 4, 1–19.

Morozov, A. V., and Karpov, V.L. (2019). Proteasomes and several aspects of their heterogeneity relevant to cancer. Front. Oncol. *9*, AUG, 761.

Mulcahy Levy, J.M., and Thorburn, A. (2020). Autophagy in cancer: moving from understanding mechanism to improving therapy responses in patients. Cell Death Differ. *27*, 3, 843–857.

Murata, S., Yashiroda, H., and Tanaka, K. (2009). Molecular mechanisms of proteasome assembly. Nat. Rev. Mol. Cell Biol. *10*, 2, 104–115.

Murata, S., Takahama, Y., Kasahara, M., and Tanaka, K. (2018). The immunoproteasome and thymoproteasome: functions, evolution and human disease. Nat. Immunol. *19*, 9, 923–931.

Narberhaus, F., Urech, C., and Hennecke, H. (1999). Characterization of the *Bradyrhizobium japonicum ftsH* gene and its product. J. Bacteriol. *181*, 23, 7394–7397.

Navalkar, A., Pandey, S., Singh, N., Patel, K., Datta, D., Mohanty, B., Jadhav, S., Chaudhari, P., and Maji, S.K. (2021). Direct evidence of cellular transformation by prion-like p53 amyloid infection. J. Cell Sci. *134*, 11.

Neher, S.B., Flynn, J.M., Sauer, R.T., and Baker, T.A. (2003). Latent ClpX-recognition signals ensure LexA destruction after DNA damage. Genes Dev. *17*, 9, 1084–1089.

Ninnis, R.L., Spall, S.K., Talbo, G.H., Truscott, K.N., and Dougan, D.A. (2009). Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli*. EMBO J. *28*, 12, 1732–1744.

Nishimura, K., and Van Wijk, K.J. (2015). Organization, function and substrates of the essential Clp protease system in plastids. Biochim. Biophys. Acta - Bioenerg. *1847*, 9, 915–930.

Okuno, T., Yamanaka, K., and Ogura, T. (2006). An AAA protease FtsH can initiate proteolysis from internal sites of a model substrate, apo-flavodoxin. Genes to Cells *11*, 3, 261–268.

Olivares, A.O., Baker, T.A., and Sauer, R.T. (2015). Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. Nat. Rev. Microbiol. *14*, 1, 33–44.

de Oliveira, G.A.P., Petronilho, E.C., Pedrote, M.M., Marques, M.A., Vieira, T.C.R.G., Cino, E.A., and Silva, J.L. (2020). The status of p53 oligomeric and aggregation states in cancer. Biomolecules *10*, 4.

Patteson, J.B., Lescallette, A.R., and Li, B. (2019). Discovery and biosynthesis of azabicyclene, a conserved nonribosomal peptide in *Pseudomonas aeruginosa*. Org. Lett. *21*, 13, 4955–4959.

Piatkov, K.I., Vu, T.T.M., Hwang, C.S., and Varshavsky, A. (2015). Formyl-methionine as a degradation signal at the N-termini of bacterial proteins. Microb. Cell *2*, 10, 376–393.

Poole, K. (2012). Stress responses as determinants of antimicrobial resistance in Gram-negative bacteria. Trends Microbiol. *20*, 5, 227–234.

Pressler, K., Vorkapic, D., Lichtenegger, S., Malli, G., Barilich, B.P., Cakar, F., Zingl, F.G., Reidl, J., and Schild, S. (2016). AAA+ proteases and their role in distinct stages along the *Vibrio cholerae* lifecycle. Int. J. Med. Microbiol. *306*, 6, 452–462.

Pruteanu, M., and Baker, T.A. (2009). Proteolysis in the SOS response and metal homeostasis in *Escherichia coli*. Res. Microbiol. *160*, 9, 677–683.

Puchades, C., Rampello, A.J., Shin, M., Giuliano, C.J., Wiseman, R.L., Glynn, S.E., and Lander, G.C. (2017). Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. Science. *358*, 6363.

Puchades, C., Ding, B., Song, A., Wiseman, R.L., Lander, G.C., and Glynn, S.E. (2019). Unique structural features of the mitochondrial AAA+ protease AFG3L2 reveal the molecular basis for activity in health and disease. Mol. Cell *75*, 5, 1073-1085.e6.

Puchades, C., Sandate, C.R., and Lander, G.C. (2020). The molecular principles governing the activity and functional diversity of AAA+ proteins. Nat. Rev. Mol. Cell Biol. *21*, 1, 43–58.

Qiu, D., Eisinger, V.M., Head, N.E., Pier, G.B., and Yu, H.D. (2008). ClpXP proteases positively regulate alginate overexpression and mucoid conversion in *Pseudomonas aeruginosa*. Microbiology *154*, 7, 2119–2130.

Rabl, J., Smith, D.M., Yu, Y., Chang, S.C., Goldberg, A.L., and Cheng, Y. (2008). Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases. Mol. Cell *30*, 3, 360.

Raju, R.M., Unnikrishnan, M., Rubin, D.H.F., Krishnamoorthy, V., Kandror, O., Akopian, T.N., Goldberg, A.L., and Rubin, E.J. (2012). *Mycobacterium tuberculosis* ClpP1 and ClpP2 function together in protein degradation and are required for viability in vitro and during infection. PLoS Pathog. *8*, 2, e1002511.

Ramachandran, R., Hartmann, C., Song, H.K., Huber, R., and Bochtler, M. (2002). Functional interactions of HslV (ClpQ) with the ATPase HslU (ClpY). Proc. Natl. Acad. Sci. *99*, 11, 7396–7401.

Rao, R. V., and Bredesen, D.E. (2004). Misfolded proteins, endoplasmic reticulum stress and neurodegeneration. Curr. Opin. Cell Biol. *16*, 6, 653.

Rivera-Rivera, I. (2015). Mechanism of active substrate delivery by the AAA+ protease adaptor ClpS. Ph.D. thesis, Massachusetts Institute of Technology. Retrieved from https://dspace.mit.edu/handle/1721.1/101352

Rivera-Rivera, I., Román-Hernández, G., Sauer, R.T., and Baker, T.A. (2014). Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. Proc. Natl. Acad. Sci. U. S. A. *111*, 37, E3853-9.

Roche, E.D., and Sauer, R.T. (1999). SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. EMBO J. *18*, 16, 4579–4589.

Roche, E.D., and Sauer, R.T. (2001). Identification of endogenous ssrA-tagged proteins reveals tagging at positions corresponding to stop codons. J. Biol. Chem. *276*, 30, 28509–28515.

Román-Hernández, G., Grant, R.A., Sauer, R.T., and Baker, T.A. (2009). Molecular basis of substrate selection by the N-end rule adaptor protein ClpS. Proc. Natl. Acad. Sci. U. S. A. *106*, 22, 8888–8893.

Román-Hernández, G., Hou, J.Y., Grant, R.A., Sauer, R.T., and Baker, T.A. (2011). The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. Mol. Cell *43*, 2, 217–228.

Rousseau, A., and Bertolotti, A. (2018). Regulation of proteasome assembly and activity in health and disease. Nat. Rev. Mol. Cell Biol. *19*, 11, 697–712.

Sandate, C.R., Szyk, A., Zehr, E.A., Lander, G.C., and Roll-Mecak, A. (2019). An allosteric network in spastin couples multiple activities required for microtubule severing. Nat. Struct. Mol. Biol. *26*, 8, 671–678.

Sauer, R.T., and Baker, T.A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. Annu. Rev. Biochem. *80*, 1, 587–612.

Schmitt, E.K., Riwanto, M., Sambandamurthy, V., Roggo, S., Miault, C., Zwingelstein, C., Krastel, P., Noble, C., Beer, D., Rao, S.P.S., et al. (2011). The natural product cyclomarin kills *Mycobacterium tuberculosis* by targeting the ClpC1 subunit of the caseinolytic protease. Angew. Chemie - Int. Ed. *50*, 26, 5889–5891.

Schmitz, K.R., Carney, D.W., Sello, J.K., and Sauer, R.T. (2014). Crystal structure of *Mycobacterium tuberculosis* ClpP1P2 suggests a model for peptidase activation by AAA+ partner binding and substrate delivery. Proc. Natl. Acad. Sci. *111*, 43, E4587–E4595.

Seong, I.S., Kang, M.S., Choi, M.K., Lee, J.W., Koh, O.J., Wang, J., Eom, S.H., and Chung, C.H. (2002). The C-terminal tails of HslU ATPase act as a molecular switch for activation of HslV peptidase. J. Biol. Chem. *277*, 29, 25976–25982.

Shin, M., Watson, E.R., Song, A.S., Mindrebo, J.T., Novick, S.J., Griffin, P.R., Wiseman, R.L., and Lander, G.C. (2021). Structures of the human LONP1 protease reveal regulatory steps involved in protease activation. Nat. Commun. *12*, 1, 3239.

Shrader, T.E., Tobias, J.W., and Varshavsky, A. (1993). The N-end rule in *Escherichia coli*: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene aat. J. Bacteriol. *175*, 14, 4364.

Snider, J., Thibault, G., and Houry, W.A. (2008). The AAA+ superfamily of functionally diverse proteins. Genome Biol. *9*, 4, 1–8.

Song, H.K., and Eck, M.J. (2003). Structural basis of degradation signal recognition by SspB, a specificity-enhancing factor for the ClpXP proteolytic machine. Mol. Cell *12*, 1, 75–86.

Sousa, M.C., Trame, C.B., Tsuruta, H., Wilbanks, S.M., Reddy, V.S., and McKay, D.B. (2000). Crystal and solution structures of an HslUV protease-chaperone complex. Cell *103*, 4, 633–643.

Sousa, M.C., Kessler, B.M., Overkleeft, H.S., and McKay, D.B. (2002). Crystal structure of HslUV complexed with a vinyl sulfone inhibitor: Corroboration of a proposed mechanism of allosteric activation of HslV by HslU. J. Mol. Biol. *318*, 3, 779–785.

Stefani, M. (2004). Protein misfolding and aggregation: New examples in medicine and biology of the dark side of the protein world. Biochim. Biophys. Acta - Mol. Basis Dis. *1739*, 1, 5–25.

Stein, B.J., Grant, R.A., Sauer, R.T., and Baker, T.A. (2016). Structural basis of an N-Degron adaptor with more stringent specificity. Structure *24*, 2, 232–242.

Striebel, F., Hunkeler, M., Summer, H., and Weber-Ban, E. (2010). The mycobacterial Mpa-proteasome unfolds and degrades pupylated substrates by engaging Pup's N-terminus. EMBO J. *29*, 7, 1262–1271.

Striebel, F., Imkamp, F., Özcelik, D., and Weber-Ban, E. (2014). Pupylation as a signal for proteasomal degradation in bacteria. Biochim. Biophys. Acta - Mol. Cell Res. *1843*, 1, 103–113.

Suno, R., Niwa, H., Tsuchiya, D., Zhang, X., Yoshida, M., and Morikawa, K. (2006). Structure of the whole cytosolic region of ATP-dependent protease FtsH. Mol. Cell *22*, 5, 575–585.

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

Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991). The N-end rule in bacteria. Science. *254*, 5036, 1374–1377.

Tomita, T., and Matouschek, A. (2019). Substrate selection by the proteasome through initiation regions. Protein Sci. *28*, 7, 1222–1232.

Tomoyasu, T., Yuki, T., Morimura, S., Mori, H., Yamanaka, K., Niki, H., Hiraga, S., and Ogura, T. (1993). The *Escherichia coli* FtsH protein is a prokaryotic member of a protein family of putative ATPases involved in membrane functions, cell cycle control, and gene expression. J. Bacteriol. *175*, 5, 1344–1351.

Torres-Delgado, A., Kotamarthi, H.C., Sauer, R.T., and Baker, T.A. (2020). The intrinsically disordered N-terminal extension of the ClpS adaptor reprograms its partner AAA + ClpAP protease. J. Mol. Biol. *432*, 17, 4908–4921.

Trentini, D.B., Suskiewicz, M.J., Heuck, A., Kurzbauer, R., Deszcz, L., Mechtler, K., and Clausen, T. (2016). Arginine phosphorylation marks proteins for degradation by a Clp protease. Nature *539*, 7627, 48.

Tryggvesson, A., Ståhlberg, F.M., Töpel, M., Tanabe, N., Mogk, A., and Clarke, A.K. (2015). Characterization of ClpS2, an essential adaptor protein for the cyanobacterium *Synechococcus elongatus*. FEBS Lett. *589*, 24, 4039–4046.

Valastyan, J.S., and Lindquist, S. (2014). Mechanisms of protein-folding diseases at a glance. Dis. Model. Mech. *7*, 1, 9.

Varshavsky, A. (1991). Naming a targeting signal. Cell *64*, 1, 13–15.

Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. Protein Sci. *20*, 8, 1298–1345.

Varshavsky, A. (2017). The ubiquitin system, autophagy, and regulated protein degradation. Annu. Rev. Biochem. *86*, 123–128.

Varshavsky, A. (2019). N-degron and C-degron pathways of protein degradation. Proc. Natl. Acad. Sci. U. S. A. *116*, 2, 358–366.

Vasudevan, D., Rao, S.P.S., and Noble, C.G. (2013). Structural basis of mycobacterial inhibition by cyclomarin A. J. Biol. Chem. *288*, 43, 30883–30891.

Wah, D.A., Levchenko, I., Rieckhof, G.E., Bolon, D.N., Baker, T.A., and Sauer, R.T. (2003). Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the AAA+ ClpXP protease. Mol. Cell *12*, 2, 355–363.

Wang, M., and Kaufman, R.J. (2014). The impact of the endoplasmic reticulum protein-folding environment on cancer development. Nat. Rev. Cancer *14*, 9, 581–597.

Wang, J., Hartling, J.A., and Flanagan, J.M. (1997). The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis. Cell *91*, 4, 447–456.

Wang, K.H., Oakes, E.S.C., Sauer, R.T., and Baker, T.A. (2008a). Tuning the strength of a bacterial N-end rule degradation signal. J. Biol. Chem. *283*, 36, 24600–24607.

Wang, K.H., Roman-Hernandez, G., Grant, R.A., Sauer, R.T., and Baker, T.A. (2008b). The molecular basis of N-end rule recognition. Mol. Cell *32*, 3, 406–414.

Wang, T., Darwin, K.H., and Li, H. (2010). Binding-induced folding of prokaryotic ubiquitin-like protein on the *Mycobacterium* proteasomal ATPase targets substrates for degradation. Nat. Struct. Mol. Biol. 2010 1711 *17*, 11, 1352–1357.

Wang, Y., Cao, W., Merritt, J., Xie, Z., and Liu, H. (2021). Characterization of FtsH essentiality in *Streptococcus mutans* via genetic suppression. Front. Genet. *12*, 659220.

Wendler, P., Ciniawsky, S., Kock, M., and Kube, S. (2012). Structure and function of the AAA+ nucleotide binding pocket. Biochim. Biophys. Acta - Mol. Cell Res. *1823*, 1, 2–14.

Whitby, F.G., Masters, E.I., Kramer, L., Knowlton, J.R., Yao, Y., Wang, C.C., and Hill, C.P. (2000). Structural basis for the activation of 20S proteasomes by 11S regulators. Nature *408*, 6808, 115–120.

Witt, S., Kwon, Y. Do, Sharon, M., Felderer, K., Beuttler, M., Robinson, C. V, Baumeister, W., and Jap, B.K. (2006). Proteasome assembly triggers a switch required for active-site maturation. Structure *14*, 7, 1179–1188.

Wu, W.F., Zhou, Y., and Gottesman, S. (1999). Redundant *in vivo* proteolytic activities of *Escherichia coli* Lon and the ClpYQ (HslUV) Protease. J. Bacteriol. *181*, 12, 3681.

Yang-Hartwich, Y., Soteras, M.G., Lin, Z.P., Holmberg, J., Sumi, N., Craveiro, V., Liang, M., Romanoff, E., Bingham, J., Garofalo, F., et al. (2015). p53 protein aggregation promotes platinum resistance in ovarian cancer. Oncogene *34*, 27, 3605–3616.

Yang, Y., and Bedford, M.T. (2012). Titivated for destruction:the methyl degron. Mol. Cell *48*, 4, 487–488.

Yoo, S.J., Seol, J.H., Shin, D.H., Rohrwild, M., Hang, M.S., Tanaka, K., Goldberg, A.L., and Chung, C.H. (1996). Purification and characterization of the heat shock proteins HslV and HslU that form a new ATP-dependent protease in *Escherichia coli*. J. Biol. Chem. *271*, 24, 14035–14040.

Yu, A.Y.H., and Houry, W.A. (2007). ClpP: A distinctive family of cylindrical energy-dependent serine proteases. FEBS Lett. *581*, 19, 3749–3757.

Zehr, E., Szyk, A., Piszczek, G., Szczesna, E., Zuo, X., and Roll-Mecak, A. (2017). Katanin spiral and ring structures shed light on power stroke for microtubule severing. Nat. Struct. Mol. Biol. *24*, 9, 717–725.

Zehr, E.A., Szyk, A., Szczesna, E., and Roll-Mecak, A. (2020). Katanin grips the β-tubulin tail through an electropositive double spiral to sever microtubules. Dev. Cell *52*, 1, 118-131.e6.

Zerbib, E., Schlussel, S., Hecht, N., Bagdadi, N., Eichler, J., and Gur, E. (2021). The prokaryotic ubiquitin-like protein presents poor cleavage sites for proteasomal degradation. Cell Rep. *36*, 4, 109428.

Zeth, K., Ravelli, R.B., Paal, K., Cusack, S., Bukau, B., and Dougan, D.A. (2002). Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. Nat. Struct. Biol. *9*, 12, 906–911.

Zhongming, G.E., and Taylor, D.E. (1996). Sequencing, expression, and genetic characterization of the *Helicobacter pylori ftsH* gene encoding a protein homologous to members of a novel putative ATPase family. J. Bacteriol. *178*, 21, 6151–6157.

Ziemski, M., Leodolter, J., Taylor, G., Kerschenmeyer, A., and Weber-Ban, E. (2021). Genome-wide interaction screen for *Mycobacterium tuberculosis* ClpCP protease reveals toxin–antitoxin systems as a major substrate class. FEBS J. *288*, 1, 111–126.

Zuromski, K.L., Kim, S., Sauer, R.T., and Baker, T.A. (2021). Division of labor between the pore-1 loops of the D1 and D2 AAA+ rings coordinates substrate selectivity of the ClpAP protease. J. Biol. Chem. *297*, 6, 101407.

CHAPTER 2

ClpAP proteolysis does not require rotation of the ClpA unfoldase relative to ClpP

This chapter has been previously published as:
Kim S., Zuromski K.L., Bell T.A., Sauer R.T., and Baker T.A. ClpAP proteolysis does not require rotation of the ClpA unfoldase relative to ClpP. (2020). eLife *9*, e61451. doi: 10.7554/eLife.61451.

Author contributions:
I prepared the crosslinked ClpA-ClpP complexes and performed λ cI^{N}-ssrA degradation and ^{cp7}GFP-ssrA degradation experiments. K.L. Zuromski performed gel quantification of crosslinking efficiency, ATPase, FITC-casein degradation, and ^{5IAF}V13P titin^{I27}-ssrA degradation experiments.

S. Kim, K.L. Zuromski: Conceptualization, Resources, Formal analysis, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing – original draft, Project administration, Writing – review and editing

T.A. Bell: Conceptualization, Resources, Formal analysis, Validation, Investigation, Methodology, Writing – original draft, Writing – review and editing

R.T. Sauer: Conceptualization, Resources, Formal analysis, Supervision, Funding acquisition, Validation, Methodology, Writing – original draft, Project administration, Writing – review and editing

T.A. Baker: Conceptualization, Resources, Supervision, Funding acquisition, Validation, Methodology, Writing – original draft, Project administration, Writing – review and editing

ABSTRACT

AAA+ proteases perform regulated protein degradation in all kingdoms of life and consist of a hexameric AAA+ unfoldase/translocase in complex with a self-compartmentalized peptidase. Based on asymmetric features of cryo-EM structures and a sequential hand-over-hand model of substrate translocation, recent publications have proposed that the AAA+ unfoldases ClpA and ClpX rotate with respect to their partner peptidase ClpP to allow function. Here, we test this model by covalently crosslinking ClpA to ClpP to prevent rotation. We find that crosslinked ClpAP complexes unfold, translocate, and degrade protein substrates *in vitro*, albeit modestly slower than uncrosslinked enzyme controls. Rotation of ClpA with respect to ClpP is therefore not required for ClpAP protease activity, although some flexibility in how the AAA+ ring docks with ClpP may be necessary for optimal function.

INTRODUCTION

The AAA+ (ATPases Associated with diverse cellular Activities) protease subfamily uses the energy of ATP hydrolysis to disassemble and degrade proteins that are misfolded, deleterious, or unneeded (Sauer and Baker, 2011). AAA+ proteases are composed of a hexameric single- or double-ringed AAA+ unfoldase/translocase and a self-compartmentalized partner peptidase. The AAA+ rings form a shallow helix and stack with planar peptidase rings (Puchades *et al*., 2020). After protein substrate recognition by the unfoldase, repeated cycles of ATP hydrolysis power conformational changes in the AAA+ motor, promoting substrate unfolding and processive translocation of the resulting polypeptide into the proteolytic chamber of the peptidase for degradation. Recent structural and biochemical studies have illuminated some aspects of this process, but the molecular nature of the stepwise cycles these proteolytic machines use to carry out mechanical unfolding and translocation of protein substrates is still being actively explored (Puchades *et al*., 2020).

The ClpAP protease consists of the $ClpA_6$ AAA+ unfoldase, a double-ring AAA+ enzyme with two AAA+ modules per subunit, and the tetradecameric $ClpP_{14}$ peptidase, which contains two heptameric rings (Sauer and Baker, 2011; **Figure 2.1A**). Thus, the interface between ClpA and ClpP involves an asymmetric six-to-seven subunit mismatch. The ClpXP protease, composed of the single-ring AAA+ $ClpX_6$ unfoldase and the $ClpP_{14}$ peptidase, also has a six-seven mismatch, as do proteasomal AAA+ enzymes. How such mismatches are accommodated structurally and whether the mismatches play important roles in the mechanisms of these ATP-dependent proteases has long been a subject of interest. Recent near-atomic-resolution cryo-EM structures of ClpAP and ClpXP reveal that each unfoldase has six flexible peptidase-binding loops protruding from the bottom face of the AAA+ ring that can interact with $ClpP_{14}$ (Fei *et al.,* 2020a; 2020b; Ripstein *et al*., 2020; Lopez *et al.,* 2020). Part of each loop containing a conserved tripeptide motif (IGL in ClpA; IGF in ClpX) docks into hydrophobic clefts on the top of the $ClpP_7$ ring, engaging a total of five or six of the seven clefts and leaving one or two clefts unoccupied (**Figure 2.1A**) (Fei *et al.,* 2020a; 2020b; Ripstein *et al*., 2020; Lopez *et al.,* 2020).

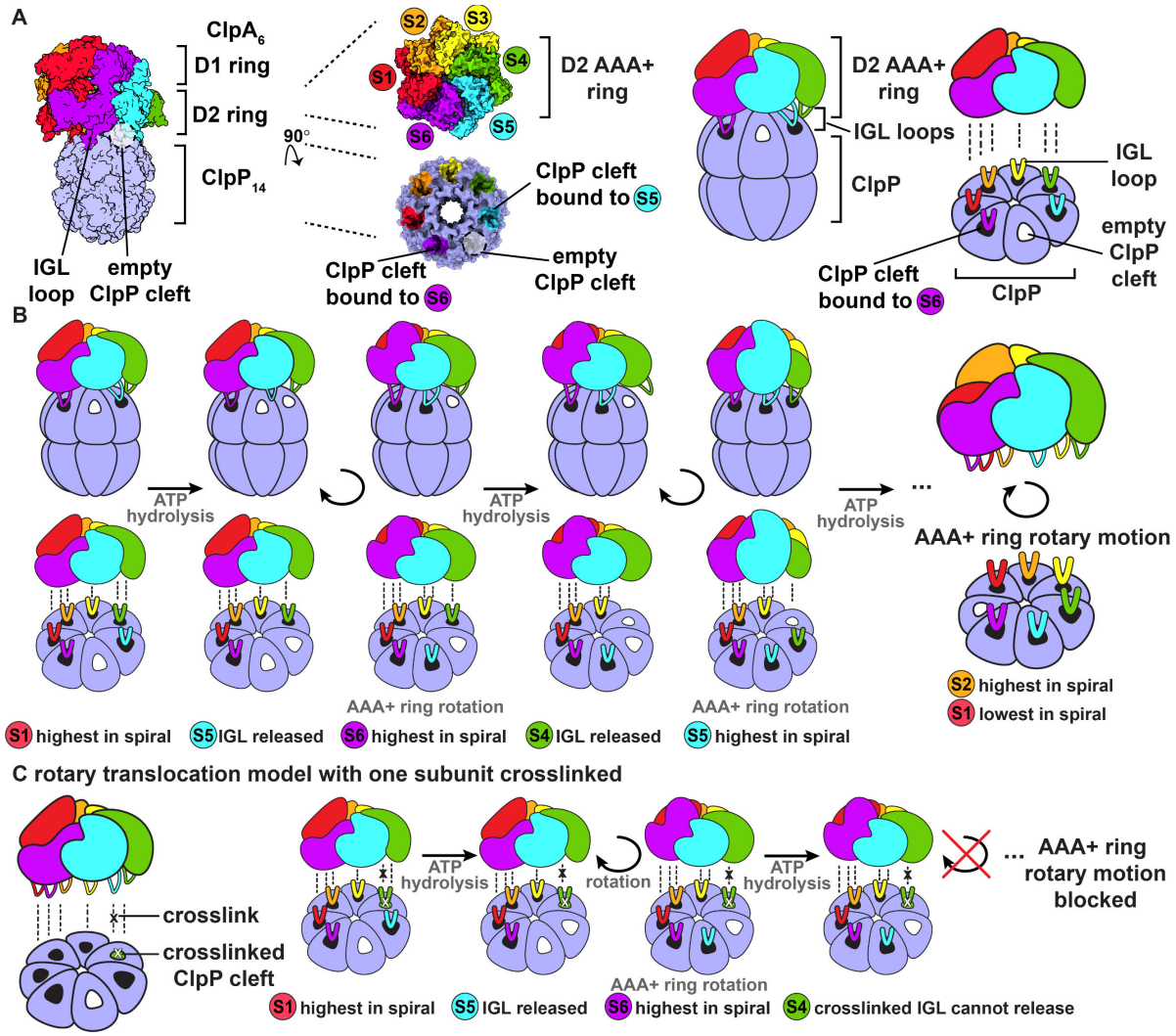


Figure 2.1 ClpAP structure and rotary translocation model. (**A**) Complex of ClpP with ClpA (PDB 6UQO). Subunits of ClpA, labeled S1 through S6, are ordered from the highest to the lowest position in the spiral relative to ClpP at the beginning of the mechanical cycle. The IGL loops of ClpA hexamers dock into a subset of the seven clefts in a heptameric ClpP ring. In this structure, there is an empty cleft between the second lowest and lowest subunits in the spiral (S5 and S6, respectively). The coloring of the ClpP clefts represents the docked position of the IGL loops from the corresponding AAA+ subunits; empty clefts are colored white. The rightmost panel is a generalized model of the ClpA D2 AAA+ ring docking into the ClpP interface. (**B**) Rotary translocation model with clockwise around-the-ring ATP hydrolysis and IGL loop release and rebinding (Ripstein *et al.*, 2020; Lopez *et al.,* 2020). When subunit S1 is highest in the spiral, ATP hydrolysis releases the IGL loop of subunit S5 and the AAA+ ring rotates clockwise with respect to ClpP. During rotation, subunit S6 moves to the top in of the spiral, and the IGL loop of subunit S5 takes a clockwise "step" and rebinds to the adjacent empty ClpP cleft. Repetition of this sequence of ATP hydrolysis and IGL loop release and rebinding results in rotary motion of the AAA+ ring with respect to the ClpP ring. (**C**) Rotary translocation model with at least one crosslinked IGL loop. If one ClpA subunit is crosslinked to a ClpP cleft, the rotary motion of the AAA+ ring with respect to the ClpP ring is blocked. The crosslinked ClpA subunit cannot be released from the ClpP cleft and cannot sequentially move to each position in the ClpA spiral.

Subunits in both the ClpA and ClpX hexamers adopt a shallow helical conformation with axial pore loops that interact with an extended substrate polypeptide to form a structure reminiscent of a spiral staircase. By contrast, ClpP subunits are arranged in near-planar rings that enclose a chamber with luminal peptidase active sites. Other AAA+ proteases have similar architectures, with spiral AAA+ rings and planar peptidase rings (Puchades *et al*., 2020). In structures of heterohexameric AAA+ protease motors in which the positions of unique subunits can be determined, different subunits can occupy the highest and lowest spiral positions, suggesting that dynamic rearrangement of subunits within the spiral is part of the ATP-fueled mechanical cycle that powers substrate translocation (de la Peña *et al.*, 2018; Dong *et al.*, 2019). In one model for this cycle, an enzyme power stroke is initiated when the second lowest subunit in the spiral hydrolyzes ATP (S5 at the beginning of the cycle, **Figure 2.1A**), resulting in a rearrangement that moves this subunit and higher subunits, together with bound substrate, each down one position in the spiral, at the same time that the lowest subunit (S6) disengages from substrate and moves to the top of the spiral (**Figure 2.1B**; Puchades *et al*., 2020). Intriguingly, in recent cryo-EM structures of ClpAP and ClpXP, an empty ClpP cleft is always flanked by clefts that interact with the IGL/IGF loops of the second lowest and lowest subunits within the spiral (S5 and S6 in **Figure 2.1A**) (Fei *et al.,* 2020a; 2020b; Ripstein *et al*., 2020; Lopez *et al.,* 2020). If different subunits in the ClpA or ClpX hexamers pass sequentially through each position in the spiral during substrate translocation and the empty cleft in ClpP is always between clefts that interact with specific subunits in the spiral, then the AAA+ ring should rotate with respect to ClpP during protein translocation (**Figure 2.1B**; Ripstein *et al*., 2020; Lopez *et al.,* 2020).

Here, we test the effects of preventing rotation of the ClpA ring with respect to the ClpP ring by covalently crosslinking multiple IGL loops of ClpA to ClpP (**Figure 2.1C**). We find that an enzyme containing multiple covalent crosslinks between ClpA and ClpP retains substantial proteolytic activity against unfolded and metastable native substrates but displays defects in degrading more stably folded proteins. We conclude that rotation of ClpA with respect to ClpP is not required for

substrate translocation or unfolding, but some freedom of movement at the ClpA-ClpP interface is likely to be important for optimal mechanical activity.

RESULTS AND DISCUSSION

Crosslinking ClpA to ClpP

For crosslinking studies, we used cysteine-free genetic backgrounds for ClpA (C47S, C203S, C243S; Zuromski *et al*., 2020) and ClpP (C91V, C113A; Amor *et al.*, 2016). We then introduced an E613C mutation into the IGL loop of otherwise cysteine-free ClpA (^{E613C}ClpA‡) and appended a cysteine after Asn^{193}, the C-terminal residue of otherwise cysteine-free ClpP (ClpP^{+C}). Based on cryo-EM structures of ClpAP (Lopez *et al.*, 2020), the cysteines introduced by these mutations should be close enough to allow crosslinking of specific subunits of ClpA to neighboring subunits of ClpP. For example, **Figure 2.2A** shows that Glu^{613} in each subunit of the ClpA hexamer is close to a ClpP Arg^{192} residue, the last ClpP amino acid visible in the ClpAP structure, in six of the seven ClpP protomers. We mixed ^{E613C}ClpA‡ with ClpP^{+C} in the presence of a homobifunctional cysteine crosslinker and then separated covalently joined ^{E613C}ClpA‡–ClpP^{+C} complexes (peak 1) from uncrosslinked ClpP^{+C} (peak 2) by size-exclusion chromatography (**Figure 2.2B**). After pooling fractions containing ^{E613C}ClpA‡–ClpP^{+C} complexes (**Figure 2.2B**), quantification by SDS-PAGE revealed that 90 ± 1% of the ClpA was crosslinked to ClpP (designated A–P) (**Figure 2.2C** lane 7). Based on this crosslinking efficiency, the vast majority of ClpA hexamers should contain one or more crosslinked A–P subunits (>99.99%, assuming independent crosslinking), and ~98% of hexamers should contain four, five, or six ClpA subunits crosslinked to ClpP (**Figure 2.2—S1**). The ^{E613C}ClpA‡–ClpP^{+C} pool also contained uncrosslinked ClpP^{+C}, as expected, and some crosslinked ClpP^{+C} dimers (**Figure 2.2C**, lane 7).

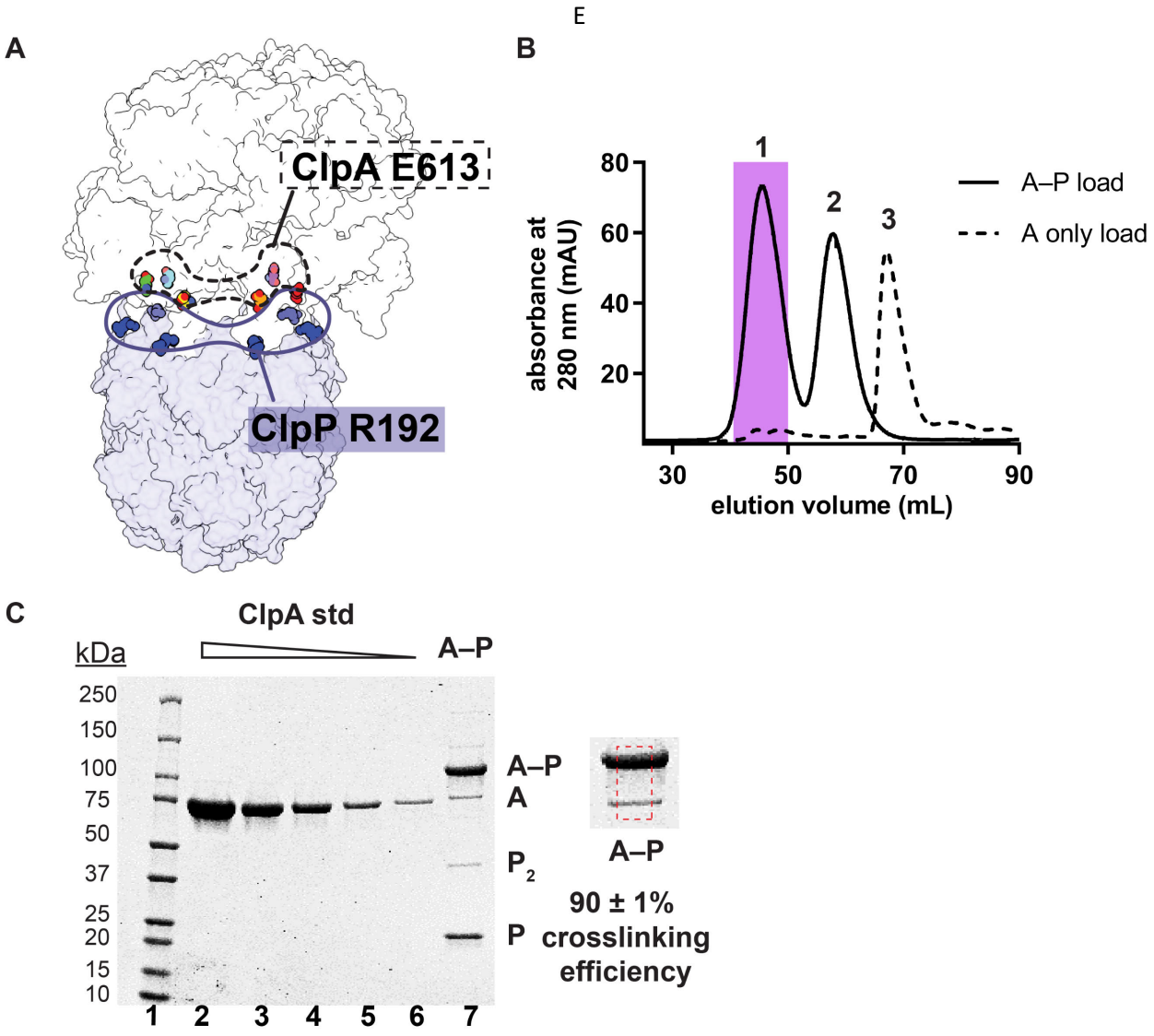


Figure 2.2 ClpA–ClpP crosslinking and purification. (**A**) Proximity of Glu^{613} residues in six subunits of the hexameric ClpA ring (shown in the dashed outline) to Arg^{192} residues in six of the seven subunits of a heptameric ClpP ring (shown in the solid outline). (**B**) Size-exclusion chromatograms of $^{E613C}ClpA^{‡}$–$ClpP^{+C}$ following crosslinking (solid line; peaks 1 and 2) or uncrosslinked $^{E613C}ClpA^{‡}$ (dashed line; peak 3), which is largely monomeric under the chromatography conditions. As shown in panel C, most ClpA in peak 1 is crosslinked to ClpP. The shaded area in peak 1 represents the crosslinked A–P that was pooled and used in all experiments in this study. Peak 2 corresponds to uncrosslinked $ClpP^{+C}$ remaining after the crosslinking reaction and chromatographs at the position expected for a tetradecamer. (**C**) Reducing SDS-PAGE of the peak-1 pool. Lanes 1-6 are MW standards or different concentrations of purified $^{E613C}ClpA^{‡}$. Lane 7 is an aliquot of the peak-1 pool. The shift to higher molecular weight from uncrosslinked ClpA (A) to crosslinked ClpA–ClpP (A–P) is consistent with covalent linkage of a single ClpA monomer (~83 kDa) to a single ClpP monomer (~23 kDa). The dashed red box is a zoomed-in view of lane 7 used to calculate crosslinking efficiency of $^{E613C}ClpA^{‡}$ to $ClpP^{+C}$. Crosslinking efficiency was calculated as the mean ± 1 SD of four independent replicates.

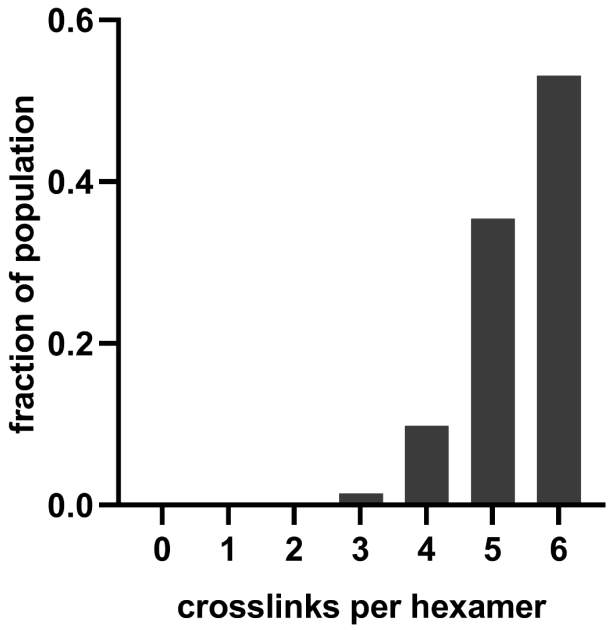


Figure 2.2—S1 Histogram of the expected number of crosslinks between ^{E613C}ClpA‡ and ClpP^{+C} in the A–P pool assuming independent crosslinking of individual ClpA and ClpP subunits with 90 ± 1% efficiency.

Crosslinked complexes degrade model substrates

To test if ClpA rotation relative to ClpP is required for ATP-fueled proteolysis, we measured ATP hydrolysis and degradation of model substrates by the purified crosslinked A–P pool *in vitro* compared to an A•P control consisting of assembled but uncrosslinked ^{E613C}ClpA‡ and wild-type ClpP (ClpP^{WT}). The A–P pool hydrolyzed ATP at a steady-state rate of 412 ± 40 min^{-1} enz^{-1}, whereas this rate was 1035 ± 78 min^{-1} enz^{-1} for the A•P control. The A•P control hydrolyzed ATP and degraded ssrA-tagged proteins at comparable rates to wild-type ClpAP (ClpA^{WT}ClpP^{WT}) and a cysteine-free ClpA (ClpA^{CF}) variant (ClpA^{CF}ClpP^{WT}), demonstrating that the C47S, C203S, C243S, and E613C mutations do not impair activity (**Figure 2.3—S1**). We also observed similar rates of ATP hydrolysis and ssrA-tagged protein degradation when ClpP^{+C}, the ClpP variant used for crosslinking, was paired with ClpA^{WT}, ClpA^{CF}, or ^{E613C}ClpA‡.

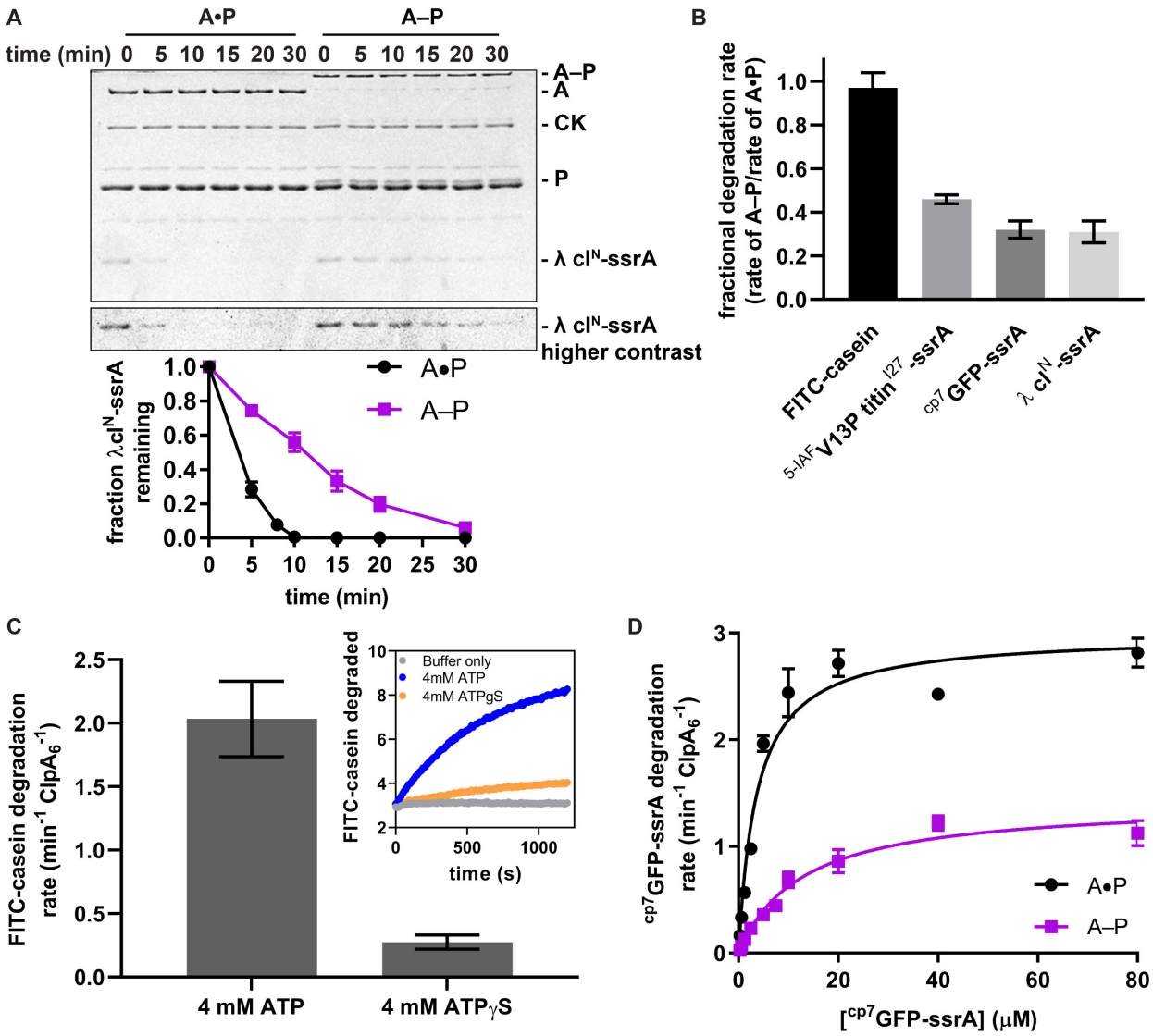


Figure 2.3 Substrate degradation by crosslinked ClpAP (A–P) and uncrosslinked ClpAP (A•P). (**A**) Top, SDS-PAGE assay of the kinetics of λ cI^N-ssrA degradation by A–P and the A•P control (CK is creatine kinase). Bottom, quantification of λ cI^N-ssrA degradation. Values are means ± 1 SD (n = 3). (**B**) Degradation of substrates of varying thermodynamic stability (18 μM FITC-casein, 5 μM $^{5\text{-IAF}}$titin^{V13P}-ssrA, 20 μM ^{cp7}GFP-ssrA, 15 μM λ cI^N-ssrA by A–P. Fractional degradation rates were calculated by dividing the degradation rates of A–P by the A•P rates. Values are means ± propagated error (n ≥ 3). (**C**) Degradation of FITC-casein (18 μM) by A–P in the presence of ATP or ATPγS. FITC-casein degradation was quantified by normalizing the relative fluorescence units to the total FITC-casein degraded upon porcine elastase addition at the endpoint of the assay and subtracting the contributions of photobleaching from the buffer-only control. Values are means ± 1 SD (n = 3). The inset shows representative degradation kinetics. (**D**) Michaelis-Menten analysis of ^{cp7}GFP-ssrA degradation kinetics by A–P and the A•P control. Values are means ± 1 SD (n = 3). For A–P degradation, V_{max} was 1.4 ± 0.07 min^{-1} $ClpA_6^{-1}$, K_M was 13 ± 1.6 μM, and R^2 was 0.96; for the A•P control, V_{max} was 3.0 ± 0.10 min^{-1} $ClpA_6^{-1}$, K_M was 3.7 ± 0.5 μM, and R^2 was 0.96, where the errors are those of non-linear least-squares fitting to the Michaelis-Menten equation.

To test the effects of crosslinking ClpA to ClpP on proteolysis, we measured the rate at which the A–P and A•P enzymes degraded proteins with a range of native stabilities. These substrates included the N-terminal domain of the phage λ cI repressor with an ssrA tag (λ cI^{N}-ssrA; Gottesman *et al.,* 1998), ^{cp7}GFP-ssrA (Nager *et al.*, 2011), $^{5\text{-IAF}}$V13P titin^{I27}-ssrA (Kenniston *et al.*, 2003; Iosefson *et al.,* 2015), and FITC-casein (Twining, 1984; Thompson *et al.*, 1994). Under the conditions of these assays *in vitro*, the A–P sample degraded the folded substrates (λ cI^{N}-ssrA and ^{cp7}GFP-ssrA) at rates that were 31 ± 5% and 32 ± 4%, respectively, of the A•P control, and degraded unfolded $^{5\text{-IAF}}$V13P titin^{I27}-ssrA and FITC-casein at 46 ± 2% and 97 ± 7%, respectively, of the control rates (**Figure 2.3A-B**). The rate of degradation of FITC-casein by the A–P pool was reduced ~6-fold when ATPγS was substituted for ATP (**Figure 2.3C**), indicating that robust degradation of this molten-globule substrate requires ATP hydrolysis. We also determined steady-state kinetic parameters for degradation of ^{cp7}GFP-ssrA by the A–P pool and A•P control (**Figure 2.3D**). Compared to the A•P control, V_{max} was ~50% and K_M was ~3-fold weaker for degradation of this substrate by the A–P pool. This reduction in V_{max} for the A–P pool was roughly comparable to its reduced ATP-hydrolysis activity, suggesting that slower degradation of folded substrates by the A–P pool results from slower ATPase activity. Thus, our results show that multiple crosslinks between ClpA and ClpP in the A–P pool cause modest slowing of the rates of ATP hydrolysis and protein degradation compared to the uncrosslinked controls (**Figure 2.3—S1**), with more prominent degradation defects for native substrates. Notably, however, crosslinks between ClpA and ClpP do not prevent the protein unfolding or translocation steps required for proteolysis. Only the crosslinked A–P sample exhibited substantially lower ATP-hydrolysis and protein degradation activity compared to the uncrosslinked controls; thus, the reduced A–P enzymatic activities are likely to be direct consequences of introducing specific covalent crosslinks between ClpA and ClpP rather than other modifications introduced in the experimental design.

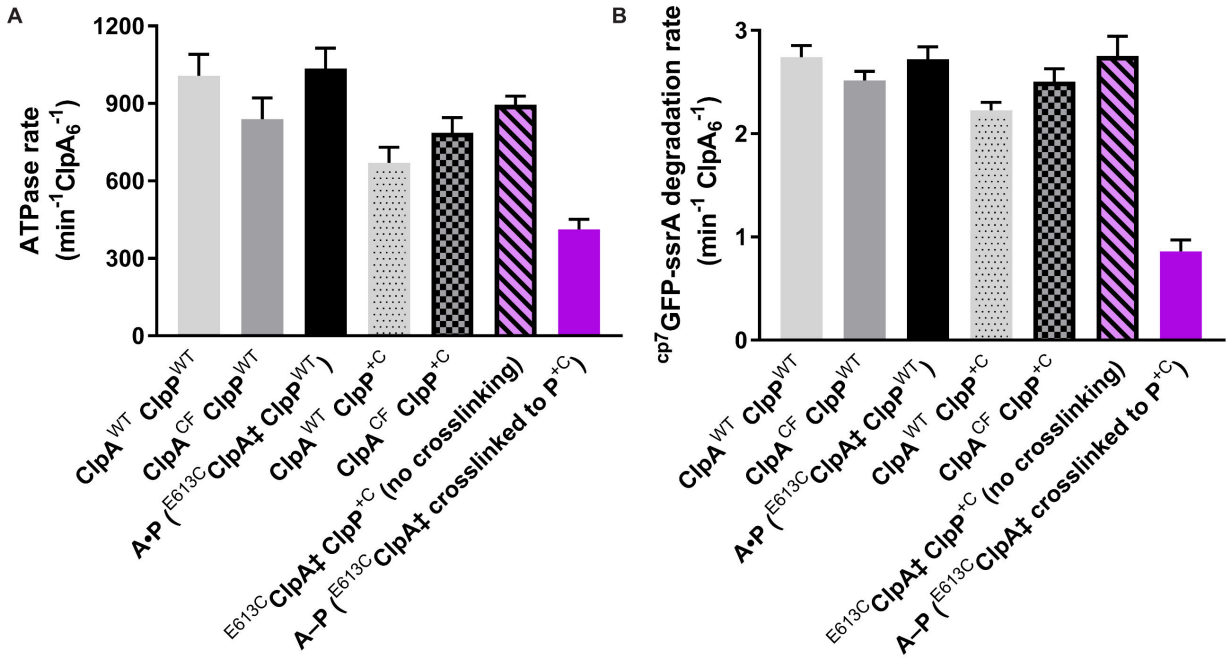


Figure 2.3—S1 ATPase and degradation activities of uncrosslinked ^{E613C}ClpA‡ClpP^{+C} and ^{E613C}ClpA‡ClpP^{WT} (A•P) complexes are comparable to those of wild-type ClpAP. (**A**) Hydrolysis of ATP (5 mM) by crosslinked ClpAP (A–P) and uncrosslinked controls consisting of wild-type ClpA (ClpA^{WT}), cysteine-free ClpA (ClpA^{CF}), or ^{E613C}ClpA‡ and either wild-type ClpP (ClpP^{WT}) or ClpP^{+C}. (**B**) Degradation of ^{cp7}GFP-ssrA (20 μM) by crosslinked (A–P) and uncrosslinked ClpAP (A•P) variants, as shown in panel (A). A•P and uncrosslinked ^{E613C}ClpA‡ClpP^{+C} degrade ^{cp7}GFP-ssrA (20 μM) at similar rates as wild-type ClpAP and other uncrosslinked controls, and faster than crosslinked A–P. Values are means ± 1 SD (n ≥ 3).

Mechanistic implications of crosslinked complex activity

Models in which ClpA or ClpX must rotate with respect to ClpP to allow substrate translocation (**Figure 2.1B-C**) predict that crosslinking ClpA or ClpX to ClpP would abolish protein degradation by stopping rotation and linked sequential movements of ClpA/ClpX subunits through each position in the spiral. Our experimental results do not support these models. Rather, we find that preventing rotation by "riveting" the ClpA ring to the ClpP ring still permits substantial degradation of native and denatured protein substrates *in vitro*. ClpXP complexes in which one IGF loop is crosslinked to ClpP can also degrade folded and unfolded substrates, albeit at lower rates than uncrosslinked controls (Bell, 2020). The high degree of crosslinking in our ClpAP experiments, where 98% of complexes contain at least four covalent crosslinks between ClpA and ClpP, and

~50% of complexes are predicted to contain six crosslinks (**Figure 2.2—S1**), would be expected to hinder each ClpA subunit from adopting each position in the spiral by affecting conformational accessibility, especially near the ClpP interface. Moreover, in approximately half of the crosslinked enzymes, it would not be possible to have two empty ClpP clefts. Hence, the proposal that this intermediate is a requisite step in translocation, as proposed for ClpXP (Ripstein *et al.*, 2020), is also inconsistent with our results. Lopez *et al.* (2020) proposed that ClpA and ClpP might rotate in defined contexts, for example during the degradation of very stable substrates. Although we cannot exclude this possibility, we prefer simpler models in which the basic mechanism of AAA+ protease function does not change in a substrate-specific manner. As we observe reduced rates of degradation of folded substrates when ClpA and ClpP are crosslinked, conformational flexibility between the unfoldase and protease appears to be important for optimal unfolding. However, rotation of the ClpA or ClpX rings with respect to ClpP is clearly not a strict requirement for degradation. We suggest, therefore, that ring-ring rotation models be considered to be both unproven and unlikely in the absence of direct evidence for such rotation.

AAA+ proteases in the FtsH/Yme1/Agf3l2 and Lon families have AAA+ and peptidase modules that are genetically tethered as part of the same polypeptide chain and therefore also must operate without rotation between the unfoldase and protease components (Glynn, 2017). For ClpAP, a non-rotary mechanism could be explained by the sequential hand-over-hand mechanism if the empty ClpP cleft can localize between any pair of ClpA subunits. Alternative mechanisms, such as the reciprocating action of one or two AAA+ subunits, might also explain both the observed pattern of unfoldase-protein interactions seen in cryo-EM structures and the robust degradation activity of genetically or biochemically tethered AAA+ proteases against multiple substrates. Further experiments will be required to discriminate between these models.

MATERIALS AND METHODS

KEY RESOURCES TABLE				
Reagent type (species) or resource	**Designation**	**Source or reference**	**Identifiers**	**Additional information**
gene (*Escherichia coli*)	clpA			UniProtKB - P0ABH9
gene (*Escherichia coli*)	clpP			UniProtKB - P0A6G7
strain, strain background (*Escherichia coli*)	T7 Express	New England Biolabs	C2566I	Chemically competent cells
recombinant DNA reagent	pT7 ClpP^{+C} (plasmid)	This paper		For overexpression of C-terminally His_6-tagged ClpP (C91V, C113A) with extra Cys residue for crosslinking. Progenitor: pT7 ClpP-TEV-cHis_6 (Stinson *et al*. 2013, Amor *et al.* 2016)
recombinant DNA reagent	pET23b His_7SumoFLAG ^{E613C}ClpA‡ (plasmid)	This paper		For overexpression of ClpA with Cys substitution and C47S, C203S, C243S background for crosslinking. Progenitor: pET23b

				His_7Sumo $ClpA^{cf}$ΔC9 (Zuromski *et al.* 2020)
recombinant DNA reagent	pET23b His_7Sumo $ClpA^{cf}$ΔC9 (plasmid)	Zuromski *et al.* 2020		For overexpression of cysteine-free ClpA ($ClpA^{CF}$) harbouring C47S, C203S, C243S mutations
recombinant DNA reagent	WT ClpA (plasmid)	Seol *et al.* (1994), Hou *et al.* (2008)		WT ClpA (M169T background) for overexpression
recombinant DNA reagent	ClpP-His_6 (plasmid)	Kim *et al*. (2000)		WT ClpP for overexpression
recombinant DNA reagent	^{cp7}GFP-ssrA (plasmid)	Nager *et al.* (2011)		Circularly permutated variant of superfolder GFP-ssrA for overexpression
recombinant DNA reagent	V13P titin^{I27}-ssrA (plasmid)	Kenniston *et al.* 2003		ssrA-tagged I27 domain variant for overexpression
recombinant DNA reagent	His_6SUMO λ cI^{N}-ssrA (plasmid)	This paper		ssrA-tagged residues 1-93 of λ cI (UniProtKB - P03034) for overexpression
chemical compound, drug	bismaleimidoethane	Thermo Fisher Scientific	Cat # 22323	

chemical compound, drug	Adenosine 5'-O-(3-Thiotriphosphate), Tetralithium Salt	Millipore Sigma	Cat# 119120-25MG	
chemical compound, drug	5-Iodoacetamidofluorescein	Thermo Fisher Scientific	Cat# I30451	
chemical compound, drug	Casein fluorescein isothiocyanate from bovine milk (FITC-casein)	Sigma-Aldrich	Cat# C0528-10MG	

Proteins

The gene encoding *E. coli* $ClpP^{+C}$ was generated using PCR mutagenesis, and the corresponding protein was purified by established protocols (Martin *et al.*, 2005) and stored in buffer containing 0.5 mM dithiothreitol (DTT). Wild-type ClpP ($ClpP^{WT}$) was purified by established protocols (Kim *et al.*, 2001). The plasmid for $^{E613C}ClpA^{‡}$ was generated by PCR mutagenesis of *E. coli* $ClpA^{\Delta C9}$ fused to the 3´-end of His_7SumoFLAG cloned into pET23b (Novagen). His_7SumoFLAG-$^{E613C}ClpA^{‡}$ was overexpressed in T7Express (New England Biolabs), and initially purified by Ni-NTA chromatography. Following ULP-1 cleavage to remove His_7SumoFLAG, $^{E613C}ClpA^{‡}$ was further purified by SP-Sepharose cation-exchange chromatography and stored in 50 mM HEPES-KOH (pH 7.5), 300 mM NaCl, 20 mM $MgCl_2$, 10% glycerol, and 2 mM TCEP. Cysteine-free ($ClpA^{CF}$) and wild-type ($ClpA^{WT}$) were purified by established protocols (Zuromski *et al*, 2020; Hou *et al.*, 2008). The ^{cp7}GFP-ssrA and V13P titin^{I27}-ssrA proteins were purified as described (Nager *et al.*, 2011; Kenniston *et al.*, 2003). V13P titin^{I27}-ssrA was labeled with 5-iodoacetamidofluorescein (5-IAF) for fluorescent assays as described (Iosefson *et al.*, 2015). The plasmid for λ cI^{N}-ssrA was generated by PCR mutagenesis of a gene encoding amino acids 1-93 the bacteriophage λ cI repressor. This construct was fused to the 3´-end of His_6Sumo cloned into pET23b and appended

with the C-terminal ssrA degron. His_6Sumo-λ-cI^N-ssrA was purified by Ni-NTA chromatography, ULP-1 cleavage, Ni^{++}-NTA chromatography to remove the His_6Sumo fragment, Mono-Q anion-exchange chromatography, and Superdex-75 size-exclusion chromatography, and stored in 25 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 10% glycerol, and 1 mM DTT. ClpA and variant concentrations were calculated as hexamer equivalents, and ClpP and variant concentrations were calculated as tetradecamer equivalents.

Crosslinking ClpA to ClpP

^{E613C}ClpA‡ (4 μM) and $ClpP^{+C}$ (9.6 μM) were mixed in a total volume of 2.5 mL and desalted into 50 mM HEPES-KOH (pH 7), 300 mM NaCl, 20 mM $MgCl_2$, 10% glycerol, and 5 mM EDTA using a Sephadex G-25 PD-10 column (GE Healthcare). After diluting to a final volume of 5 mL, crosslinking was initiated by addition of 5 mM ATPγS and 200 μM bismaleimidoethane (BMOE; Thermo Fisher) and allowed to proceed at room temperature for 45 min. The reaction was quenched by addition of 50 mM DTT at room temperature for 20 min before purification by Superdex-200 size-exclusion chromatography in 50 mM HEPES-KOH (pH 7.5), 300 mM NaCl, 20 mM $MgCl_2$, 10% glycerol, and 2 mM TCEP. The purified A–P pool was used for all subsequent biochemical assays. Quantification of crosslinking and the concentration of A–P were measured by quantifying Coomassie-stained SDS-PAGE bands relative to ^{E613C}ClpA‡ standards. The area under the curve (AUC) corresponding to pixel intensities of the crosslinked and uncrosslinked species were quantified by ImageQuant (GE Healthcare) after scanning Coomassie-stained SDS-PAGE using a Typhoon FLA 9500 (GE Healthcare). Crosslinking efficiency was measured in four independent replicates by multiplying the calculated concentration of each species by the volume loaded in each lane, and calculated as:

$$\text{Efficiency} = \frac{\text{picomoles A–P}_{\text{Crosslinked}}}{\text{picomoles A–P}_{\text{Crosslinked}} + \text{picomoles } {}^{E613C}\text{ClpA}^{\ddagger}_{\text{Uncrosslinked}}}$$

Biochemical Assays

We determined the concentration of ^{E613C}ClpA‡–ClpP^{+C} by a standard-curve comparison to ^{E613C}ClpA‡ (Figure 2.2C). We calculated the concentration of the uncrosslinked ^{E613C}ClpA‡ species by measuring the absorbance at 280 nm (ε = 32890 M^{-1} cm^{-1}) using a NanoDrop One UV-Vis Spectrophotometer (Thermo-Fisher Scientific), ATP-hydrolysis assays were performed using an NADH-coupled assay (Martin *et al.*, 2005) at 30 °C in Buffer HO (25 mM HEPES-KOH, pH 7.5, 300 mM NaCl, 20 mM $MgCl_2$, 10% glycerol, 2 mM TCEP) with 5 mM ATP and 0.25 μM ^{E613C}ClpA‡ and 0.75 μM ClpP^{WT} for the A•P control; 0.25 μM ^{E613C}ClpA‡–ClpP^{+C} for the A–P pool; or the combinations of 0.25 μM ClpA^{WT}, ClpA^{CF}, or ^{E613C}ClpA‡ and 0.75 μM ClpP^{WT} or ClpP^{+C} listed in Figure 2.3—S1A. Degradation reactions were performed at 30 °C in Buffer HO with 4 mM ATP and an ATP-regeneration system consisting of 50 μg/mL creatine kinase (Millipore-Sigma) and 5 mM creatine phosphate (Millipore-Sigma). Degradation of ^{cp7}GFP-ssrA was monitored by loss of substrate fluorescence (excitation 467 nm; emission 511 nm) using a SpectraMax M5 plate reader (Molecular Devices) (Nager *et al.*, 2011). The ^{cp7}GFP-ssrA concentration was 20 μM in Figure 2.3B and Figure 2.3—S1B; concentrations varying from 0.31 to 80 μM in Figure 2.3D) contained 0.25 μM ^{E613C}ClpA‡ and 0.75 μM ClpP^{WT} for the A•P control or 0.25 μM ^{E613C}ClpA‡–ClpP^{+C} for the A–P pool. In the ^{cp7}GFP-ssrA degradation assays shown in Figure 2.3—S1B, degradation reactions for uncrosslinked controls included the indicated combinations of 0.25 μM ClpA^{WT}, ClpA^{CF}, or ^{E613C}ClpA‡ and 0.75 μM ClpP^{WT} or ClpP^{+C}, in addition to 0.25 μM ^{E613C}ClpA‡ and 0.75 μM ClpP^{WT} for the A•P control or 0.25 μM ^{E613C}ClpA‡–ClpP^{+C} for the A–P pool. Degradation of FITC-casein (18 μM, Sigma-Aldrich) containing 0.25 μM ^{E613C}ClpA‡ and 0.75 μM ClpP^{WT} for the A•P control or 0.25 μM ^{E613C}ClpA‡–ClpP^{+C} for the A–P pool was monitored by increase in fluorescence (excitation 340 nm; emission 520 nm); to determine the endpoint of complete FITC-casein degradation, 0.5 μL of 5 mg/mL porcine elastase (Sigma-Aldrich) was added to each well and incubated for 30 min. ClpAP degradation reactions with FITC-casein (18 μM) were performed at 30 °C in Buffer HO with 4 mM ATP or ATPγS (Millipore-Sigma). Degradation of $^{5\text{-IAF}}$V13P titin^{I27}-

ssrA (5 μM) containing 0.2 μM ^{E613C}ClpA‡ and 0.75 μM ClpP^{WT} for the A•P control or 0.2 μM ^{E613C}ClpA‡–ClpP^{+C} for the A–P pool was monitored by increase in fluorescence (excitation 494 nm; emission 518 nm). Gel degradation of λ cI^{N}-ssrA (15 μM monomer) containing 0.2 μM ^{E613C}ClpA‡ and 0.4 μM ClpP^{WT} for the A•P control or 0.2 μM ^{E613C}ClpA‡–ClpP^{+C} for the A–P pool was performed in triplicate by taking samples of each reaction at specific time points, stopped by addition of SDS-PAGE loading sample and boiling at 100 °C before loading on Tris-Glycine-SDS gels. Bands were visualized by staining with colloidal Coomassie G-250 and quantified by ImageQuant (GE Healthcare) after scanning by Typhoon FLA 9500 (GE Healthcare). The fraction of λ cI^{N}-ssrA remaining was calculated by dividing the intensity of this band at a given time point by the density at time zero, after normalization by the creatine kinase density. The biochemical assays were performed with A–P from a single preparation to ensure that crosslinking efficiency was the same throughout all assays. All experiments were performed in at least three independent replicates and values reported were calculated as the mean ± 1 SD of independent replicates or the ratio of means ± propagated error of independent replicates. Propagated error for "fractional degradation rate (rate of A–P / rate of A•P" of A–P mean activity compared to A•P mean activity was computed as:

$$\text{Propagated error of } \frac{\text{mean}_{A-P}}{\text{mean}_{A\bullet P}} = \frac{\text{mean}_{A-P}}{\text{mean}_{A\bullet P}} \sqrt{\left(\text{SD}_{A-P} / \text{mean}_{A-P}\right)^2 + \left(\text{SD}_{A\bullet P} / \text{mean}_{A\bullet P}\right)^2}$$

Acknowledgments: We are grateful to X. Fei (MIT) for helpful advice on preparation of crosslinked samples and A. Torres-Delgado (MIT) and H.C. Kotamarthi (MIT) for providing materials. We thank J.R. Kardon, A.O. Olivares, and K.R. Schmitz for thoughtful comments on the manuscript. This work was supported by NIH grants GM-101988 (RTS), AI-016892 (RTS, TAB), and the Howard Hughes Medical Institute (TAB). T.A. Bell and S.K. were supported by NIH training grant 5T32GM-007287; S.K. and K.L.Z. were supported by NSF grant GRFP 1745302.

REFERENCES

Amor, A. J., Schmitz, K. R., Sello, J. K., Baker, T. A., & Sauer, R. T. (2016). Highly dynamic interactions maintain kinetic stability of the ClpXP protease during the ATP-fueled mechanical cycle. *ACS Chemical Biology*, **11**(6), 1552–1560.

Bell, T.A (2020). Intersubunit communication and coordinated mechanical activity in the AAA+ protease ClpXP (PhD thesis). Massachusetts Institute of Technology, Cambridge, MA, USA.

de la Peña, A.H., Goodall, E.A., Gates, S.N., Lander, G.C., and Martin, A. (2018). Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis–driven translocation. *Science* **362**(6418), eaav0725.

Dong, Y., Zhang, S., Wu, Z., Li, X., Wang, W.L., Zhu, Y., Stoilova-McPhie, S., Lu, Y., Finley, D., and Mao, Y. (2019). Cryo-EM structures and dynamics of substrate-engaged human 26S proteasome. *Nature* **565**, 49–55.

Fei X., Bell T.A., Jenni S., Stinson B.M., Baker T.A., Harrison S.C., Sauer R.T. (2020a). Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. *eLife* **9**:e52774.

Fei X., Bell T.A., Barkow, S.R., Baker T.A., Harrison S.C., Sauer R.T. (2020b). Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. *eLife* **9**:e61496.

Glynn, S.E. (2017). Multifunctional mitochondrial AAA proteases. *Frontiers in Molecular Biosciences* **4**(34).

Gottesman, S., Roche, E., Zhou, Y., & Sauer, R. T. (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes & Development* **12**(9), 1338–1347.

Hou J.Y., Sauer R.T., Baker T.A. Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. (2008). *Nature Structural & Molecular Biology* **15**(3), 288-294.

Iosefson, O., Nager, A.R., Baker, T.A., and Sauer, R.T. (2015). Coordinated gripping of substrate by subunits of a AAA+ proteolytic machine. *Nature Chemical Biology* **11**, 201–206.

Kenniston, J.A., Baker, T.A., Fernandez, J.M., and Sauer, R.T. (2003). Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* **114**, 511–520.

Kim, Y., Levchenko, I., Fraczkowska, K., Woodruff R.V., Sauer R.T., Baker T.A. (2001). Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nature Structural Biology* **8,** 230–233.

Lopez, K.E., Rizo, A.N., Tse, E., Lin, J., Scull, N.W., Thwin, A.C., Lucius, A.L., Shorter, J., Southworth, D.R. (2020). Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis. *Nature Structural & Molecular Biology* **27**, 406–416.

Martin, A., Baker, T.A., and Sauer, R.T. (2005). Rebuilt AAA + motors reveal operating principles for ATP-fueled machines. *Nature* **437**, 1115–1120.

Nager, A.R., Baker, T.A., and Sauer, R.T. (2011). Stepwise unfolding of a β barrel protein by the AAA+ ClpXP protease. *Journal of Molecular Biology* **413**, 4–16.

Puchades, C., Sandate, C.R., Lander, G.C. (2020). The molecular principles governing the activity and functional diversity of AAA+ proteins. *Nature Reviews Molecular Cell Biology* **21**, 43-58.

Ripstein, Z. A., Vahidi, S., Houry, W. A., Rubinstein, J. L., & Kay, L. E. (2020). A processive rotary mechanism couples substrate unfolding and proteolysis in the ClpXP degradation machinery. *eLife* **9**, e52158.

Sauer, R.T., and Baker, T.A. (2011). AAA+ Proteases: ATP-fueled machines of protein destruction. *Annual Review of Biochemistry* **80**, 587–612.

Seol J.H., Yoo S.J., Kim K.I., Kang M.S., Ha D.B., Chung C.H. (1994). The 65-kDa protein derived from the internal translational initiation site of the *clpA* gene inhibits the ATP-dependent protease Ti in *Escherichia coli*. *The Journal of Biological Chemistry* **269**, 29468–29473.

Stinson, B. M., Nager, A. R., Glynn, S. E., Schmitz, K. R., Baker, T. A., & Sauer, R. T. (2013). Nucleotide binding and conformational switching in the hexameric ring of a AAA+ machine. *Cell* **153**(3), 628–639.

Thompson, M.W., Singh S.K., and Maurizi M.R. (1994). Processive degradation of proteins by the ATP-dependent Clp protease from *Escherichia coli. The Journal of Biological Chemistry* **269**, 18209–18215.

Twining, S.S. (1984). Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes. *Analytical Biochemistry* **143**, 30-34.

Zuromski, K.L., Sauer, R.T., and Baker, T.A. (2020). Modular and coordinated activity of AAA+ active sites in the double-ring ClpA unfoldase of the ClpAP protease. *Proc. Natl. Acad. Sci. USA.* **117**(41), 25455-25463.

CHAPTER 3

AAA+ protease-adaptor structures reveal altered conformations and ring specialization

This chapter has been written as a manuscript for publication.

Sora Kim[1,3], Xue Fei[1-3], Robert T. Sauer[1], and Tania A. Baker[1]

[1] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA
[2] Present address: Kymera Therapeutics Inc., Watertown, MA, USA
[3] These authors contributed equally: Sora Kim, Xue Fei

Author contributions:

S. Kim: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review and editing, Visualization, Funding acquisition

X. Fei: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review and editing, Visualization

R.T. Sauer: Conceptualization, Methodology, Validation, Writing – review & editing, Supervision, Funding acquisition

T.A. Baker: Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition

ABSTRACT

ClpAP, a two-ring AAA+ protease, degrades N-end-rule proteins bound by the ClpS adaptor. Here, we present high-resolution cryo-EM structures of ClpAPS complexes showing how ClpA pore loops interact with the ClpS N-terminal extension (NTE), which is normally intrinsically disordered. In two classes, the NTE is bound by a spiral of pore-1 and pore-2 loops in a manner similar to substrate-polypeptide binding by many AAA+ unfoldases. Kinetic studies reveal that pore-2 loops of the ClpA D1 ring catalyze protein remodeling required for substrate delivery by ClpS. In a third class, D2 pore-1 loops are rotated, tucked away from the channel, and do not bind the NTE, demonstrating asymmetry in engagement by the D1 and D2 rings. These studies show new structures and functions for key AAA+ elements. Pore-loop tucking may be used broadly by AAA+ unfoldases, for example during enzyme pausing/unloading.

INTRODUCTION

Regulated proteolysis by energy-dependent AAA+ (ATPases associated with diverse cellular activities) proteases maintains cellular homeostasis of proteins in all organisms. AAA+ proteases harness the energy of ATP hydrolysis to degrade regulatory proteins and proteins that are damaged, misfolded, or no longer needed (Sauer and Baker, 2011). AAA+ proteases, such as ClpAP, consist of a hexameric AAA+ unfoldase (*e.g.* $ClpA_6$) and a self-compartmentalized peptidase (*e.g.* $ClpP_{14}$). In the recognition step of degradation, a peptide sequence (degron) in a protein substrate is engaged by pore loops lining the axial channel of the AAA+ hexamer. Through conformational changes powered by ATP-hydrolysis cycles, native structure in the bound substrate is unfolded and processively translocated through the channel and into the peptidase chamber, where the polypeptide is cleaved into fragments. In addition to binding and engaging degrons directly, AAA+ proteases interact with adaptor proteins that modify their substrate specificity, often by tightening the affinity of the enzyme•substrate complex (Sauer and Baker, 2011; Mahmoud and Chien, 2018).

Prokaryotes and eukaryotes use the N-end-rule pathway to target proteins bearing specific N-terminal residues (called N-degrons) for rapid degradation (Varshavsky, 2019). In *Escherichia coli*, the ClpS adaptor promotes ClpAP degradation of proteins containing Leu, Phe, Tyr, or Trp residues at the N-terminus (Tobias et al., 1991; Dougan et al., 2002; Zeth et al., 2002a, 2002b; Erbse et al., 2006; Varshavsky, 2019). ClpS (~10 kDa) docks with the N-terminal domain of ClpA and contains a hydrophobic pocket that binds the N-end-rule residue (Zeth et al., 2002a, 2002b; Wang et al., 2008a; Román-Hernández et al., 2009). ClpS functions as a specificity switch for ClpAP, promoting degradation of N-degron substrates while inhibiting degradation of ssrA-tagged and related substrates (Dougan et al., 2002; Guo et al., 2002; Erbse et al., 2006; Wang et al., 2008a). Unlike some adaptors that simply act as molecular matchmakers between the substrate and enzyme, ClpS also modulates the rate of ATP hydrolysis and substrate unfolding by ClpA.

Interestingly, ClpS is proposed to interact with ClpA as a 'pseudo-substrate' (Dougan et al., 2002; De Donatis et al., 2010; Román-Hernández et al., 2011; Rivera-Rivera et al., 2014; Torres-Delgado et al., 2020). Specifically, the N-terminal extension (NTE) of free ClpS is exposed as an unstructured peptide, mimicking a degron. The NTE is poorly conserved among orthologs, with the exception of a short junction sequence adjacent to the ClpS core that typically contains a few tandem prolines (Hou et al., 2008; Román-Hernández et al., 2011). During ClpS-assisted degradation, a ClpS•N-degron substrate complex initially binds to ClpA. Subsequently, the N-degron substrate is transferred to ClpA for degradation and ClpS escapes destruction by mechanisms that are poorly understood.

Each ClpA subunit has two AAA+ modules, called D1 and D2, that associate in the hexamer to form two stacked rings (Grimaud et al., 1998). The D1 and D2 modules belong to different AAA+ subfamilies and have distinct biochemical functions (Erzberger and Berger, 2006). The D2 ring, a member of the HCLR AAA+ clade, is the principal ATPase motor responsible for unfolding and translocating substrates, including proteins with high thermodynamic stabilities (Kress et al., 2009; Kotamarthi et al., 2020; Zuromski et al., 2020). In contrast, the D1 ring, a classic AAA+ clade member, assists the D2 ring as an auxiliary motor, improves enzyme processivity, and plays a major role in substrate recognition (Kress et al., 2009; Kotamarthi et al., 2020; Zuromski et al., 2020, 2021). ClpS differentially regulates the activities of the D1 and D2 rings (Kress et al., 2009; Zuromski et al., 2021) via interactions of its NTE that we characterize here. Previous cryo-EM structures of ClpAP elucidate how the axial channel of the D1 and D2 rings engages the polypeptide of a directly recognized substrate (Lopez et al., 2020). Pore-1 and pore-2 loops in both rings form spiral-staircase-like arrangements that bind the substrate polypeptide, in a similar manner to those in structures of other double-ring AAA+ enzymes, such as Hsp104, ClpB, Cdc48/p97, and NSF (Deville et al., 2017; Gates et al., 2017; White et al., 2018; Yu et al., 2018; Cooney et al., 2019; Lo et al., 2019; Rizo et al., 2019; Twomey et al., 2019). However, these

previous structures do not provide insight into the distinct, specialized functions of each AAA+ ring of ClpA or the mechanism of ClpS-assisted degradation of N-degron substrates.

To characterize ring specialization and ClpS-ClpA collaboration, we solved cryo-EM structures of ClpAPS complexes that show how the normally disordered ClpS NTE assumes an extended conformation when bound in the ClpA axial channel. These structures reveal marked conformational differences from prior ClpAP structures (Lopez et al., 2020). We identify multiple conformations of ClpS-bound ClpA, including an arrangement in which the pore-1 loops of the D2 ring are tucked-in and face away from the channel, allowing only the D1 ring to interact strongly with the ClpS NTE. Mutagenesis and biochemical experiments establish that the pore-2 loops of the ClpA D1 ring are essential for ClpS delivery of an N-degron substrate but contribute little to docking of ClpS with ClpA. Our results demonstrate structural and functional plasticity among the ClpA pore loops, provide a structural basis for the functions of ClpS during N-degron substrate degradation, and contribute more broadly to understanding the operational modes available to AAA+ enzymes as they perform diverse biological processes.

RESULTS

Distinct conformations of ClpAPS delivery complexes

We used size-exclusion chromatography in the presence of ATPγS to purify a complex of ClpA, ClpP, ClpS, and the N-end-rule substrate YLFVQELA-GFP (**Figure 3.1a,b**). Based on SDS-PAGE, the YLFVQELA-GFP substrate appeared to be sub-stoichiometric compared to ClpS (**Figure 3.1b**). Because ATPγS does not support degradation (Thompson et al., 1994; Hoskins et al., 1998; Ishikawa et al., 2001; Effantin et al., 2010; Miller and Lucius, 2014; Lopez et al., 2020), these complexes should represent early stages in ClpS-mediated delivery of N-degron substrates.

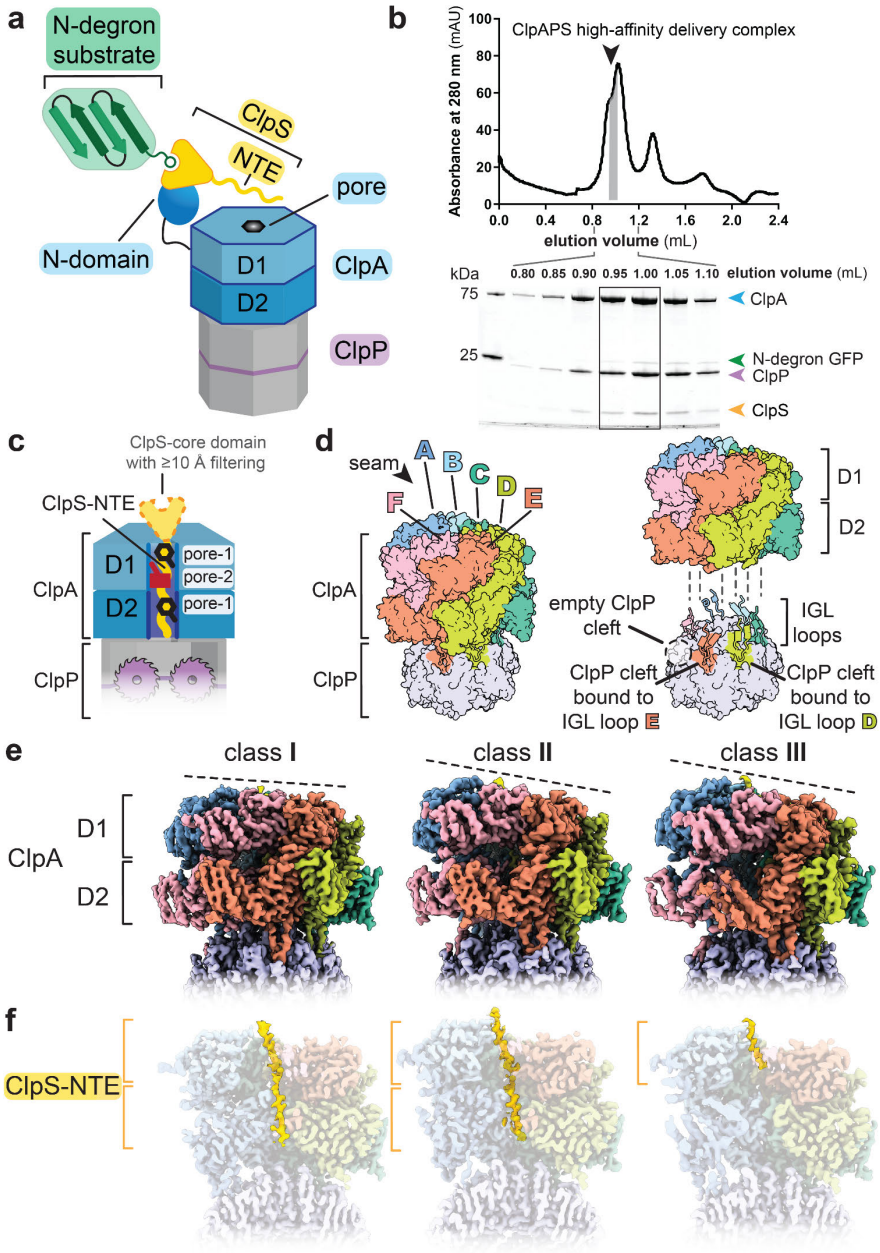


Figure 3.1 | Architectures of ClpS-bound ClpAP.
a, Cartoon of ClpS delivery of an N-degron substrate for ClpAP degradation. Native ClpS (yellow wedge) binds to the N-degron substrate (green) and also binds to an N-terminal domain of ClpA (blue oval). **b**, Size-exclusion chromatography (top panel) of a complex of ClpA, ClpP, ClpS, and YLFVQELA-GFP (an N-degron substrate) assayed by SDS-PAGE (bottom panel). Gray shaded area and boxed area indicate the fractions pooled for cryo-EM. **c**, Cartoon of pore loops that interact with the NTE and proteins resolved in cryo-EM structures. **d**, (left panel) ClpA subunit nomenclature in right-hand spiral hexamer, where the seam interface is between the lowest (F) subunit and the highest (A) subunit. The ClpA hexamer docks into clefts in the $ClpP_7$ ring via IGL loops (right panel). The empty ClpP cleft is located between the clefts occupied by subunits E and F. **e**, Side views of the cryo-EM maps of classes **I**, $\mathbf{II_c}$, and $\mathbf{III_b}$. The dashed line indicates the relative height of ClpA subunits within the spiral. **f**, Cutaway views of panel (e) showing density for the ClpS NTE colored yellow.

Table 3.1. Cryo-EM data collection, processing, model building, and validation statistics.

Class	**I**	**II**			**III**	
Subclass name	**I**	**II_a**	**II_b**	**II_c**	**III_a**	**III_b**
PDB ID	7UIX	7UIV	7UIW	7UIZ	7UIY	7UJ0
EMDB ID	26556	26554	26555	26558	26557	26559
Data collection and processing						
microscope	Talos Arctica					
camera	K3					
magnification	45,000X					
voltage (kV)	200					
total electron dose ($e^-/Å^2$)	34.71					
defocus range (μm)	-0.5 to -2.5					
pixel size (Å)	0.435					
micrographs collected	9169					
initial particles	1,043,033					
final particles	51,750	156,677	43,431	37,530	37,885	31,453
symmetry	C1	C1	C1	C1	C1	C1
map resolution (Å) at 0.143 FSC threshold	3.24	3.38	3.33	3.24	3.22	3.26
Model composition						
non-hydrogen atoms	38021	37900	37776	38131	37880	38006
protein residues	4820	4805	4792	4837	4819	4823
nucleotides	12	11	10	12	12	12
Refinement						
map–model CC	0.80	0.77	0.75	0.81	0.81	0.82
RMSD bond lengths (Å)	0.005	0.005	0.005	0.005	0.006	0.006
RMSD bond angles (°)	1.115	1.117	1.100	1.095	1.156	1.142
Validation						
MolProbity score	1.09	1.06	1.06	1.07	1.14	1.10
clash score	2.98	2.74	2.74	2.78	3.48	3.05
Cβ deviation (%)	0.00	0.00	0.00	0.00	0.00	0.00
rotamer outliers (%)	0.00	0.00	0.02	0.00	0.00	0.00
Ramachandran favored (%)	99.94	99.73	99.75	99.88	99.77	99.50
Ramachandran disallowed (%)	0.00	0.00	0.00	0.00	0.00	0.00

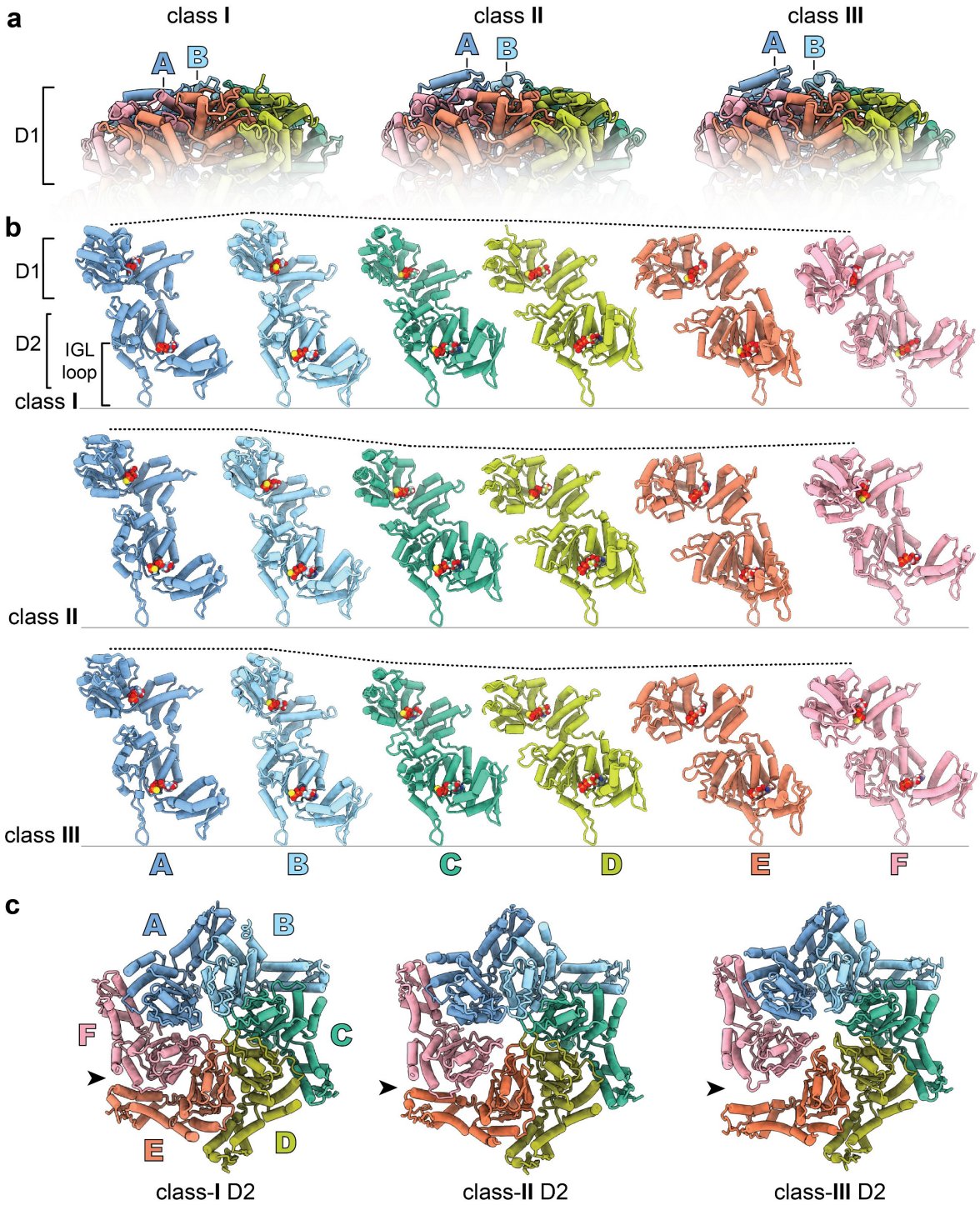
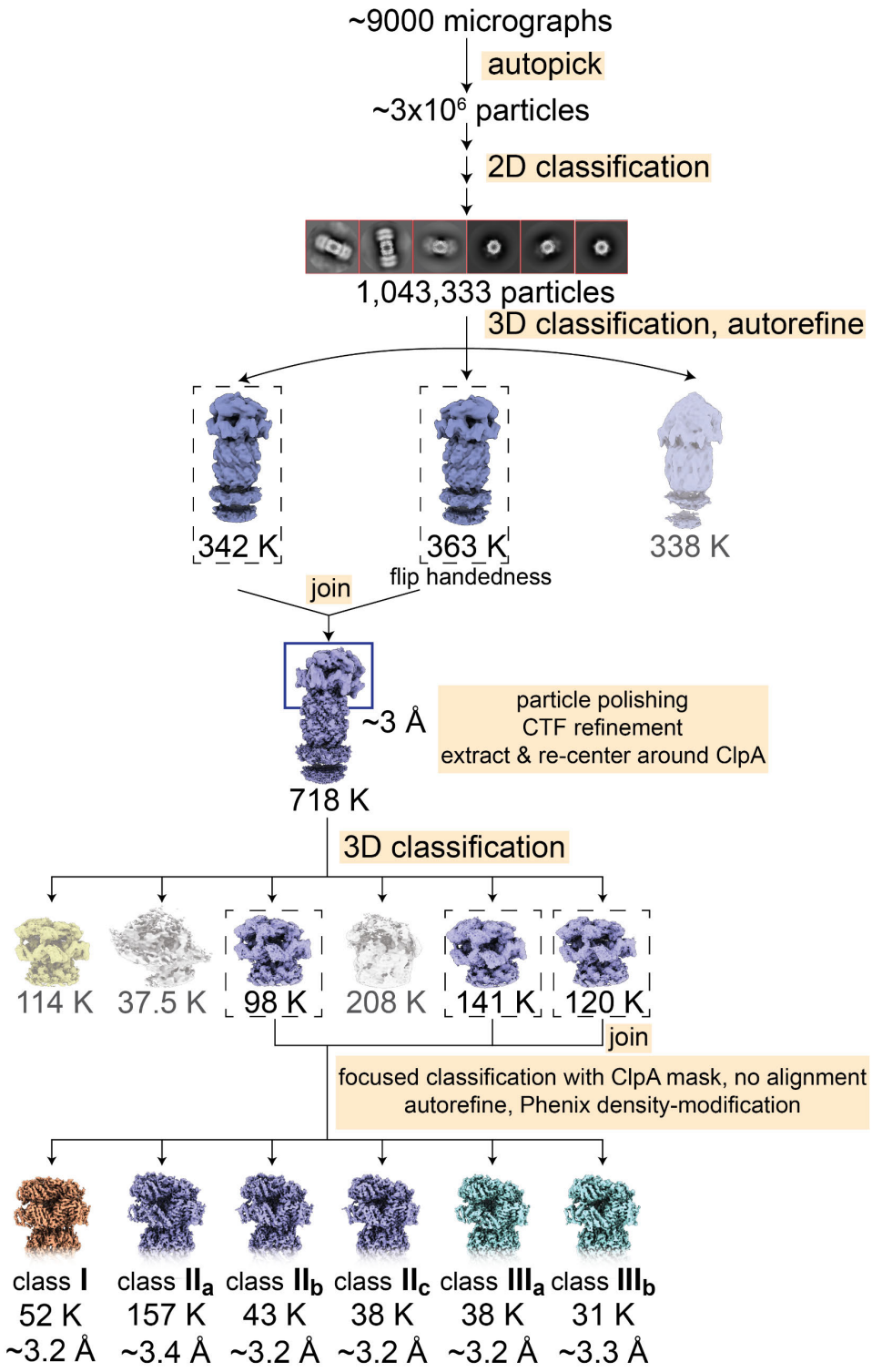
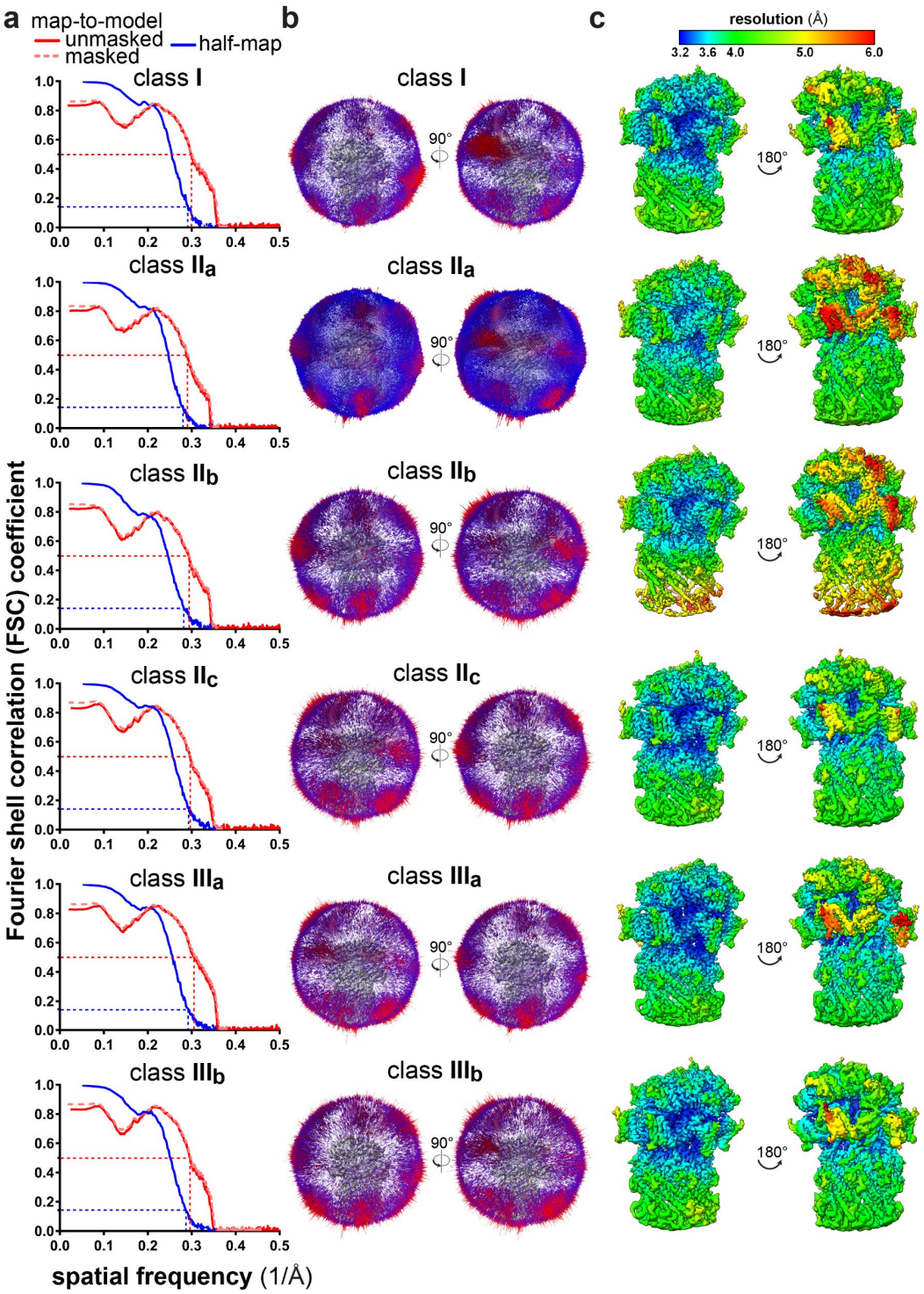


Figure 3.2 | Conformational differences in ClpA subunits and hexamers.
a, Atomic models of class **I**, **II**$_c$, and **III**$_a$ D1 rings, showing position of subunits A and B.
b, Individual subunits of class **I**, **II**$_c$, and **III**$_a$ atomic models. The dashed line indicates the relative height of each subunit, following alignment to the bottom of the IGL loop.
c, D2 ring rigid-body interface between subunits E and F of classes **I**, **II**$_c$, and **III**$_a$. In class **III**$_a$, the small AAA+ domain of subunit E in the D2 ring swings out and loses contact (arrow) with the neighboring large D2 AAA+ domain of subunit F.



Extended Data Figure 3.1 | Cryo-EM data-processing workflow diagram.
EM micrographs containing doubly-capped ClpAP complexes (two ClpA hexamers bound to one ClpP 14-mer) were processed in RELION-3. The final 3D classes were refined using density-modification in PHENIX.

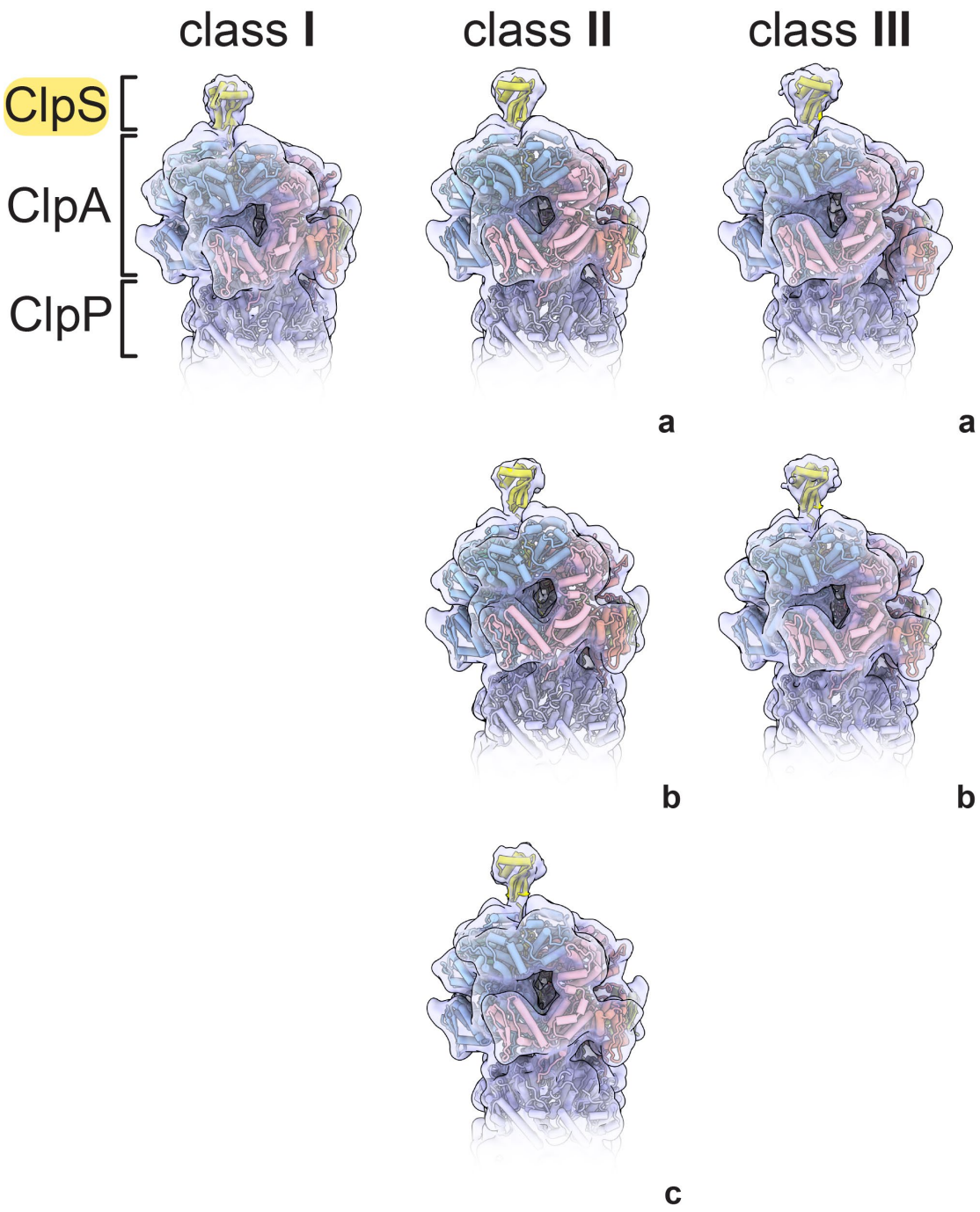


Extended Data Figure 3.2 | Cryo-EM validation.

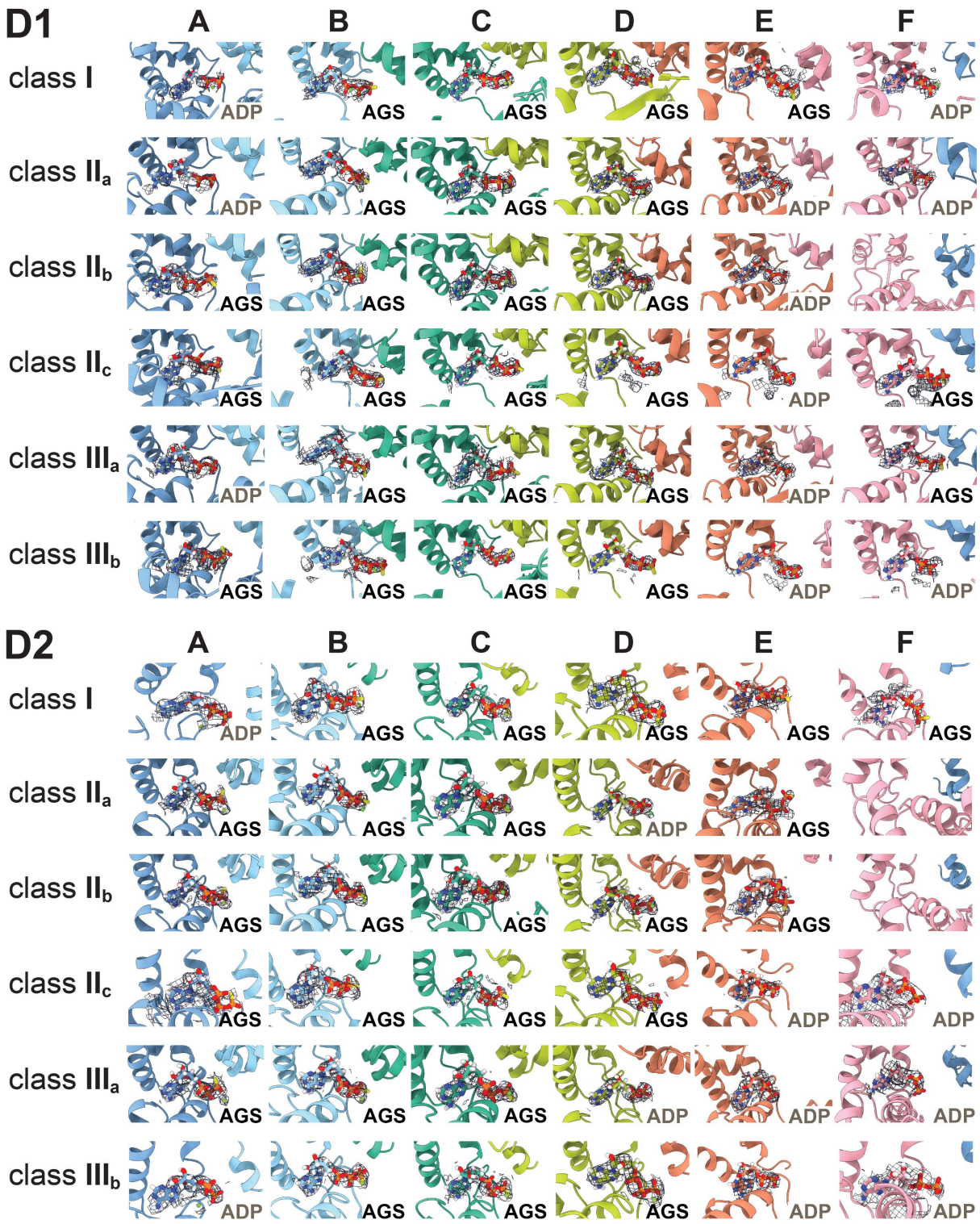
a, Fourier Shell Correlation (FSC) plots of half maps (shown in blue) or model-map (shown in red) resolution. The dashed lines indicate the cut-off values at FSC=0.5 (model-map) or FSC=0.143 (half-map).

b, Euler angle distribution plots of the particles used in the final reconstruction of class **I**, **II**$_{a\text{-}c}$, and **III**$_{a,b}$ structures.

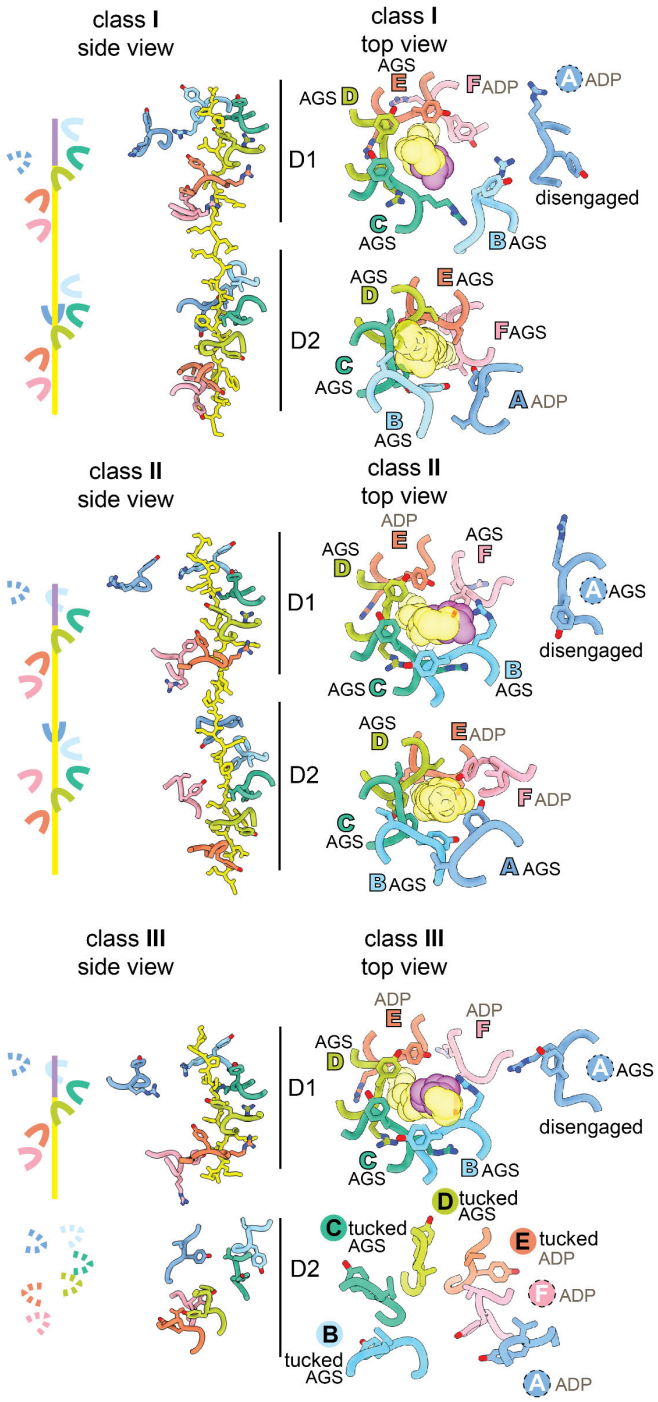
c, Local resolution maps of final reconstructions, colored according to RELION-3 calculations.



Extended Data Figure 3.3 | ClpS docking to cryo-EM maps.
The ClpS core domain (PDB 3O2B; res. 27-106) was docked to final reconstructions of class **I**, $\mathbf{II}_{\mathbf{a\text{-}c}}$, and $\mathbf{III}_{\mathbf{a,b}}$ that were low-pass filtered to 10 Å.



Extended Data Figure 3.4 | Nucleotide occupancy.
EM density (shown as mesh) of nucleotides in each D1 and D2 binding site of class **I**, **II**$_{a\text{-}c}$, and **III**$_{a,b}$ structures. Nucleotide density is not observed in the D1 site of F subunit in **II**$_b$ or the D2 site of the F subunit of **II**$_a$ or **II**$_b$.



Extended Data Figure 3.5 | Pore-1 loop interactions with ClpS NTE.
The leftmost panel in each structure is a diagram of pore-1 loops in the D1 and D2 rings, where solid lines represent the presence of NTE contacts and dashed lines represent their absence. Middle panel shows lateral views of contacts between ClpA pore-1 loops and the ClpS NTE in class **I**, **II$_c$**, and **III$_b$** atomic models. The rightmost panel in each structure is a closer view of the pore-1 loops with the NTE shown as transparent spheres and the corresponding nucleotide from each subunit. Labels in colored text denote NTE engagement; the dotted circle denotes lack of NTE engagement, with Tyr^{540} pointing towards the channel; labels in black text indicate the tucked conformation (Tyr^{540} away from the channel).

Following single-particle cryo-EM analyses (**Extended Data Figure 3.1**), three-dimensional (3D) classification and reconstruction using RELION-3 yielded six density maps (3.22–3.38 Å), representing three general structural classes (**I**, **II**, and **III**) with the latter classes being subdivided into $\mathbf{II_a}$/$\mathbf{II_b}$/$\mathbf{II_c}$ or $\mathbf{III_a}$/$\mathbf{III_b}$ subclasses (**Table 3.1**; **Extended Data Figure 3.2**). In low-pass filtered maps, the ClpS core domain (res. 27–106) could be docked into each of the six maps (**Extended Data Figure 3.3**). In unfiltered maps, there was good density for all or part of the NTE of ClpS, for the D1 and D2 rings of ClpA, and for ClpP (**Figure 3.1c**). There was no substantial density for the core domain of ClpS, the N-domains of ClpA, or YLFVQELA-GFP, suggesting that these domains/proteins are not present in fixed conformations relative to the remaining parts of the complex or are potentially absent (YLFVQELA-GFP).

In our structures, the six subunits of the ClpA hexamer, which we label A through F (clockwise direction with subunit F at the bottom), formed a shallow spiral, as expected from prior cryo-EM structures (Lopez et al., 2020). Six flexible IGL loops (res. 610–628) in each ClpA hexamer docked into clefts in the heptameric ring of ClpP, leaving one empty cleft (**Figure 3.1d**). Differences between classes **I**, **II**, and **III** include the relative positions of subunits in the ClpA spiral, density for the ClpS NTE in the ClpA channel, and changes within individual ClpA subunits (**Figure 3.1e**,**f** and **3.2**). For example, density for the ClpS NTE was present in both the D1 and D2 rings of ClpA in classes **I** and **II**, but only in the D1 ring of class **III** (**Figure 3.1f**). In classes **II** and **III**, the relative height of ClpA subunits in the spiral was A (highest) > B > C > D > E > F (lowest), whereas in class-**I** subunit B was higher than subunit A, resulting in a more planar D1 ring (**Figure 3.2a**,**b**). Additionally, in the D2 ring of class-**III** structures, the small AAA+ domain of subunit E swings outward, breaking the rigid-body interface with its large AAA+ domain neighbor from subunit F (**Figure 3.2c**). By contrast, structures of ClpAP with RepA-GFP and ATPγS (Lopez et al., 2020) did not display this feature, suggesting that it arises as a consequence of ClpS binding. The subclasses ($\mathbf{II_a}$/$\mathbf{II_b}$/$\mathbf{II_c}$ or $\mathbf{III_a}$/$\mathbf{III_b}$) differed from each other largely in the detailed interactions between

ClpA and the ClpS NTE, the visibility of individual NTE residues, and the nucleotide occupancy of each ATPase site (ATPγS, ADP, or empty) (**Extended Data Figure 3.4**,**3.5**).

Binding of the ClpS NTE within the axial channel of ClpA

Each of our structures contained clear main-chain and side-chain density corresponding to all or part of the ClpS NTE (res. 2–26) in the ClpA channel (**Figure 3.3a**). The register of this ClpS peptide is very similar in each structure, with the C-terminal portion of the NTE (Pro^{24}–Pro^{25}–Ser^{26}) near the top of the ClpA channel, and the N-terminal portion near the bottom of the channel in classes **I** and **II**. The ClpA channel is lined by the D1 KYR pore-1 loops (res. 258–260) and pore-2 loops (res. 292–302) and by the D2 GYVG pore-1 loops (res. 539–542) and pore-2 loops (res. 526–531). Pore-1 loops of AAA+ unfoldases and protein-remodeling machines contain a key, conserved aromatic side chain (usually tyrosine; underlined in KYR and GYVG) that contacts the substrate polypeptide in the channel and functions in the binding, unfolding, and translocation of target proteins (Schlieker et al., 2004; Weibezahn et al., 2004; Hinnerwisch et al., 2005a; Martin et al., 2008; Doyle et al., 2012; Iosefson et al., 2015; Lopez et al., 2020; Zuromski et al., 2021). The ClpS NTE was bound by many KYR and GYVG pore-1 loops and also by the D1 pore-2 loops of ClpA. Neighboring pore-1 loops interacted with two-residue segments of the NTE, as observed for substrate polypeptides bound to multiple AAA+ unfoldases and protein-remodeling machines (Puchades et al., 2020).

Despite this overall resemblance to substrate engagement, there were deviations in individual pore-1 loop interactions from those in prior structures of ClpAP and some Hsp100 family members. For example, the D2 GYVG pore-1 loops of all six ClpA subunits contacted the ClpS NTE in classes **I** and $\mathbf{II_c}$ (**Extended Data Figure 3.5**), whereas previous ClpAP structures and subclasses $\mathbf{II_a}$ and $\mathbf{II_b}$ show four or five engaged GYVG loops (**Extended Data Figure 3.6b**) (Lopez et al., 2020). The configuration of pore-1 loops in classes **I** and $\mathbf{II_c}$ was also different from

an extended Hsp104•casein structure in which loops from both the top and bottom AAA+ rings of all six protomers contact substrate in a split 'lock-washer' conformation (Gates et al., 2017). In classes **I** and **II$_c$**, we observed five bound and one unbound pore-1 loops in D1 and six bound pore-1 loops in D2, an arrangement found in the high-affinity state of *Mycobacterium tuberculosis* ClpB (Yu et al., 2018). In many AAA+ structures, only the pore loops of ATP-bound subunits contact substrate (Puchades et al., 2020). By contrast and as reported for ClpAP•substrate complexes (Lopez et al., 2020), the pattern of engaged vs. disengaged pore loops in our structures did not strictly correlate with the nucleotide present in the corresponding ATPase active site (**Extended Data Figure 3.4**,**3.5**). For instance, ADP is bound to the class-**II$_c$** D2 nucleotide sites in subunits E and F, but the GYVG loops from these domains contact the NTE. The presence of 11 engaged pore-1 loops (five D1 and six D2) likely contributes to the high-affinity of ClpAPS•N-degron complexes assembled in ATPγS (Román-Hernández et al., 2011).

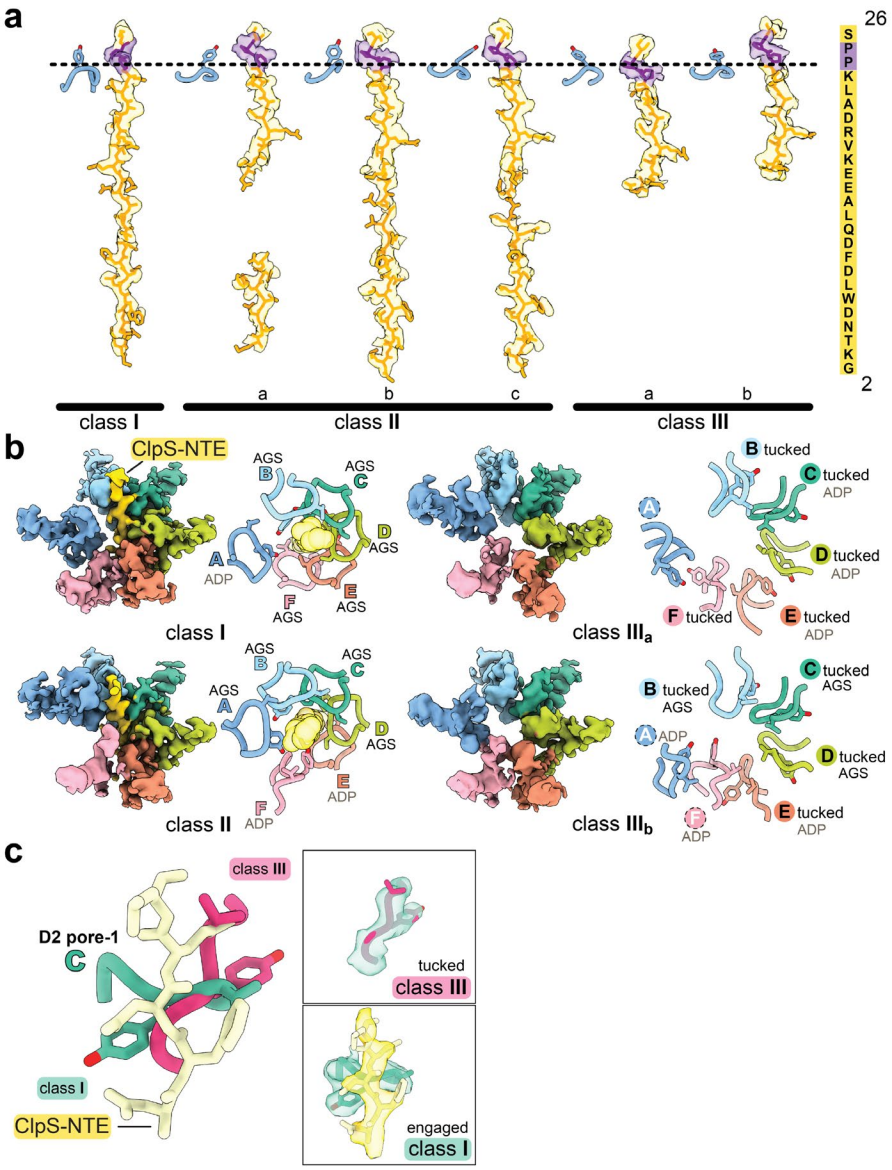
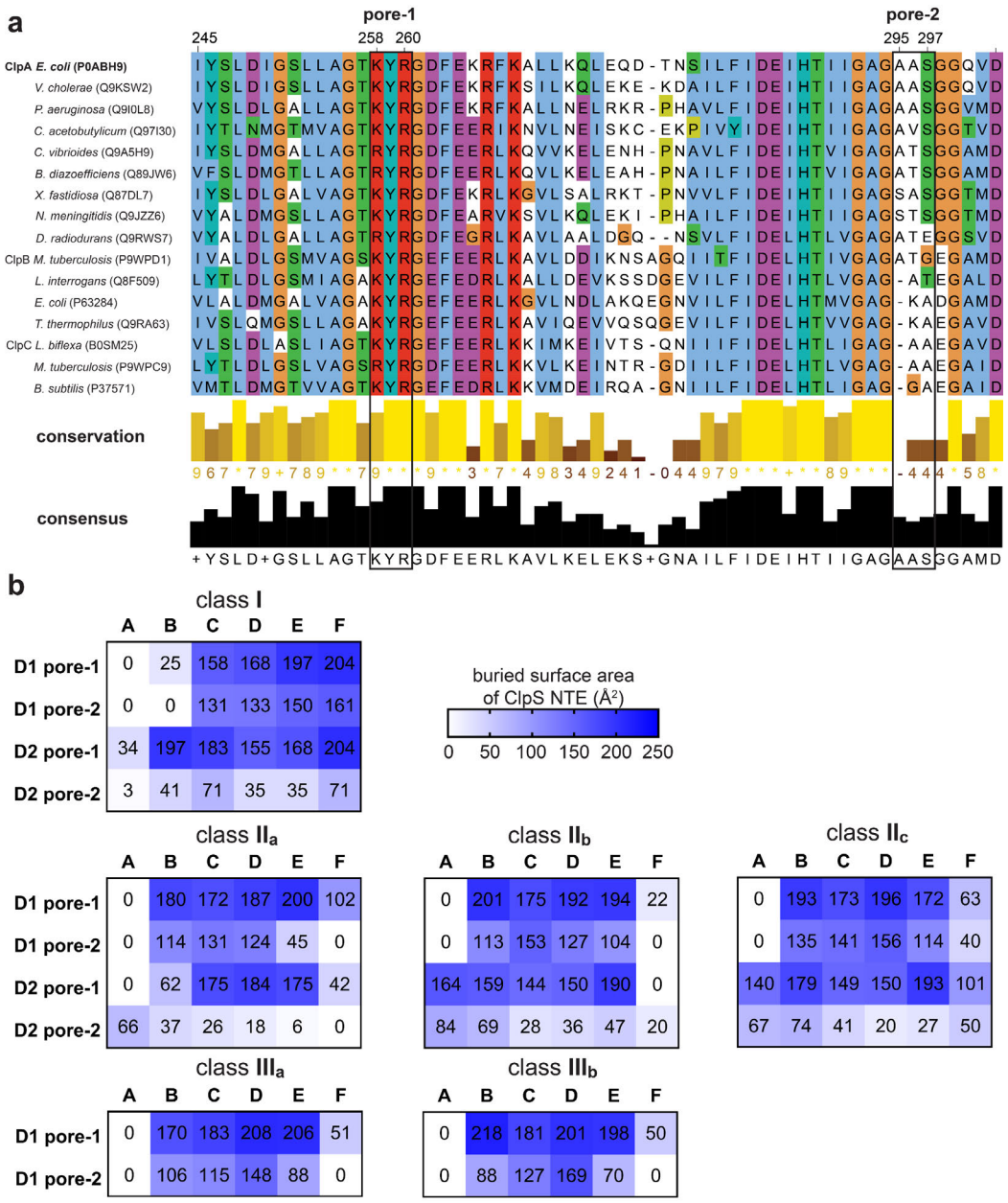


Figure. 3.3 | Conformations of the ClpS NTE and D2 pore-1 loops of ClpA.
a, Density of the ClpS NTE (transparent surface with modeled residues in sticks) in the ClpA axial channel in classes **I**, **II**, and **III**. Pro^{24} and Pro^{25} (colored purple) are part of the junction sequence between the NTE and the ClpS core domain. The ClpS NTE sequence is shown on the right. The D1 pore-1 loop of subunit A is shown as a reference point for the top of ClpA.
b, D2 ring ClpA pore-1 loops and the ClpS NTE in classes **I**, $\mathbf{II_c}$, and $\mathbf{III_{a,b}}$. The left panel in each structure depicts cryo-EM density for ClpA res. 528–555 and the ClpS NTE res. 2–15. The right panels are a zoomed-in view of the pore-1 loops (res. 538–542) and the NTE (transparent spheres), which is absent in class-**III**, in the atomic models. Subunit labels indicate nucleotide and interaction with the ClpS NTE. Labels in colored text denote NTE engagement; the dotted circle denotes lack of NTE engagement, with Tyr^{540} pointing towards the channel; labels in black text indicate the tucked conformation (Tyr^{540} away from the channel). **c**, Atomic models (sticks) and density (transparent surfaces in boxed area) of the subunit C D2 pore-1 loop in classes **I** (green) or $\mathbf{III_b}$ (pink) and the class-**I** ClpS NTE (yellow). In class **III**, density for the NTE is not observed.



Extended Data Figure. 3.6 | Comparison of pore-1 and pore-2 loop sequences and contacts to ClpS NTE.

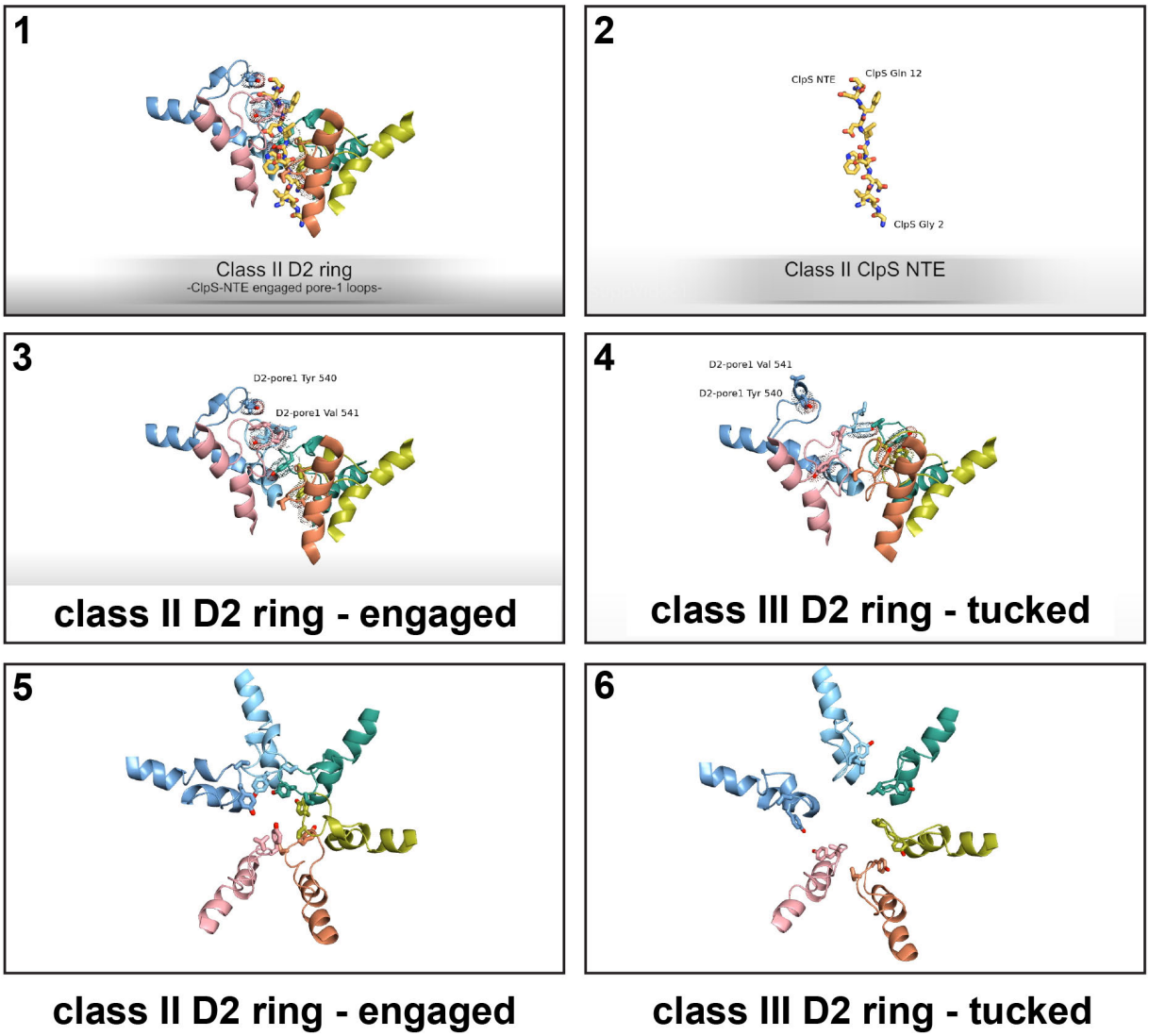
a, Multiple sequence alignment of ClpABC family members, corresponding to pore-1 (res. 258–260, *E. coli* ClpA) and D1 pore-2 loops of *E. coli* ClpA (res. 292–302), created using MUSCLE alignment. UniProt accession numbers are listed in parentheses. The alignment at each position is colored according to ClustalX (orange=Gly, blue=hydrophobic, green=polar, magenta/purple=positive charge, white=unconserved). Conservation scores are calculated in Jalview from the amino acid properties in the alignment. Conserved columns that have the highest conservation score are indicated by '*' symbols (corresponding to a numeric score of 11), followed next by mutations that conserve all physico-chemical properties, indicated by '+' symbols. Gaps are indicated by '-', and the lowest conservation score is zero.

b, Buried Surface Area (BSA) of ClpS NTE. Contacts between ClpS NTE and pore-1 or pore-2 loops in the D1 and D2 rings of classes **I**, **II$_{a-c}$**, and **III$_{a,b}$** were evaluated using PISA. Raw BSA values are provided in each box that correspond to coloring by the heat map scale.

Pore-1 loops of D2 ring rotate out of the channel to alter polypeptide contacts

In classes **I** and **II**, residues 2–15 of the ClpS NTE were built into density in the D2 portion of the channel, but this NTE region was not visible in class **III**, presumably as a consequence of its conformational heterogeneity. We infer that these NTE residues are within D2, as the more C-terminal NTE segment (res. 16–26) is bound by the D1 ring of class **III** in the same manner as in classes **I** and **II**. Thus, the two AAA+ rings of ClpA can differ in their engagement with the NTE, a feature not observed in substrate-bound ClpAP structures (Lopez et al., 2020).

This absence of density for the N-terminal portion of the NTE in class **III** correlated with distinct structural features within the axial channel. Most surprisingly, the D2 pore-1 loops in class **III** were rotated ~90° compared to their orientation in classes **I** and **II**, and the key Tyr^{540} side chains were tucked-in and turned away from the axial channel (**Figure 3.3b,c** and **Supplementary Video 1**). In both class-**III** subclasses, at least four of the six pore 1 loops were convincingly in this new tucked conformation. In many AAA+ unfoldases and protein-remodeling machines, one or two pore-1 loops, usually at the top and bottom of the spiral, are disengaged from the substrate polypeptide as a result of translational displacement of the corresponding subunit(s) (Deville et al., 2017; Gates et al., 2017; Puchades et al., 2017, 2019, 2020; Ripstein et al., 2017, 2020; Yu et al., 2018; Cooney et al., 2019; Dong et al., 2019; Han et al., 2019, 2020; Lo et al., 2019; Twomey et al., 2019; Fei et al., 2020a, 2020b; Lopez et al., 2020). This ‘canonical’ disengaged state of pore-1 loops in one or two subunits is very different than the tucked and rotated orientations of the class-**III** D2 pore-1 loops, in which no interactions with the polypeptide in the channel were present in the D2 ring. Pore-1 tyrosine contacts with the polypeptide within the AAA+ channel are considered essential for substrate binding and translocation. Thus, rotation of most (or all) Tyr^{540} side chains in the class-**III** D2 ring is sufficient to explain the lack of initial engagement of the N-terminal segment of the NTE and/or loss of binding that may occur during ClpS-assisted degradation of N-degron substrates (see *Discussion*).



Supplementary Video 1. Altered pore-1 loop conformations in the ClpA D2 ring.
Comparison of class-**II$_c$** (pore-1 loops are engaged to ClpS NTE) and class-**III$_a$** atomic models (showing tucked pore loops). Frames are labeled in the upper left corner.

Three additional features of the class-**III** D2 ring are noteworthy. Coincident with the pore-1-loop rotation, the ClpA channel in the D2 ring of class **III** was wider than in classes **I** and **II** (**Figure 3.2c**). Second, as noted above, the D2 rigid-body interface between the small AAA+ domain of subunit E and its neighboring large AAA+ domain in subunit F was broken in class **III**. This rearrangement may facilitate the accompanying conformational changes that result in loss of NTE contacts by the D2 pore-1 loops. Finally, the D2 ring contained ADP in three adjacent subunits in class **III**, whereas classes **I** and **II** contained no more than two ADPs in the D2 ring (**Extended**

Data Figure 3.4). Thus, ClpA has the ability to bind all of a polypeptide in the axial channel tightly using pore loops in both rings or by disrupting coordinated activity of the D1 and D2 rings, to specifically bind only the C-terminal portion of this sequence within the D1 ring.

Pore-2 loops in the D1 ring form a second network of NTE-engaging contacts

In addition to the pore-1 loop interactions described above, our structures show that at least four pore-2 loops (res. 292–302) in the D1 ring of ClpA contacted the ClpS NTE (**Figure 3.4**). In each subunit, these pore-2 contacts were positioned below the corresponding D1 KYR contacts and were offset by ~60°. The Ala^{295}-Ala^{296}-Ser^{297} tripeptide (AAS) at the tip of the D1 pore-2 loops contacted the opposing face of the ClpS NTE compared to the contacts made by the D1 pore-1 loops (compare orientation of D1 pore-2 loops on left vs. D1 pore-1 loops on right side of channel in **Figure 3.4a**). In contrast to the well-defined KYR motif in the D1 pore-1 loop, which is conserved among Hsp104/ClpABC protein-remodeling enzymes and contains the invariant aromatic residue present in all AAA+ unfoldases, the key residues and functions of the pore-2 loops have been poorly delineated to date (Puchades et al., 2020). Among ClpABC family members, the pore-2 loops are more variable in sequence and length (**Figure 3.4b** and **Extended Data Figure 3.6a**).

To quantify the extent of pore-1 vs. pore-2 loop interactions in the D1 ring, we calculated the buried surface area (BSA) of the ClpS–NTE interface with each class of pore loops using PISA (Krissinel and Henrick, 2007). Mirroring the pattern of D1 pore-1 loops bound to the ClpS NTE, multiple pore-2 loops made significant ClpS NTE interactions in all class **I**, **II**, and **III** structures (**Figure 3.4c**). The D1 pore-2 loops made substantially larger contributions to the interface with the ClpS NTE than the pore-2 loops of the D2 ring, as the buried surface area contributed by the D1 pore-2 loops was comparable to those from either the D1 KYR or D2 GYVG pore-1 loops (**Extended Data Figure 3.6b**). For example, in class $\mathbf{II_c}$, the BSA values for the D1 KYR loops range from 63 to 196 $Å^2$, the D1 pore-2 loops range from 40 to 156 $Å^2$, and D2 GYVG pore-1

loops range from 101 to 179 Å^2. In contrast, the D2 pore-2 loops only weakly contacted the NTE, as BSA values of these interactions range from 20 to 74 Å^2. Thus, pore loops in the D1 ring make a greater total number of NTE interactions than pore loops in the D2 ring. The extensive network of NTE-engaging residues in the D1 ring suggests that it has more specific polypeptide binding/recognition 'capacity' than the D2 ring, as predicted by biochemical studies (Hinnerwisch et al., 2005a; Zuromski et al., 2021).

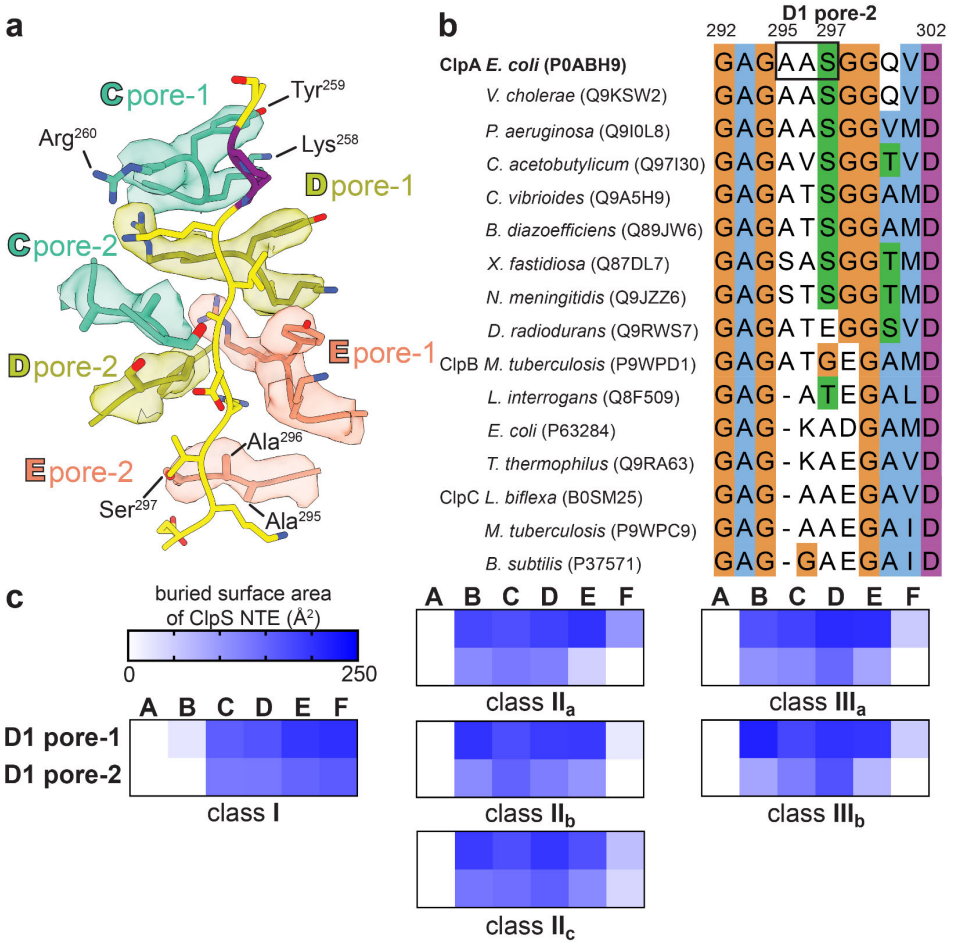


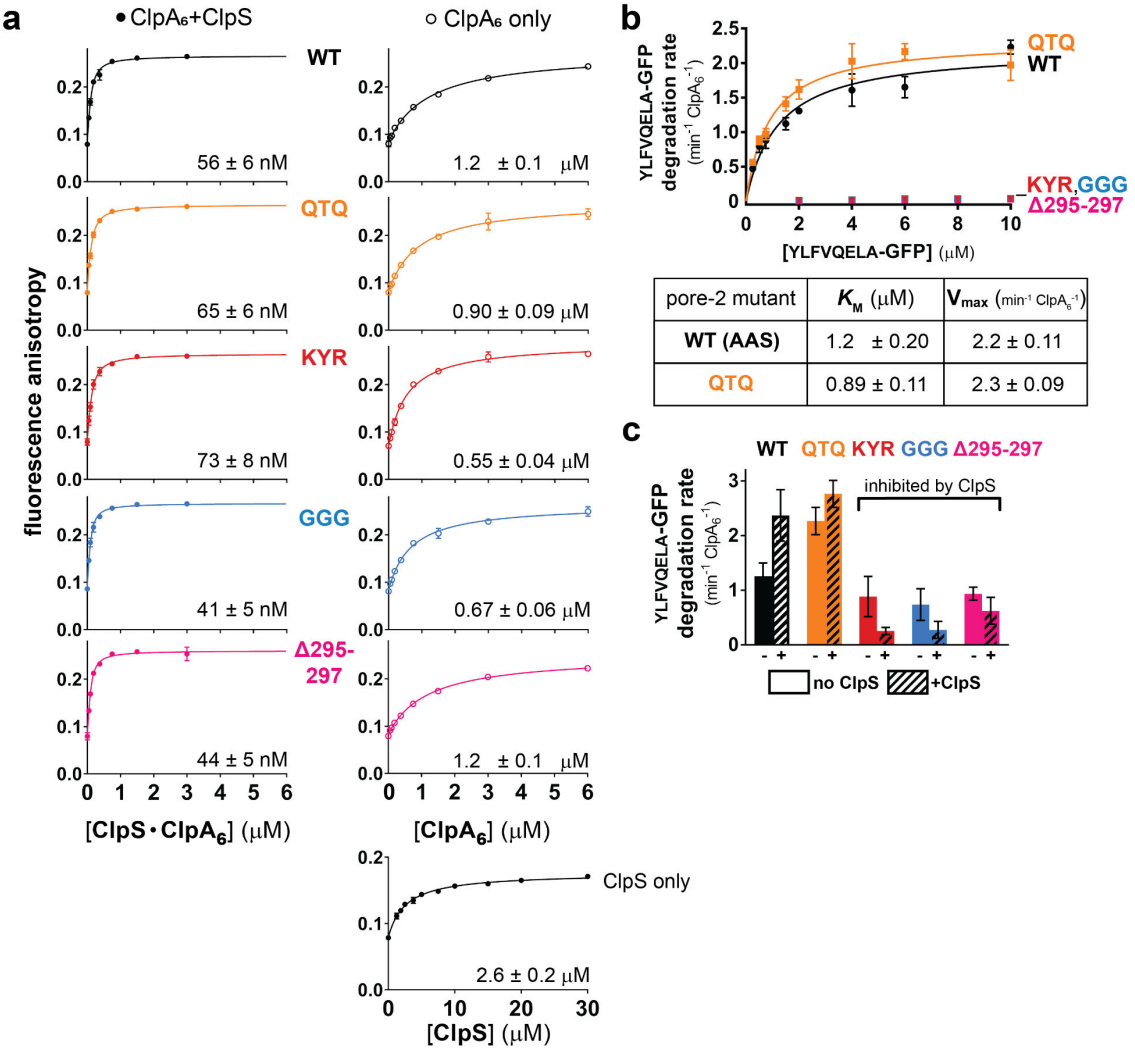
Figure. 3.4 | Interaction of ClpA pore-2 loops with the ClpS NTE in the D1 ring.
a, Three pairs of pore-1 (KYR, res. 258–260) and pore-2 (AAS, res. 295–297) loops in the ClpA D1 ring of class **I**, shown as sticks in representative subunit coloring, and the ClpS NTE (yellow and purple sticks). Cryo-EM density of each pore loop is shown with the respective transparent surfaces. **b**, Multiple sequence alignment of ClpABC family members corresponding to D1 pore-2 loops of *E. coli* ClpA (res. 292–302). UniProt accession numbers are listed in parentheses. The alignment (Edgar et al., 2004) at each position is colored according to ClustalX (orange=Gly, blue=hydrophobic, green=polar, magenta/purple=positive charge, white=unconserved). **c**, Buried surface area of the ClpS NTE by the pore-1 or pore-2 loops of the D1 ring in the atomic models of classes **I**, **II**, and **III**. See also Extended Data Fig. 3.6.

D1 pore-2 loops mediate substrate unfolding and mechanical remodeling of ClpS

To test the functional importance of the D1 pore-2 loops, we mutated the AAS sequence (res. 295–297) to increase bulkiness (QTQ), to mimic the pore-1 loop (KYR), to increase flexibility (GGG), or to delete this tripeptide (Δ295–297). As a defect in ClpS binding with these mutants was one reasonable hypothesis based on our structures, we first assayed assembly of ternary $ClpA_6$•ClpS•N-degron peptide complexes using fluorescence anisotropy (**Figure 3.5a**). Strikingly, all pore-2 loop variants maintained tight affinity for the ClpS•N-degron complexes and behaved similarly to wild-type ClpA (^{WT}ClpA) in the control experiment that monitored the binary affinity of ClpA to the N-degron peptide. We then used these variants to assay ClpAPS degradation of the N-degron substrate YLFVQELA-GFP (**Figure 3.5b**). Notably, all of the D1 pore-2-loop variants except QTQ were unable to degrade this substrate. These defects could arise from an inability to unfold or translocate YLFVQELA-GFP or from failure to transfer the YLFVQELA-GFP substrate from ClpS to ClpA.

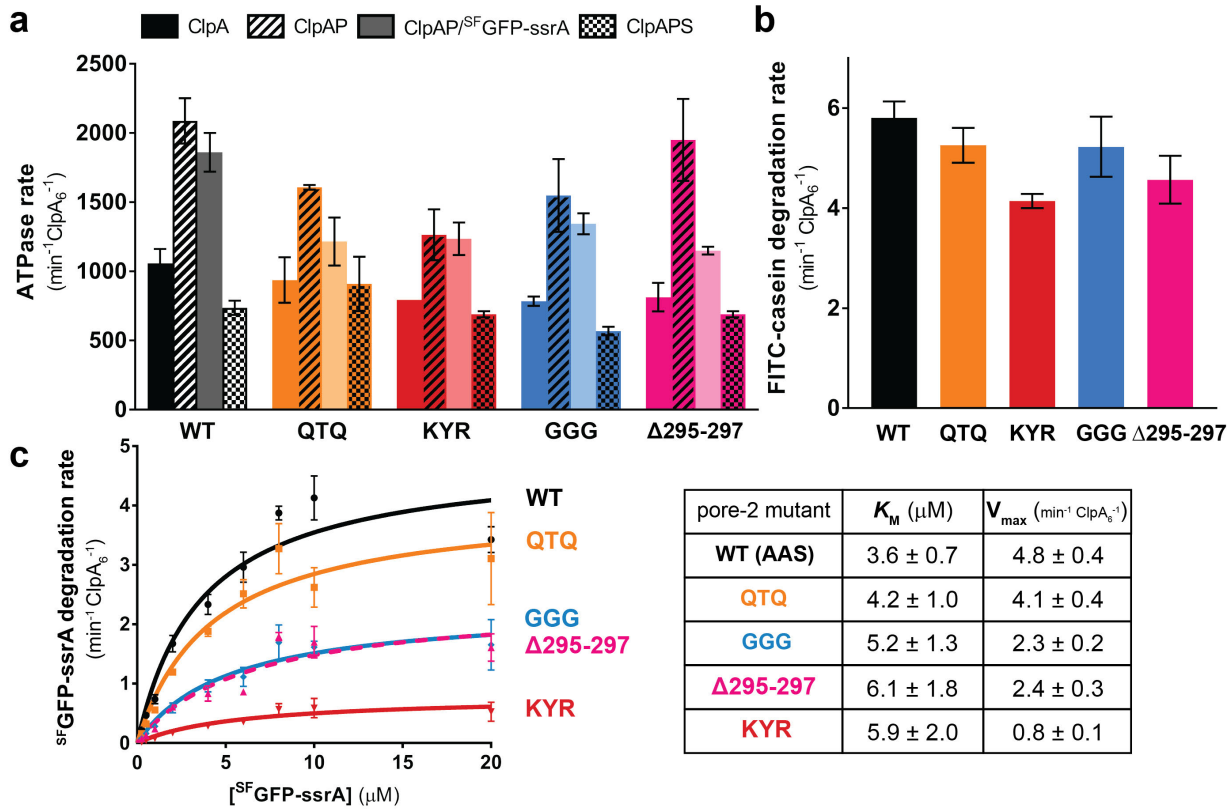
For each variant, we then determined the ATP-hydrolysis rate of ClpA alone and in the presence of ClpP, ClpS, and/or a directly recognized protein substrate. The ATPase rate serves as an indirect readout of functional ClpA assembly with its binding partners, which differentially modulate ATP hydrolysis by ClpA. For instance, ClpP binding stimulates the ATPase rate of ^{WT}ClpA ~two-fold, whereas ClpS suppresses the ATPase activity of ClpAP to a rate similar to that of ClpA alone (Hinnerwisch et al., 2005b; Hou et al., 2008). All of the pore-2 variants had basal ATPase rates comparable to ^{WT}ClpA and exhibited ATPase modulation by ClpP and ClpS that was generally similar to wild-type (**Extended Data Figure 3.7a**). Furthermore, in the presence of the super-folder GFP substrate (^{SF}GFP-ssrA), which does not require ClpS for recognition and degradation, the ATPase rate of each ClpAP variant (with the exception of ^{KYR}ClpAP) was moderately reduced during substrate processing, as expected from a previous study reporting ~20% suppression of ATP hydrolysis by GFP-ssrA (Kress et al., 2009). We conclude based on

these studies that our D1 pore-2 loop mutations do not grossly alter ClpA ATPase activity and are also unlikely to substantially change ClpA assembly with ClpP, ClpS, or ^{SF}GFP-ssrA.



pore-2 mutant	K_M (μM)	V_{max} (min^{-1} ClpA$_6^{-1}$)
WT (AAS)	1.2 ± 0.20	2.2 ± 0.11
QTQ	0.89 ± 0.11	2.3 ± 0.09

Figure. 3.5 | D1 pore-2 loops are critical for ClpS-mediated degradation.
a, Fluorescence anisotropy of ClpA pore-2 variants alone (ClpA$_6$, open circles), with an equimolar mixture with ClpS (+ClpS, filled circles), or ClpS only (bottom right panel), titrated in increasing concentrations against a fixed concentration of fluorescein-labeled N-degron peptide. Values are mean fluorescence anisotropy values from triplicates with error bars representing ± 1 S.D, and were fit to equations listed in the *Methods*. The K_d values are reported on the lower right on each isotherm, with (±) the standard error of nonlinear least-squares and R^2 values=0.99 for all fits. **b**, Kinetic analysis of YLFVQELA-GFP degradation by ClpA D1 pore-2 variants (see *Methods* for concentrations). Values are mean degradation rates (min^{-1} ClpA$_6^{-1}$) of triplicates with error bars representing ± 1 S.D. In the table, K_M and V_{max} values ± errors were obtained by non-linear least squares fitting to the Michaelis-Menten equation. Degradation rates of GGG, Δ295–297, and KYR could not be fit to the Michaelis-Menten equation. **c**, YLFVQELA-GFP (20 μM) degradation rates of ClpA D1 pore-2 variants (0.1 μM ClpA$_6$) and ClpP (0.2 μM ClpP$_{14}$), in the absence and presence of ClpS (0.6 μM ClpS). Summary data are mean degradation rates (min^{-1} ClpA$_6^{-1}$) of triplicates with error bars representing ± 1 S.D.



pore-2 mutant	K_M (μM)	V_{max} (min^{-1} ClpA$_6^{-1}$)
WT (AAS)	3.6 ± 0.7	4.8 ± 0.4
QTQ	4.2 ± 1.0	4.1 ± 0.4
GGG	5.2 ± 1.3	2.3 ± 0.2
Δ295-297	6.1 ± 1.8	2.4 ± 0.3
KYR	5.9 ± 2.0	0.8 ± 0.1

Extended Data Figure 7 | D1 pore-2 loop variants can support ATP hydrolysis and substrate unfolding, translocation, and degradation.
a, ATP hydrolysis rates of ClpA D1 pore-2 variants alone or in the presence of ClpP, ClpP and ^{SF}GFP-ssrA substrate, or ClpP and ClpS; see *Methods* for concentrations.
b, FITC-casein (20 μM) degradation rates of ClpA D1 pore-2 variants (0.2 μM $ClpA_6$) and ClpP (0.4 μM $ClpP_{14}$). Summary data are mean degradation rates (min^{-1} $ClpA_6^{-1}$) of triplicates with error bars representing ± 1 S.D.
c, Kinetic analysis of ^{SF}GFP-ssrA degradation by ClpA D1 pore-2 variants (0.2 μM $ClpA_6$, 0.4 μM $ClpP_{14}$). Values are mean degradation rates (min^{-1} $ClpA_6^{-1}$) of triplicates with error bars representing ± 1 S.D. In the table, K_M and V_{max} values ± errors were obtained by non-linear least squares fitting to the Michaelis-Menten equation.

Next, we assayed the ability of these D1 pore-2-loop variants to degrade FITC-casein, a molten-globule protein that does not require ClpS for recognition or robust ClpAP unfolding activity for degradation (Thompson et al., 1994). ^{KYR}ClpAP degraded FITC-casein ~30% slower than ^{WT}ClpAP, but the remaining D1 pore-2 variants degraded this substrate at roughly the wild-type rate (**Extended Data Figure 3.7b**), indicating that recognition and translocation of this substrate are not substantially affected by the D1 pore-2 loop mutations. We then assayed the effects of the ClpA D1 pore-2 loop mutations on the steady-state kinetics of ^{SF}GFP-ssrA degradation

(**Extended Data Figure 3.7c**). K_M values for degradation of this highly stable native substrate by ^{WT}ClpAP and the D1 pore-2 loop variants were within error, suggesting that the D1 pore-2 loops play little, if any, role in recognition of ^{SF}GFP-ssrA. V_{max} for ^{SF}GFP-ssrA degradation was unaffected by the QTQ mutation, reduced ~two-fold by the GGG and Δ295–297 mutations, and reduced ~six-fold for the KYR mutant. Based on these results, we conclude that the D1 pore-2 loops can promote, but are not essential for, a reaction step after initial substrate recognition, presumably GFP unfolding, which is rate limiting for degradation (Singh et al., 2000). Importantly however, these partial defects in unfolding by the GGG, Δ295–297, and KYR variants are insufficient to explain the complete inability of these mutants to degrade YLFVQELA-GFP when delivered by ClpS.

In comparison to FITC-casein and ^{SF}GFP-ssrA, which are directly recognized by ClpAP, degradation of ClpS-dependent substrates require an additional protein-remodeling step. That is, ClpA must remodel ClpS, to allow substrate transfer to ClpA, and then unfold the N-degron substrate. Concurrently, ClpS reduces the ClpA ATPase rate, which in turn slows unfolding and translocation (Dougan et al., 2002; Hou et al., 2008; De Donatis et al., 2010; Román-Hernández et al., 2011; Rivera-Rivera et al., 2014; Torres-Delgado et al., 2020). Therefore, ClpS should inhibit N-degron substrate degradation by the ClpA D1 pore-2 loop variants that we infer lack sufficient unfolding activity to remodel ClpS and transfer the substrate from ClpS to ClpA. We tested this hypothesis by measuring the degradation rates of YLFVQELA-GFP in the absence and presence of ClpS (**Figure 3.5c**). Although recognition of N-end-rule substrates by ClpAP alone is intrinsically weak and normally enhanced by ClpS (Wang et al., 2007), the addition of ClpS hindered YLFVQELA-GFP degradation by the KYR, GGG, and Δ295–297 variants, but not by ^{WT}ClpA or the QTQ variant, as predicted if the D1 pore-2 mutants are specifically defective in a ClpS remodeling step required for efficient N-degron substrate degradation.

In summary, these data suggest that the sequence identity of the AAS tripeptide (res. 295–297) alone is not critical for D1 pore-2 loop activity, as substituting these residues with QTQ had little effect on ATP hydrolysis and degradation of all substrates tested (**Figure 3.5b**,**c** and **Extended Data Figure 3.7**). Instead, changing the chemical/conformational properties of this loop by altering charge/aromaticity (KYR) or flexibility (GGG and Δ295–297) had more profound effects. The severe defects in ClpAPS degradation conferred by the deleterious D1 pore-2 mutations but unchanged ClpS•N-degron assembly support the conclusion that D1 pore-2 loops assist in mechanical work needed to transfer the N-degron substrate from the adaptor to the protease (and perhaps also for subsequent reaction steps) but are not required for adaptor/substrate docking with ClpAP.

DISCUSSION

The ClpS NTE is a "degron mimic"

Our ClpAP-ClpS structures, taken with previous ClpAP structures and those of additional AAA+ family members, illustrate the variety of functional conformations AAA+ unfoldases can adopt to perform their biological functions. Importantly in all our structures, interactions between the ClpS NTE and pore loops in the ClpA channel mimic contacts observed with a polypeptide segment of the protein substrate in prior ClpA structures (Lopez et al., 2020). Specifically, the conserved tyrosines from adjacent pore-1 loops in the D1 (K<u>Y</u>R) and D2 (G<u>Y</u>VG) ring contact every second residue of the NTE polypeptide (**Extended Data Figure 3.5**), with additional contacts mediated by the pore-2 loops of the D1 ring. Thus, in addition to its interaction with the ClpA N-domain, ClpS uses its NTE to dock tightly with the ClpA channel during substrate delivery.

Our ClpAPS structures were assembled in ATPγS, which ClpA does not hydrolyze, demonstrating that binding of the entire ClpS NTE within both ClpA rings does not require hydrolysis-dependent power strokes. Together with biochemical studies and structures of substrate complexes with

ATPγS-bound ClpA (Hoskins et al., 1998, 2000; Lopez et al., 2020), these results suggest that any polypeptide in an unfolded/misfolded protein could passively enter an open ClpA channel, enabling ClpAP to function broadly in general protein quality control. Indeed, most of the ClpS NTE sequence is poorly conserved among orthologs and can be changed without compromising delivery of N-end-rule substrates (Hou et al., 2008), suggesting that ClpA can engage many different sequences. By contrast, the full axial channel of the ClpXP protease is blocked by a pore-2 loop prior to initiation of unfolding and translocation, probably limiting binding to proteins bearing highly specific ClpX degrons (Fei et al., 2020b).

Our structures also reveal that ClpA pore loops bind and engage the ClpS NTE, and thus can apply mechanical force to the ClpS core domain during the N-degron delivery process. Previous studies demonstrate that the ClpS NTE enters the ClpA axial channel during assembly of delivery complexes and also can independently function as a degron for ClpAP, providing biochemical evidence that the ClpA pore loops can 'pull' on the NTE to remodel ClpS (Román-Hernández et al., 2011; Rivera-Rivera et al., 2014). Blocking the ClpS NTE from entering the channel inhibits ClpS-assisted substrate degradation, reinforcing the importance of ClpA 'pulling' on the ClpS NTE during N-degron delivery (Rivera-Rivera et al., 2014). The degron-like binding of the NTE provides a structural basis for the delivery mechanism depicted in **Figure 3.6**, in which ClpA pore loops engage the NTE and power strokes resulting from ATP hydrolysis transmit force to mechanically remodel ClpS and thereby promote transfer of the N-end-rule substrate from ClpS to ClpAP for degradation.

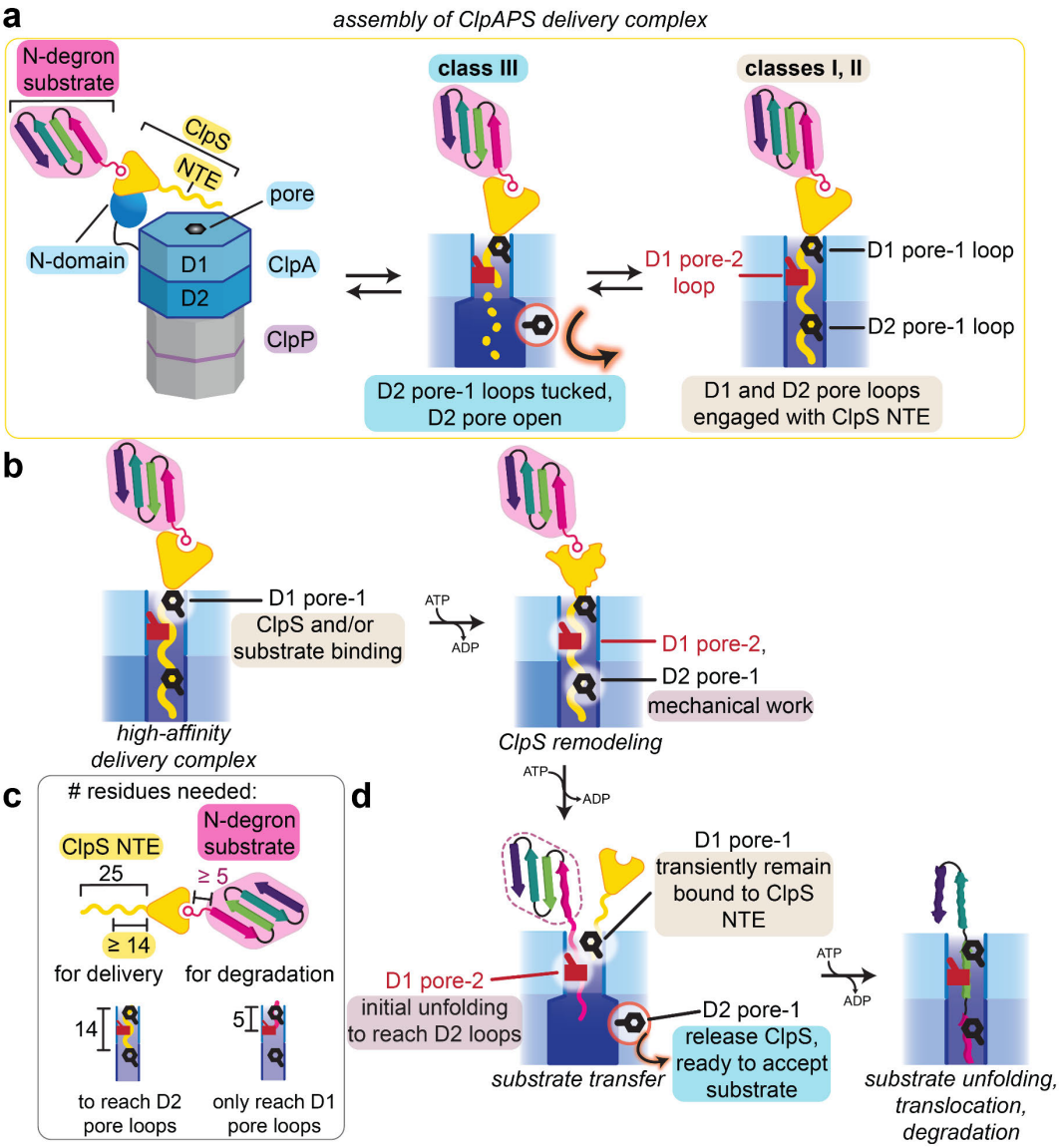


Figure 3.6 | Models of ClpS-mediated degradation by ClpAP.
a, Summary of observed interactions of ClpS NTE and pore-1 and pore-2 loops in D1 and D2 from cryo-EM structures. In a dynamic equilibrium of low-affinity (not observed in this study) and high-affinity delivery complexes (classes **I, II,** and **III**), pore-1 loops in the D2 ring (*i*) are tucked-in and turned away from the axial channel (class **III**), correlated with a loss of observed ClpS NTE density, as indicated by the dotted line or (*ii*) contact the ClpS NTE (classes **I**, **II**).
b, Pore loop functions in ClpS remodeling. Pore-1 loops, especially in D1, are required for ClpS and substrate binding to form delivery complexes. ATP hydrolysis powers ClpS remodeling, allowing D1 pore-2 and D2 pore-1 loops to translocate and tug on the ClpS NTE to promote substrate transfer and degradation. **c**, Summary from biochemical studies (Erbse et al., 2006; Wang et al., 2008b) of number of residues required for ClpS NTE and linker between N-end residue and folded domain in N-degron substrate. **d**, Proposed function of pore loops during substrate transfer. Following ClpS remodeling, D2 pore-1 loops may release the ClpS NTE, allowing for unfolding and translocation of the N-degron substrate to proceed from the D1 pore-2 loops.

Despite the degron-like interactions of the NTE with ClpA, ClpS is not degraded (Dougan et al., 2002; Román-Hernández et al., 2011). Interestingly, mutation of Pro^{24}-Pro^{25} to Ala^{24}-Ala^{25} near the ClpS NTE-core junction generates a ClpS variant that can be degraded by ClpAP (Rivera-Rivera, 2015; Zuromski et al., 2021). In our structures, Pro^{24}-Pro^{25} binds near the top of the ClpA channel, and the ClpS core domain is flexibly positioned directly above the pore. Although our structures only capture the initial docking of ClpS with ClpA, we propose that during subsequent stages of ClpS delivery, ClpA pore loops may not grip Pro^{24}-Pro^{25} strongly enough to fully unfold ClpS, leading to 'back-slipping' in the channel and thus ClpS release. Such slipping is likely a consequence of the unique chemical properties of proline, which lacks an amide hydrogen and cannot form the extended peptide conformation adopted by the rest of the NTE in our structures and by substrate polypeptides in the channels of many other AAA+ unfoldases (Puchades et al., 2020). ClpXP translocates poly-proline at an even slower rate and with a higher cost of ATP hydrolysis cycles than poly-glycine, another homopolymer that leads to pore-loop 'slipping' (Barkow et al., 2009; Bell et al., 2019). Other types of 'slippery' sequences adjacent to folded domains have been shown to cause release of truncated degradation products by a number of AAA+ proteases (Lin and Ghosh, 1996; Levitskaya et al., 1997; Sharipo et al., 2001; Hoyt et al., 2006; Daskalogianni et al., 2008; Kraut et al., 2012; Kraut, 2013; Too et al., 2013; Vass and Chien, 2013; Bell et al., 2019). Partial ClpAP processing of native ClpS does not occur because its NTE is not long enough to enter the ClpP peptidase chamber (**Figure 3.1f**), but is observed for a variant bearing a duplicated NTE of ~50 residues (Rivera-Rivera et al., 2014).

Implications of tucked pore-1 loops in the D2 ring

In the D2 ring of our class-**III** structures, many G<u>Y</u>VG pore-1 loops assume a 'tucked' conformation in which they rotate away from the center of the ClpA axial channel and do not engage the ClpS NTE (**Figure 3.3b**,**c**), presumably weakening ClpS•ClpA binding. There are several functional implications. First, the D2-disengaged/D1-engaged species could represent an intermediate in

the assembly of higher-affinity ClpAPS complexes in which both rings engage the NTE (**Figure 3.6a**). Second, tucked D2 GYVG pore loops could be important during latter steps in ClpS-dependent substrate delivery (**Figure 3.6b**), which require conformational remodeling of the ClpS core to weaken its interactions with the N-degron substrate and promote its transfer to ClpA (Hou et al., 2008; Román-Hernández et al., 2011; Rivera-Rivera et al., 2014). For example, after failed attempts by ClpA to fully unfold the ClpS core, release of the NTE from the D2 ring could increase the probability that ClpS dissociates completely from ClpAP, freeing the D2 pore loops to engage the N-degron substrate for degradation (see **Figure 3.6d**). Finally, pore-loop tucking does not require ATP hydrolysis, suggesting that under certain conditions (*e.g.*, when bound to the ClpS NTE), ClpA readily adopts the class-**III** structures, which constitute ~20% of particles in our final dataset.

More broadly, pore-loop tucking may be used during the process of enzyme pausing and/or unloading by ClpA and other AAA+ unfoldase motors. For example, the ClpA D1 ring functions as a 'back up' motor to prevent pausing when the principal D2-ring motor fails (Kotamarthi et al., 2020; Zuromski et al., 2021). Transiently breaking contacts with the polypeptide via pore-loop tucking in only the D2 ring would allow the weaker D1 ring to continue unfolding/translocation without working against the stalled D2 motor. Subsequent 'untucking' of the D2 pore-1 loops once the sequence causing the pause is cleared would allow the D2 motor to re-engage, restarting robust translocation by both rings. More generally, concerted loss of peptide contacts by all AAA+ domains within a ring via pore-loop rotation and tucking would be an efficient mechanism for an unfoldase either to transiently disengage from a bound polypeptide or facilitate full enzyme dissociation upon failure of a AAA+ motor to unfold, translocate, or remodel a bound protein. By contrast, dissociation of a AAA+ enzyme from its polypeptide track by transitioning from a closed, substrate-bound right-handed spiral to an open, left-handed 'lock-washer' observed in some

Hsp100 family members (Yokom et al., 2016; Gates et al., 2017; Yu et al., 2018) requires much larger, global conformational changes throughout the AAA+ hexamer.

Specialized functions of pore-2 loops

The pore-2 loops of other AAA+ unfoldases/remodeling enzymes have been shown to contact substrate polypeptides (Johjima et al., 2015; Deville et al., 2017; Gates et al., 2017; Puchades et al., 2017, 2019; Alfieri et al., 2018; Yu et al., 2018; Sandate et al., 2019; Zehr et al., 2020; Fei et al., 2020a, 2020b; Han et al., 2020; Lopez et al., 2020; Shin et al., 2021; Kavalchuk et al., 2022). The AAS residues of the ClpA D1 pore-2 loops make substantial contacts with the ClpS NTE. Nevertheless, we find that these interactions are not critical for ClpA•ClpS binding but instead help mediate mechanical work needed during ClpS-assisted N-end-rule degradation (**Figure 3.5** and **3.6b**). We propose that the D1 pore-2 loops of ClpA collaborate with the D2 pore-1 loops, which are also required for ClpS delivery (Zuromski et al., 2021), in mechanical remodeling of ClpS and/or substrate transfer to ClpA. Both sets of loops could contribute to coordinated pulling on the NTE to apply force to and remodel ClpS. Next, the pore-2 loops could capture and initiate unfolding of the 'released' N-degron substrate, and generate a sufficiently long polypeptide 'tail' to reach the more powerful D2 pore-1 loops (**Figure 3.6c**,**d**). Meanwhile, the D2 pore-1 loops could release the ClpS NTE via concerted loop-tucking, but then 'untuck' to grab this 'tail' for processive substrate unfolding and translocation. Future studies parsing the interaction of pore-1 and pore-2 loops in both ClpA rings are needed to further elucidate the mechanistic steps of N-degron substrate delivery and degradation, as well as to understand why pore-2 loops are critical in ClpS-mediated degradation but less important for other classes of substrates.

ClpA functions using both coordinated and independent action of the D1 and D2 rings

Loss of D2 pore-1 loop engagement with the ClpS NTE is a major feature distinguishing our class-**III** structures from classes **I** and **II**. Although the asymmetric engagement of substrate in the D1

but not the D2 ring of class **III** has some parallels with substrate-bound structures of NSF and Pex1•Pex6 (Blok et al., 2015; Gardner et al., 2018; White et al., 2018), the substrate-binding rings of these other enzymes adopt a ‘canonical’ right-hand spiral organization, whereas the ring that does not bind substrate assumes a planar conformation. In contrast, the non-binding, ClpA D2 ring in class **III** remains in the right-handed spiral conformation. Moreover, the portion of the ClpS NTE in the D1 ring of class-**III** structures is bound in the same fashion as our class-**I** and class-**II** structures. This structural snapshot of ‘divided’ NTE engagement between the D1 and D2 pore-1 loops reinforces biophysical and biochemical experiments that reveal a division of labor between the two AAA+ modules of ClpA (Kress et al., 2009; Kotamarthi et al., 2020; Zuromski et al., 2021). Multiple studies of other double-ring remodeling/unfoldase enzymes, including ClpB, Hsp104, ClpC, Cdc48/p97/VCP, and the ribosomal assembly factor Rix7, report the separation of substrate binding/recognition functions in one ring from the role of the second ring as the principal motor performing mechanical work (Hattendorf and Lindquist, 2002; Mogk et al., 2003; Wang et al., 2011; Doyle et al., 2012; Bodnar and Rapoport, 2017a, 2017b; Lo et al., 2019). These results illustrate that functional specialization of individual rings is emerging as a theme shared by many double-ring AAA+ unfoldases and protein-remodeling enzymes.

Acknowledgments

We thank E. Brignole and P. Dip for support in preparing and screening cryo-EM grids at the MIT.nano Automated Cryogenic Electric Microscopy Facility on a Talos Arctica microscope, which was a gift from the Arnold and Mabel Beckman Foundation, and C. Xu, K. Song, and K. Lee for data collection at the Cryo-EM Core Facility at the University of Massachusetts Chan Medical School. We thank I. Levchenko for advice preparing ClpAPS complexes and S. Bell, T. Bell, J. Park Morehouse, T. Shih, J. Zhang, and K. Zuromski for helpful advice and feedback.

This work was supported by NIH grant AI-016892 (R.T.S., T.A.B.), the Howard Hughes Medical Institute (T.A.B.), and the National Science Foundation Graduate Research Fellowship grant 1745302 (S.K.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions

Conceptualization, S.K., X.F., R.T.S., and T.A.B.; Methodology, S.K., X.F., T.A.B., and R.T.S; Validation, S.K., X.F., and R.T.S.; Formal Analysis, S.K. and X.F.; Investigation, S.K. and X.F.; Resources, S.K.; Writing – Original Draft, S.K. and X.F.; Writing – Review & Editing, S.K., X.F., R.T.S., and T.A.B.; Visualization, S.K. and X.F.; Supervision, R.T.S. and T.A.B.; Funding Acquisition, S.K., R.T.S., and T.A.B.

Ethics Declarations

The authors declare no competing interests.

METHODS

Proteins and peptides

ClpA pore-2 mutations were introduced using round-the-horn mutagenesis with T4 polynucleotide kinase and Q5 high-fidelity polymerase (New England Biolabs) into pET9a-^{M169T}ClpA, the plasmid used to express ClpA (gift from J. Flanagan, Hershey Medical Center). The M169T substitution helps overexpression of full-length ClpA (Seol et al., 1995) and is present in our lab version of 'wild-type' ClpA. ClpA and pore-2 variants were purified as described (Hou et al., 2008) and stored in HO buffer (50 mM HEPES-KOH, pH 7.5, 300 mM NaCl, 20 mM $MgCl_2$, 10% [w/v] glycerol, 0.5 mM DTT). ClpP-His_6 was expressed in *E. coli* strain JK10 (*clpP::cat*, Δ*lon*, *slyD::kan*, λDE3; Kenniston et al., 2003), purified as described (Kim et al., 2000), and stored in 50 mM Tris-HCl (pH 8), 150 mM KCl, 10% glycerol, 0.5 mM EDTA, and 1 mM DTT. ClpS and YLFVQELA-GFP were expressed in *E. coli* strain BL21(DE3) and purified as described for His_6-SUMO-fusion proteins (Hou et al., 2008; Román-Hernández et al., 2011). ^{SF}GFP-ssrA was expressed and purified as described (Nager et al., 2011). ClpS, ^{SF}GFP-ssrA, and YLFVQELA-GFP were stored in 25 mM HEPES-KOH, pH 7.5, 150 mM KCl, 10% glycerol, and 1 mM DTT. FITC-casein (Sigma-Aldrich C0528) was dissolved in HO buffer and used freshly for biochemical assays; an extinction coefficient at 280 nm (11,460 $M^{-1}cm^{-1}$) and absorbance values at 280 nm and 494 nm (to calculate and correct for overlap from the fluorescence of the FITC moiety) were used to calculate its concentration. The LLYVQRDSKEC-fluorescein N-degron synthetic peptide (21st Century Biochemicals [Marlborough, MA], molecular weight 1779.9 g/mol) was dissolved and stored at 100 μM in 15% DMSO.

Sample preparation and EM data acquisition

$ClpA_6$ (4 μM), $ClpP_{14}$ (8 μM), ClpS (13 μM), and YLFVQELA-GFP (13 μM) were mixed in 70 μL of assembly buffer (50 mM HEPES-KOH, pH 7.5, 300 mM KCl, 5 mM $MgCl_2$, 2 mM TCEP, 4% glycerol, 2 mM ATPγS [Calbiochem]) for 5 min at 25 °C. 25 μL of this mixture was then

chromatographed at room temperature and a flow rate of 0.04 mL/min on a Superdex-200 3.2/300 size exclusion column equilibrated in assembly buffer (GE Healthcare Ettan). A 50 µL fraction containing the largest molecular weight complex was assessed by SDS-PAGE (stained with SYPRO Red [Thermo Fisher]) and pooled for cryo-EM. After diluting the sample two-fold in assembly buffer, a 3 µL aliquot of the mixture was applied to glow-discharged R1.2/1.3 300 mesh holey carbon gold grids (Quantifoil). After a 15 s incubation, grids were blotted for 4 s at 4 °C, 100 % humidity, using Whatman grade 595 filter paper, and plunged into liquid ethane using a Vitrobot Mark IV system (Thermo Fisher Scientific).

A single grid was imaged for data collection using a Talos Arctica with a Gatan K3 direct electron detector (University of Massachusetts Chan Medical School Cryo-EM Microscopy Facility, Worcester, MA) in super-resolution mode, operated at 200 keV to collect high-resolution movies at (0.435 Å per pixel; uncalibrated magnification 45,000X) with a defocus range of -0.5 µm to -2.5 µm, with a total dose of 34.71 $e^-/Å^2$ over 26 frames (200 ms per frame).

Cryo-EM data processing

Each movie was binned by a factor of 2, aligned, corrected for beam-induced motion using MotionCor2 (Zheng et al., 2017), and CTF estimation was calculated by CTFFIND4 (Rohou and Grigorieff, 2015). A total of 9,169 micrographs were analyzed using RELION 3.0.8 (Zivanov et al., 2018) for data processing, classification, and 3D reconstruction. The majority of auto-picked particles were doubly capped complexes consisting of two ClpA hexamers per ClpP 14-mer. Following three rounds of 2D classification, 1,043,033 particles were used for 3D reconstruction. The cryo-EM map of ClpAP (EMD-20406), which was collected under similar parameters as our dataset, was low-pass filtered to 60 Å to generate an initial model for reconstruction. After the first round of 3D classification, two of three high-quality classes were combined, totaling 717,833 particles; the second class closely resembled the first with the exception of handedness and was flipped to correct handedness before being combined. The two classes were also utilized for per-

particle CTF refinement and motion correction. The combined class had a resolution of ~3 Å. The fulcrum was shifted to the center of ClpA, and particles were re-boxed to improve the resolution of this region of ClpA before performing the second round of 3D classification (T=4) with alignment to generate six classes. Three good classes were selected and combined (358,726 particles) for a third round of 3D classification with a ClpA mask, without alignment (T=20) to yield the final six classes. Each class was then subjected to 3D auto-refinement without symmetry to yield six maps with ~3.5 Å resolution. To generate the final maps, each map was density-modified and autosharpened in PHENIX (Adams et al., 2010), giving final resolutions of ranging from ~3.2-3.4 Å (**Table 3.1**).

Molecular modeling and refinement

The ClpAP cryo-EM structure (PDB code 6W23) was docked into the EM map for the class **I** structure, and the ClpAP cryo-EM structure (PDB code 6W22) was docked into all other EM maps using “fit in map” in Chimera (Pettersen et al., 2004). Real-space refinement was performed using PHENIX, and model building was performed in Coot (Emsley et al., 2004). The ClpS NTE sequence (residues 2–26) was added manually in Coot. Geometry of the final models was evaluated using MolProbity (Williams et al., 2018). Figures and movies were generated using Chimera, ChimeraX (Pettersen et al., 2021), and PyMOL (Schrödinger, LLC).

Multiple sequence alignment

The amino-acid sequences of bacterial ClpA, ClpB, and ClpC proteins were downloaded from UniProtKB (Bateman et al., 2021) and aligned using MUSCLE (Edgar, 2004) with MEGA7 (Kumar et al., 2016). The sequence alignment was visualized in Jalview (Waterhouse et al., 2009) and colored according to the Clustal X scheme.

Buried surface area calculations

The buried surface area of the ClpS NTE in all class structures was analyzed using the 'Protein interfaces, surfaces and assemblies' service PISA at the European Bioinformatics Institute. (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html). The BSA values were summed from the D1 pore-1 loop region (residues 254-264), D1 pore-2 loop region (residues 292-302), D2 pore-1 loop region (residues 536-544), and D2 pore-2 loop region (residues 525-531).

Biochemical assays

Biochemical experiments were performed with at least three technical replicates at 30 °C in HO buffer using a SpectraMax M5 Microplate Reader (Molecular Devices) to measure initial rates of absorbance or fluorescence changes or equilibrium anisotropy values. ATP-hydrolysis rates were measured over the first ~2 min by monitoring the loss of absorbance at 340 nm using a coupled NADH–ATP assay (Burton et al., 2001) with 5 mM ATP (Sigma-Aldrich), pyruvate kinase (Sigma-Aldrich; P9136 at 20 units/mL), lactate dehydrogenase (Sigma-Aldrich; L1254 at 20 units/mL), 7.5 mM phosphoenolpyruvate (Sigma-Aldrich; P0564), and 0.2 mM NADH (Roche 10107735001). ATP-hydrolysis assays were performed under four conditions: (*i*) $ClpA_6$ or variants (0.2 μM); (*ii*) $ClpA_6$ or variants (0.1 μM) and $ClpP_{14}$ (0.1 μM); (*iii*) $ClpA_6$ or variants (0.2 μM), $ClpP_{14}$ (0.2 μM), and ^{SF}GFP-ssrA (3 μM); and (*iv*) $ClpA_6$ or variants (0.2 μM), $ClpP_{14}$ (0.2 μM), and ClpS (0.6 μM).

FITC-casein degradation assays were monitored by increases in fluorescence (excitation 340 nm, emission 520 nm) as a consequence of protease-dependent unquenching over the first 5 min; reactions contained $ClpA_6$ or variants (0.2 μM), $ClpP_{14}$ (0.4 μM), and an ATP-regeneration system (4 mM ATP, 50 μg/mL creatine kinase [Sigma-Aldrich], 5 mM creatine phosphate [Sigma-Aldrich]). The endpoint fluorescence for complete FITC-casein degradation was determined by addition of porcine elastase (100 μg/mL; Sigma-Aldrich) to each well, followed by a 30-min incubation prior to reading. To determine FITC-casein degradation rates, the increase in relative fluorescence

units was normalized to the endpoint fluorescence value from fully unquenched substrate after porcine elastase incubation and the background rate was subtracted from each reaction on the basis of a no-enzyme buffer-only control.

Degradation of GFP variants was monitored by loss of fluorescence (excitation 467 nm, emission 511 nm) over the first 5-10 min. Briefly, rates were calculated by normalizing the slope values of relative fluorescence units (RFUs)/time by the fluorescence signal determined from a standard curve of RFUs versus varying concentrations of substrate in the linear range. Degradation of different concentrations of ^{SF}GFP-ssrA (0.25 - 20 μM) was assayed using $ClpA_6$ or variants (0.2 μM), $ClpP_{14}$ (0.4 μM), and the ATP-regeneration system described above. Degradation of different concentrations of YLFVQELA-GFP (0.25 - 20 μM) was assayed using $ClpA_6$ or variants (0.1 μM), $ClpP_{14}$ (0.2 μM), ClpS (0.6 μM), and the ATP-regeneration system. For degradation of YLFVQELA-GFP by the GGG, KYR, and Δ295-297 pore-2-loop mutants shown in Fig. 3.5b, concentrations were $ClpA_6$ variant (0.6 μM), $ClpP_{14}$ (1.2 μM), and ClpS (3.6 μM).

The binding of the peptide LLYVQRDSKEC-fluorescein (100 nM) to (*i*) $ClpA_6$ or variants (0.047 - 6 μM), (*ii)* ClpS (1.25 - 30 μM), or (*iii*) equimolar mixtures of $ClpA_6$ or variants and ClpS (0.047 - 3 μM) at equilibrium was assayed by fluorescence anisotropy (excitation 490 nm, emission 525 nm) in the presence of ATPγS (2 mM). Only $ClpA_6$•ClpS•peptide ternary and $ClpA_6$•peptide binary complexes have higher anisotropy levels (in comparison to ClpS•peptide binary complexes) as a result of the much larger molecular weight of $ClpA_6$ (~500 kDa) compared to that of ClpS (~10 kDa). Data were fit by a non-linear least-squares algorithm to equations for $ClpA_6$ only and ClpS only experiments:

$$\text{fluorescence anisotropy} = f_{min} + \left(\frac{f_{max} * K_d}{K_d + X}\right)$$

or to a quadratic equation for tight binding for $ClpA_6$•ClpS complexes:

$$\text{fluorescence anisotropy} = f_{min} + (f_{max} - f_{min})\frac{(L + X + K_d) - \sqrt{(L + X + K_d)^2 - 4 \cdot L \cdot X}}{2L}$$

where f_{min} is the background anisotropy value, f_{max} is the maximum anisotropy value at saturated binding, L is the concentration of peptide (100 nM), K_{d} is the dissociation equilibrium constant (in nM), and X is the concentration of $ClpA_6$•ClpS (in nM).

Data availability

Atomic models and EM maps have been deposited in the PDB and EMDB with accession codes provided in **Table 3.1**. Any additional information required to reanalyze the data reported in this study is available from the corresponding author (tabaker@mit.edu).

REFERENCES

Adams, P.D., Afonine, P. V, Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. *66*, 213–221.

Alfieri, C., Chang, L., and Barford, D. (2018). Mechanism for remodelling of the cell cycle checkpoint protein MAD2 by the ATPase TRIP13. Nature *559*, 7713, 274–278.

Barkow, S.R., Levchenko, I., Baker, T.A., and Sauer, R.T. (2009). Polypeptide translocation by the AAA+ ClpXP protease machine. Chem. Biol. *16*, 6, 605.

Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E.H., Britto, R., Bursteinas, B., et al. (2021). UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. *49*, D1, D480–D489.

Bell, T.A., Baker, T.A., and Sauer, R.T. (2019). Interactions between a subset of substrate side chains and AAA+ motor pore loops determine grip during protein unfolding. eLife *8*, e46808.

Blok, N.B., Tan, D., Wang, R.Y.-R., Penczek, P.A., Baker, D., DiMaio, F., Rapoport, T.A., and Walz, T. (2015). Unique double-ring structure of the peroxisomal Pex1/Pex6 ATPase complex revealed by cryo-electron microscopy. Proc. Natl. Acad. Sci. U. S. A. *112*, 30, E4017-25.

Bodnar, N., and Rapoport, T. (2017a). Toward an understanding of the Cdc48/p97 ATPase. F1000Research *6*, 1318.

Bodnar, N.O., and Rapoport, T.A. (2017b). Molecular mechanism of substrate processing by the Cdc48 ATPase complex. Cell *169*, 4, 722-735.e9.

Burton, R.E., Siddiqui, S.M., Kim, Y.I., Baker, T.A., and Sauer, R.T. (2001). Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine. EMBO J. *20*, 12, 3092–3100.

Cooney, I., Han, H., Stewart, M.G., Carson, R.H., Hansen, D.T., Iwasa, J.H., Price, J.C., Hill, C.P., and Shen, P.S. (2019). Structure of the Cdc48 segregase in the act of unfolding an authentic substrate. Science. *365*, 6452, 502–505.

Daskalogianni, C., Apcher, S., Candeias, M.M., Naski, N., Calvo, F., and Fåhraeus, R. (2008). Gly-Ala repeats induce position- and substrate-specific regulation of 26 S proteasome-dependent partial processing. J. Biol. Chem. *283*, 44, 30090–30100.

Deville, C., Carroni, M., Franke, K.B., Topf, M., Bukau, B., Mogk, A., and Saibil, H.R. (2017). Structural pathway of regulated substrate transfer and threading through an Hsp100 disaggregase. Sci. Adv. *3*, 8, e1701726.

De Donatis, G.M., Singh, S.K., Viswanathan, S., and Maurizi, M.R. (2010). A single ClpS monomer is sufficient to direct the activity of the ClpA hexamer. J. Biol. Chem. *285*, 12, 8771–8781.

Dong, Y., Zhang, S., Wu, Z., Li, X., Wang, W.L., Zhu, Y., Stoilova-McPhie, S., Lu, Y., Finley, D., and Mao, Y. (2019). Cryo-EM structures and dynamics of substrate-engaged human 26S proteasome. Nature *565*, 7737, 49–55.

Dougan, D.A., Reid, B.G., Horwich, A.L., and Bukau, B. (2002). ClpS, a substrate modulator of the ClpAP machine. Mol. Cell *9*, 3, 673–683.

Doyle, S.M., Hoskins, J.R., and Wickner, S. (2012). DnaK chaperone-dependent disaggregation by caseinolytic peptidase B (ClpB) mutants reveals functional overlap in the N-terminal domain and nucleotide-binding domain-1 pore tyrosine. J. Biol. Chem. *287*, 34, 28470–28479.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 5, 1792–1797.

Effantin, G., Ishikawa, T., De Donatis, G.M., Maurizi, M.R., and Steven, A.C. (2010). Local and global mobility in the ClpA AAA+ chaperone detected by cryo-electron microscopy: Functional connotations. Structure *18*, 5, 553–562.

Emsley, P., Cowtan, K., and IUCr (2004). *Coot* : model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. *60*, 12, 2126–2132.

Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D.A., and Bukau, B. (2006). ClpS is an essential component of the N-end rule pathway in Escherichia coli. Nature *439*, 7077, 753–756.

Erzberger, J.P., and Berger, J.M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. Annu. Rev. Biophys. Biomol. Struct. *35*, 93–114.

Fei, X., Bell, T.A., Jenni, S., Stinson, B.M., Baker, T.A., Harrison, S.C., and Sauer, R.T. (2020a). Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. eLife *9*, e52774.

Fei, X., Bell, T.A., Barkow, S.R., Baker, T.A., and Sauer, R.T. (2020b). Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. eLife *9*, e61496.

Gardner, B.M., Castanzo, D.T., Chowdhury, S., Stjepanovic, G., Stefely, M.S., Hurley, J.H., Lander, G.C., and Martin, A. (2018). The peroxisomal AAA-ATPase Pex1/Pex6 unfolds substrates by processive threading. Nat. Commun. *9*, 1, 135.

Gates, S.N., Yokom, A.L., Lin, J., Jackrel, M.E., Rizo, A.N., Kendsersky, N.M., Buell, C.E., Sweeny, E.A., Mack, K.L., Chuang, E., et al. (2017). Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. Science. *357*, 6348, 273–279.

Grimaud, R., Kessel, M., Beuron, F., Steven, A.C., and Maurizi, M.R. (1998). Enzymatic and structural similarities between the *Escherichia coli* ATP-dependent proteases, ClpXP and ClpAP. J. Biol. Chem. *273*, 20, 12476–12481.

Guo, F., Esser, L., Singh, S.K., Maurizi, M.R., and Xia, D. (2002). Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA. J. Biol. Chem. *277*, 48, 46753–46762.

Han, H., Fulcher, J.M., Dandey, V.P., Iwasa, J.H., Sundquist, W.I., Kay, M.S., Shen, P.S., and Hill, C.P. (2019). Structure of Vps4 with circular peptides and implications for translocation of two polypeptide chains by AAA+ ATPases. eLife *8*, e44071.

Han, H., Schubert, H.L., McCullough, J., Monroe, N., Purdy, M.D., Yeager, M., Sundquist, W.I., and Hill, C.P. (2020). Structure of spastin bound to a glutamate-rich peptide implies a hand-over-hand mechanism of substrate translocation. J. Biol. Chem. *295*, 2, 435–443.

Hattendorf, D.A., and Lindquist, S.L. (2002). Cooperative kinetics of both Hsp104 ATPase domains and interdomain communication revealed by AAA sensor-1 mutants. EMBO J. *21*, 1–2, 12–21.

Hinnerwisch, J., Fenton, W.A., Furtak, K.J., Farr, G.W., and Horwich, A.L. (2005a). Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. Cell *121*, 7, 1029–1041.

Hinnerwisch, J., Reid, B.G., Fenton, W.A., and Horwich, A.L. (2005b). Roles of the N-domains of the ClpA unfoldase in binding substrate proteins and in stable complex formation with the ClpP protease. J. Biol. Chem. *280*, 49, 40838–40844.

Hoskins, J.R., Pak, M., Maurizi, M.R., and Wickner, S. (1998). The role of the ClpA chaperone in proteolysis by ClpAP. Proc. Natl. Acad. Sci. U. S. A. *95*, 21, 12135–12140.

Hoskins, J.R., Singh, S.K., Maurizi, M.R., and Wickner, S. (2000). Protein binding and unfolding by the chaperone ClpA and degradation by the protease ClpAP. Proc. Natl. Acad. Sci. U. S. A. *97*, 16, 8892–8897.

Hou, J.Y., Sauer, R.T., and Baker, T.A. (2008). Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. Nat. Struct. Mol. Biol. *15*, 3, 288–294.

Hoyt, M.A., Zich, J., Takeuchi, J., Zhang, M., Govaerts, C., and Coffino, P. (2006). Glycine-alanine repeats impair proper substrate unfolding by the proteasome. EMBO J. *25*, 8, 1720–1729.

Iosefson, O., Olivares, A.O., Baker, T.A., and Sauer, R.T. (2015). Dissection of axial-pore loop function during unfolding and translocation by a AAA+ proteolytic machine. Cell Rep. *12*, 6, 1032–1041.

Ishikawa, T., Beuron, F., Kessel, M., Wickner, S., Maurizi, M.R., and Steven, A.C. (2001). Translocation pathway of protein substrates in ClpAP protease. Proc. Natl. Acad. Sci. U. S. A. *98*, 8, 4328–4333.

Johjima, A., Noi, K., Nishikori, S., Ogi, H., Esaki, M., and Ogura, T. (2015). Microtubule severing by katanin p60 AAA+ATPase requires the C-terminal acidic tails of both α-and β-tubulins and basic amino acid residues in the AAA+ring pore. J. Biol. Chem. *290*, 18, 11762–11770.

Kavalchuk, M., Jomaa, A., Müller, A.U., and Weber-Ban, E. (2022). Structural basis of prokaryotic ubiquitin-like protein engagement and translocation by the mycobacterial Mpa-proteasome complex. Nat. Commun. 2022 131 *13*, 1, 1–13.

Kenniston, J.A., Baker, T.A., Fernandez, J.M., and Sauer, R.T. (2003). Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. Cell *114*, 4, 511–520.

Kim, Y.I., Burton, R.E., Burton, B.M., Sauer, R.T., and Baker, T.A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. Mol. Cell *5*, 4, 639–648.

Kotamarthi, H.C., Sauer, R.T., and Baker, T.A. (2020). The non-dominant AAA+ ring in the ClpAP protease functions as an anti-stalling motor to accelerate protein unfolding and translocation. Cell Rep. *30*, 8, 2644-2654.e3.

Kraut, D.A. (2013). Slippery substrates impair ATP-dependent protease function by slowing unfolding. J. Biol. Chem. *288*, 48, 34729–34735.

Kraut, D.A., Israeli, E., Schrader, E.K., Patil, A., Nakai, K., Nanavati, D., Inobe, T., and Matouschek, A. (2012). Sequence- and species-dependence of proteasomal processivity. ACS Chem. Biol. *7*, 8, 1444–1453.

Kress, W., Mutschler, H., and Weber-Ban, E. (2009). Both ATPase domains of ClpA are critical for processing of stable protein structures. J. Biol. Chem. *284*, 45, 31441–31452.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. *372*, 3, 774–797.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. *33*, 7, 1870–1874.

Levitskaya, J., Sharipo, A., Leonchiks, A., Ciechanover, A., and Masucci, M.G. (1997). Inhibition of ubiquitin/proteasome-dependent protein degradation by the Gly-Ala repeat domain of the Epstein-Barr virus nuclear antigen 1. Proc. Natl. Acad. Sci. U. S. A. *94*, 23, 12616–12621.

Lin, L., and Ghosh, S. (1996). A glycine-rich region in NF-kappaB p105 functions as a processing signal for the generation of the p50 subunit. Mol. Cell. Biol. *16*, 5, 2248–2254.

Lo, Y.-H.H., Sobhany, M., Hsu, A.L., Ford, B.L., Krahn, J.M., Borgnia, M.J., and Stanley, R.E. (2019). Cryo-EM structure of the essential ribosome assembly AAA-ATPase Rix7. Nat. Commun. *10*, 1, 513.

Lopez, K.E., Rizo, A.N., Tse, E., Lin, J., Scull, N.W., Thwin, A.C., Lucius, A.L., Shorter, J., and Southworth, D.R. (2020). Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis. Nat. Struct. Mol. Biol. *27*, 5, 406–416.

Mahmoud, S.A., and Chien, P. (2018). Regulated proteolysis in bacteria. Annu. Rev. Biochem. *87*, 677-696

Martin, A., Baker, T.A., and Sauer, R.T. (2008). Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. Nat. Struct. Mol. Biol. *15*, 11, 1147–1151.

Miller, J.M., and Lucius, A.L. (2014). ATPγS competes with ATP for binding at Domain 1 but not Domain 2 during ClpA catalyzed polypeptide translocation. Biophys. Chem. *185*, 58–69.

Mogk, A., Schlieker, C., Strub, C., Rist, W., Weibezahn, J., and Bukau, B. (2003). Roles of individual domains and conserved motifs of the AAA+ chaperone ClpB in oligomerization, ATP hydrolysis, and chaperone activity. J. Biol. Chem. *278*, 20, 17615–17624.

Nager, A.R., Baker, T.A., and Sauer, R.T. (2011). Stepwise unfolding of a β barrel protein by the AAA+ ClpXP protease. J. Mol. Biol. *413*, 1, 4–16.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 13, 1605–1612.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. *30*, 1, 70–82.

Puchades, C., Rampello, A.J., Shin, M., Giuliano, C.J., Wiseman, R.L., Glynn, S.E., and Lander, G.C. (2017). Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. Science. *358*, 6363.

Puchades, C., Ding, B., Song, A., Wiseman, R.L., Lander, G.C., and Glynn, S.E. (2019). Unique structural features of the mitochondrial AAA+ protease AFG3L2 reveal the molecular basis for activity in health and disease. Mol. Cell *75*, 5, 1073-1085.e6.

Puchades, C., Sandate, C.R., and Lander, G.C. (2020). The molecular principles governing the activity and functional diversity of AAA+ proteins. Nat. Rev. Mol. Cell Biol. *21*, 1, 43–58.

Ripstein, Z.A., Huang, R., Augustyniak, R., Kay, L.E., and Rubinstein, J.L. (2017). Structure of a AAA+ unfoldase in the process of unfolding substrate. eLife *6*, e25754.

Ripstein, Z.A., Vahidi, S., Houry, W.A., Rubinstein, J.L., and Kay, L.E. (2020). A processive rotary mechanism couples substrate unfolding and proteolysis in the ClpXP degradation machinery. eLife *9*, e52158.

Rivera-Rivera, I. (2015). Mechanism of active substrate delivery by the AAA+ protease adaptor ClpS. Ph.D. thesis, Massachusetts Institute of Technology. Retrieved from https://dspace.mit.edu/handle/1721.1/101352

Rivera-Rivera, I., Román-Hernández, G., Sauer, R.T., and Baker, T.A. (2014). Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. Proc. Natl. Acad. Sci. U. S. A. *111*, 37, E3853-9.

Rizo, A.N., Lin, J.B., Gates, S.N., Tse, E., Bart, S.M., Castellano, L.M., DiMaio, F., Shorter, J., and Southworth, D.R. (2019). Structural basis for substrate gripping and translocation by the ClpB AAA+ disaggregase. Nat. Commun. *10*, 1.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. *192*, 2, 216–221.

Román-Hernández, G., Grant, R.A., Sauer, R.T., and Baker, T.A. (2009). Molecular basis of substrate selection by the N-end rule adaptor protein ClpS. Proc. Natl. Acad. Sci. U. S. A. *106*, 22, 8888–8893.

Román-Hernández, G., Hou, J.Y., Grant, R.A., Sauer, R.T., and Baker, T.A. (2011). The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. Mol. Cell *43*, 2, 217–228.

Sandate, C.R., Szyk, A., Zehr, E.A., Lander, G.C., and Roll-Mecak, A. (2019). An allosteric network in spastin couples multiple activities required for microtubule severing. Nat. Struct. Mol. Biol. *26*, 8, 671–678.

Sauer, R.T., and Baker, T.A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. Annu. Rev. Biochem. *80*, 1, 587–612.

Schlieker, C., Weibezahn, J., Patzelt, H., Tessarz, P., Strub, C., Zeth, K., Erbse, A., Schneider-Mergener, J., Chin, J.W., Schultz, P.G., et al. (2004). Substrate recognition by the AAA+ chaperone ClpB. Nat. Struct. Mol. Biol. *11*, 7, 607–615.

Seol, J.H., Yoo, S.J., Kang, M.-S., Ha, D.B., and Chung, C.H. (1995). The 65-kDa protein derived from the internal translational start site of the clpA gene blocks autodegradation of ClpA by the ATP-dependent protease Ti in *Escherichia coli*. FEBS Lett. *377*, 1, 41–43.

Sharipo, A., Imreh, M., Leonchiks, A., Brändén, C.I., and Masucci, M.G. (2001). *cis*-Inhibition of proteasomal degradation by viral repeats: Impact of length and amino acid composition. FEBS Lett. *499*, 1–2, 137–142.

Shin, M., Watson, E.R., Song, A.S., Mindrebo, J.T., Novick, S.J., Griffin, P.R., Wiseman, R.L., and Lander, G.C. (2021). Structures of the human LONP1 protease reveal regulatory steps involved in protease activation. Nat. Commun. *12*, 1, 3239.

Singh, S.K., Grimaud, R., Hoskins, J.R., Wickner, S., and Maurizi, M.R. (2000). Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. Proc. Natl. Acad. Sci. U. S. A. *97*, 16, 8898–8903.

Thompson, M.W., Singh, S.K., and Maurizi, M.R. (1994). Processive degradation of proteins by the ATP-dependent Clp protease from *Escherichia coli*. Requirement for the multiple array of active sites in ClpP but not ATP hydrolysis. J. Biol. Chem. *269*, 27, 18209–18215.

Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991). The N-end rule in bacteria. Science. *254*, 5036, 1374–1377.

Too, P.H.M., Erales, J., Simen, J.D., Marjanovic, A., and Coffino, P. (2013). Slippery substrates impair function of a bacterial protease ATPase by unbalancing translocation versus exit. J. Biol. Chem. *288*, 19, 13243–13257.

Torres-Delgado, A., Kotamarthi, H.C., Sauer, R.T., and Baker, T.A. (2020). The intrinsically disordered N-terminal extension of the ClpS adaptor reprograms its partner AAA+ ClpAP protease. J. Mol. Biol. *432*, 17, 4908–4921.

Twomey, E.C., Ji, Z., Wales, T.E., Bodnar, N.O., Ficarro, S.B., Marto, J.A., Engen, J.R., and Rapoport, T.A. (2019). Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. Science. *365*, 6452, eaax1033.

Varshavsky, A. (2019). N-degron and C-degron pathways of protein degradation. Proc. Natl. Acad. Sci. U. S. A. *116*, 2, 358–366.

Vass, R.H., and Chien, P. (2013). Critical clamp loader processing by an essential AAA+ protease in *Caulobacter crescentus*. Proc. Natl. Acad. Sci. U. S. A. *110*, 45, 18138–18143.

Wang, F., Mei, Z., Qi, Y., Yan, C., Hu, Q., Wang, J., and Shi, Y. (2011). Structure and mechanism of the hexameric MecA–ClpC molecular machine. Nature *471*, 7338, 331–335.

Wang, K.H., Sauer, R.T., and Baker, T.A. (2007). ClpS modulates but is not essential for bacterial N-end rule degradation. Genes Dev. *21*, 4, 403–408.

Wang, K.H., Roman-Hernandez, G., Grant, R.A., Sauer, R.T., and Baker, T.A. (2008a). The molecular basis of N-end rule recognition. Mol. Cell *32*, 3, 406–414.

Wang, K.H., Oakes, E.S.C., Sauer, R.T., and Baker, T.A. (2008b). Tuning the strength of a bacterial N-end rule degradation signal. J. Biol. Chem. *283*, 36, 24600–24607.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 9, 1189–1191.

Weibezahn, J., Tessarz, P., Schlieker, C., Zahn, R., Maglica, Z., Lee, S., Zentgraf, H., Weber-Ban, E.U., Dougan, D.A., Tsai, F.T.F., et al. (2004). Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. Cell *119*, 5, 653–665.

White, K.I., Zhao, M., Choi, U.B., Pfuetzner, R.A., and Brunger, A.T. (2018). Structural principles of SNARE complex recognition by the AAA+ protein NSF. eLife *7*, e38888.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci. *27*, 1, 293–315.

Yokom, A.L., Gates, S.N., Jackrel, M.E., Mack, K.L., Su, M., Shorter, J., and Southworth, D.R. (2016). Spiral architecture of the Hsp104 disaggregase reveals the basis for polypeptide translocation. Nat. Struct. Mol. Biol. *23*, 9, 830–837.

Yu, H., Lupoli, T.J., Kovach, A., Meng, X., Zhao, G., Nathan, C.F., and Li, H. (2018). ATP hydrolysis-coupled peptide translocation mechanism of *Mycobacterium tuberculosis* ClpB. Proc. Natl. Acad. Sci. U. S. A. *115*, 41, E9560–E9569.

Zehr, E.A., Szyk, A., Szczesna, E., and Roll-Mecak, A. (2020). Katanin grips the β-tubulin tail through an electropositive double spiral to sever microtubules. Dev. Cell *52*, 1, 118-131.e6.

Zeth, K., Dougan, D.A., Cusack, S., Bukau, B., and Ravelli, R.B. (2002a). Crystallization and preliminary X-ray analysis of the Escherichia coli adaptor protein ClpS, free and in complex with the N-terminal domain of ClpA. Acta Crystallogr. Sect. D Biol. Crystallogr. *58*, 7, 1207–1210.

Zeth, K., Ravelli, R.B., Paal, K., Cusack, S., Bukau, B., and Dougan, D.A. (2002b). Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. Nat. Struct. Biol. *9*, 12, 906–911.

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 4, 331–332.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife *7*, e42166.

Zuromski, K.L., Sauer, R.T., and Baker, T.A. (2020). Modular and coordinated activity of AAA+ active sites in the double-ring ClpA unfoldase of the ClpAP protease. Proc. Natl. Acad. Sci. U. S. A. *117*, 41, 25455–25463.

Zuromski, K.L., Kim, S., Sauer, R.T., and Baker, T.A. (2021). Division of labor between the pore-1 loops of the D1 and D2 AAA+ rings coordinates substrate selectivity of the ClpAP protease. J. Biol. Chem. *297*, 6, 101407.

CHAPTER 4

The Pro-Pro dipeptide in the ClpS adaptor is an anti-degron for the ClpAP AAA+ protease

This chapter has been written as a manuscript for publication.

Sora Kim[1], Izarys Rivera-Rivera[1,2], Robert T. Sauer[1], and Tania A. Baker[1]

[1] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139
[2] Present address: Gingko Bioworks, Boston, MA 02210

Author contributions:

S. Kim: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review and editing, Visualization, Funding acquisition

I. Rivera-Rivera: Conceptualization, Resources, Writing – review and editing

R.T. Sauer and T.A. Baker: Conceptualization, Methodology, Validation, Formal analysis, Writing – review & editing, Supervision, Funding acquisition

ABSTRACT

The bacterial protease ClpAP, composed of the ClpA unfoldase and ClpP peptidase, binds the N-terminal extension (NTE) of its adaptor ClpS in the same manner as a protein substrate. However, during N-degron substrate delivery, ClpS transfers its cargo to ClpA and escapes ClpAP degradation. Here, we elucidate the minimal structural elements needed to protect ClpS from ClpAP proteolysis. The flexible NTE of ClpS is joined to a folded core domain that binds N-degron substrates. We find that a Pro-Pro dipeptide from the junction between the ClpS NTE and core is sufficient for proteolytic resistance when fused to other stably folded protein domains. Using a panel of ssrA-tagged fusion proteins, we demonstrate that this short Pro-Pro element can inhibit degradation of the β-barrel GFP or the α-helical-bundle Rop proteins but not the metastable λ cI^N repressor domain. Pro-Pro can also antagonize ClpXP degradation of ssrA-tagged GFP, suggesting that this ClpS-derived element is protective against many AAA+ proteases. Modulating the spacing of Pro-Pro from the folded domain shows that this motif must be close to the native structure (*i.e.* less than three residues apart) to be effective. These results indicate that the Pro-Pro sequence at the NTE–core junction functions as an anti-degron, promoting mechanical remodeling of ClpS, rather than its degradation, and substrate transfer.

Keywords

ATPases associated with diverse cellular activities; protein degradation; AAA+ ATPase adaptor; adaptor remodeling

INTRODUCTION

Proteases of the AAA+ (ATPases associated with diverse cellular activities) superfamily control protein turnover during normal growth and in response to stress conditions (Varshavsky, 2017; Mahmoud and Chien, 2018). Found in all organisms, AAA+ proteases consist of a hexameric AAA+ unfoldase and a barrel-shaped partner peptidase, which assemble into stacked rings with a narrow axial channel. Protein substrates are recognized directly by the AAA+ unfoldase through interactions of its axial channel with a peptide degron sequence in the substrate or indirectly with the assistance of adaptor proteins. Using cycles of ATP binding and hydrolysis, the AAA+ unfoldase unfolds native structure in the substrate and processively translocates the substrate polypeptide into the peptidase for degradation (Sauer and Baker, 2011; Mahmoud and Chien, 2018).

The bacterial AAA+ protease ClpAP, composed of the $ClpA_6$ unfoldase and $ClpP_{14}$ peptidase, recognizes proteins bearing degrons located at their C-termini, such as the ssrA tag, and those with N-terminal degrons (Tobias et al., 1991; Gottesman et al., 1998). ClpS, a ClpAP adaptor protein, inhibits degradation of ssrA-tagged substrates and enhances recognition of a specific class of N-degrons containing N-end-rule destabilizing residues (Dougan et al., 2002; Erbse et al., 2006; Wang et al., 2007). In *Escherichia coli*, substrates bearing Phe, Trp, Tyr, or Leu N-terminal residues bind to a hydrophobic pocket in the ClpS core domain (res. 27–106), which docks with an N-terminal domain of ClpA (**Figure 4.1a**) to form an enzyme•adaptor•substrate ternary complex (Guo et al., 2002; Zeth et al., 2002a, 2002b; Wang et al., 2008). ClpS also binds to ClpAP using its N-terminal extension (NTE), which is intrinsically disordered in the free adaptor but adopts an extended conformation, similar to a substrate polypeptide, when bound in the ClpA channel (**Figure 4.1b**) (see Chapter 3, p. 83). The degron-like NTE is required for ClpS-assisted delivery of N-end-rule substrates, inhibition of ClpAP degradation of ssrA-tagged substrates, and suppression of ClpAP ATPase rate by ClpS (Dougan et al., 2002; Hou et al., 2008; De Donatis et

al., 2010; Román-Hernández et al., 2011; Rivera-Rivera et al., 2014; Torres-Delgado et al., 2020). Thus, the direct interaction of the ClpS NTE with the unfoldase machinery of ClpA differentiates ClpS from most adaptors, which usually tether substrates to auxiliary domains of a AAA+ enzyme.

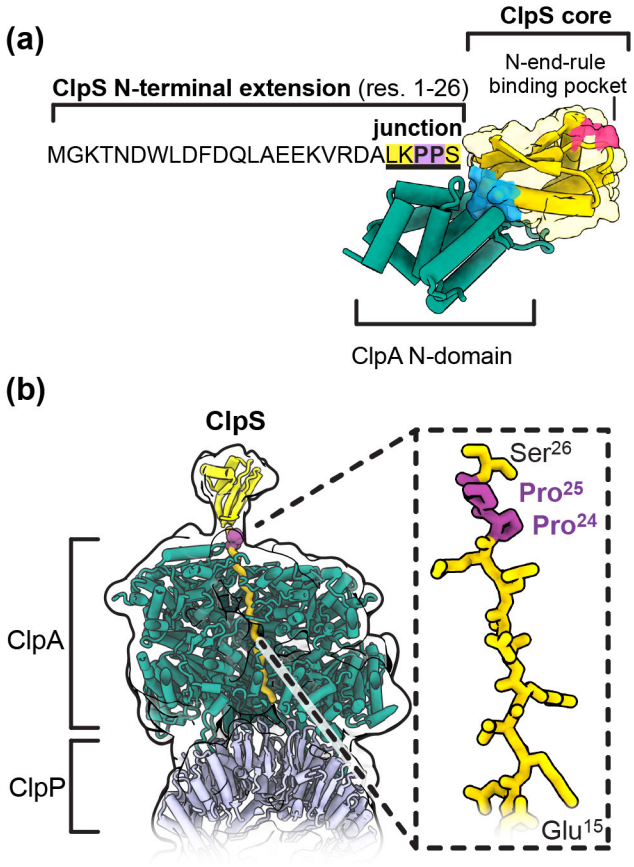


Figure 4.1 Pro^{24}-Pro^{25} in the ClpS NTE junction precede the ClpS core domain.
(a) Structure of the *E. coli* ClpS core domain in complex with the ClpA N-domain (PDB 1MBU). The NTE sequence is shown, with the Pro-Pro dipeptide colored purple.
(b) Degron-like binding of the ClpS NTE to ClpAP (PDB 7UIZ). The ClpS core domain (PDB 3O2B; res. 27–106) is docked to a 10 Å low-pass filtered cryo-EM map (black outline; EMDB 26558; PDB 7UIZ). The right panel shows a zoomed-in view of the ClpS NTE and the Pro^{24}-Pro^{25} dipeptide
For (a) and (b), ClpS is colored yellow, and ClpA is colored green.

Though the assembly of the ClpAPS•N-degron complex is well-characterized, the details of subsequent reaction steps during ClpS-assisted substrate delivery await elucidation. For example, to access the N-degron hidden in the ClpS binding pocket, ClpA appears to mechanically remodel the ClpS core domain, allowing the substrate to be transferred from ClpS to the ClpA axial channel (Román-Hernández et al., 2011; Rivera-Rivera et al., 2014). Following this substrate handoff step, ClpS presumably dissociates from the ClpA channel, such that it can bind and deliver additional substrate molecules. Although the ClpS NTE mimics a degron and can be engaged by ClpA pore loops, ClpS resists degradation by its cognate protease (Dougan et al., 2002; Román-Hernández et al., 2011; Rivera-Rivera et al., 2014). Therefore, a mechanism likely

exists to specifically promote ClpS remodeling, substrate transfer, and release, instead of global unfolding and degradation of ClpS. Rates of protein unfolding by AAA+ proteases depend on the local stability of degron-proximal structure of the substrate, as well as on the ability of the unfoldase to grip residues of the degron near the folded domain (Matouschek et al., 1999; Lee et al., 2001; Kenniston et al., 2003, 2004; Matouschek, 2003; Koodithangal et al., 2009; Olivares et al., 2017; Bell et al., 2019). ClpS contains a Pro^{24}-Pro^{25} dipeptide in the junction between its NTE and core domain (colored purple in **Figure 4.1**) that is moderately conserved, unlike the poorly conserved residues in the preceding NTE portion (Hou et al., 2008; Román-Hernández et al., 2011). Mutation of this sequence to Ala^{24}-Ala^{25} results in ClpS degradation by ClpAP, indicating that the ClpS junction modulates protein unfolding and degradation (Rivera-Rivera, 2015; Zuromski et al., 2021).

To investigate the role of the Pro-Pro motif in protecting the ClpS protein from ClpAP, we designed ssrA-tagged fusion proteins that directly test whether Pro-Pro antagonizes degron-mediated degradation. These engineered substrates isolate protection by the Pro-Pro motif from contributions by the ClpS core domain and the rest of the ClpS NTE. We find that a Pro-Pro dipeptide adjacent to a stably folded domain is necessary and sufficient to inhibit degradation by ClpAP. Importantly, in addition to Pro-Pro itself, our experiments reveal that the location of this sequence, positioned within a few residues of the beginning of the folded domain, is also important. Protection from degradation is not limited to ClpAP, as GFP bearing Pro-Pro and an ssrA tag also resists degradation by ClpXP. We conclude that the intrinsic Pro-Pro motif counteracts the substrate-like interactions of the ClpS NTE with the ClpA channel, allowing ClpS to escape degradation and instead promote multiple rounds of N-degron substrate delivery, thereby functioning as an ‘anti-degron’.

RESULTS AND DISCUSSION

Pro-Pro protects ssrA-tagged GFP from degradation by ClpAP

To interrogate the ability of Pro-Pro to inhibit ClpAP degradation of another protein, we fused a variable sequence and the ssrA tag to the C-terminus of GFP derived from *Aequeorea victoria* (**Figure 4.2a**). GFP-ssrA is a highly stable protein and an efficiently degraded ClpAP substrate (Singh et al., 2000). Single-molecule experiments have characterized the rates of GFP unfolding, which is rate-limiting for degradation, and translocation by ClpAP and ClpXP (Martin et al., 2008; Maillard et al., 2011; Sen et al., 2013; Olivares et al., 2014). We truncated GFP at Ile^{229}, the last structurally ordered residue in crystal structures, generating a minGFP scaffold to probe the effect of residues adjacent to a folded domain on ClpAP unfolding and degradation (Ormö et al., 1996; Yang et al., 1996; Bell et al., 2019). The variable sequence between the minGFP domain and ssrA tag consisted of Pro-Pro, Ala-Ala, or Ala-Ala-Ala-Pro-Pro, in which the Pro-Pro is farther away from the beginning of the minGFP structured region.

We determined rates of steady-state ClpAP degradation of different concentrations of our minGFP variants (**Figure 4.2b**). Notably, minGFP-PP-ssrA was protected from ClpAP proteolysis, whereas minGFP-AA-ssrA and minGFP-AAAPP-ssrA were degraded efficiently and with similar kinetic parameters. Although minGFP-AAAPP-ssrA contains the Pro-Pro motif, these residues alone are clearly not sufficient to inhibit ClpAP degradation. Analogous to the location of Pro^{24}-Pro^{25} in the ClpS core junction, protection of ssrA-tagged minGFP requires placement of the Pro-Pro residues next to the GFP β-barrel. These data also explain why the ClpS NTE-GFP fusion protein is degraded by ClpAP as a result of the Pro-Pro motif being too far from the GFP folded structure (Román-Hernández et al., 2011). Furthermore, the protection of minGFP-PP-ssrA shows that the Pro-Pro sequence can shield the adjacent minGFP domain from ClpAP degradation in the C-to-N direction, the opposite polarity of the Pro-Pro sequence in native ClpS.

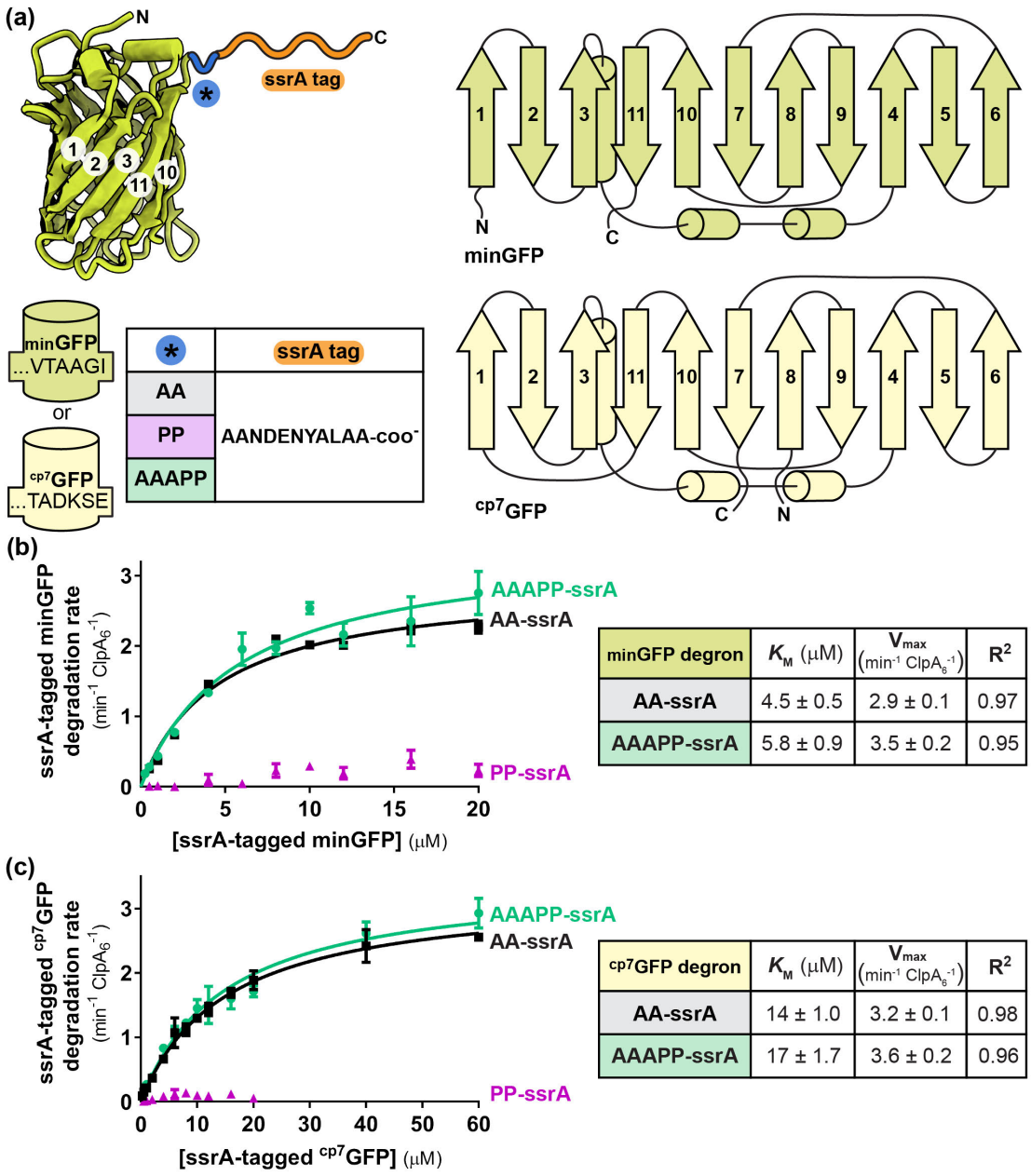


Figure 4.2 Pro-Pro adjacent to minGFP inhibits ClpAP degradation. (a) Design of ssrA-tagged minGFP and ^{cp7}GFP substrates. minGFP substrates contained residues 1–229 *A. victoria* GFP (PDB 1GFL) with S2R, S65G, S72A, and R80Q mutations, followed by a variable sequence indicated by the blue asterisk and ssrA tag. ^{cp7}GFP substrates were composed of ^{cp7}GFP that ends with the β7 sequence, TADKSE, the blue asterisk sequence, and the ssrA tag. Secondary structure cartoons of minGFP and ^{cp7}GFP are shown on the right. (b) Kinetic analysis of ClpAP degradation of ssrA-tagged minGFP variants. (c) Kinetic analysis of ClpAP degradation of ssrA-tagged ^{cp7}GFP variants. For (b) and (c), values are mean degradation rates (min^{-1} ClpA$_6^{-1}$) of triplicates with ± 1 SD error bars. K_M and V_{max} values were obtained by non-linear least squares fitting to the Michaelis-Menten equation ± fitting errors. Degradation rates of minGFP-PP-ssrA or ^{cp7}GFP-PP-ssrA were extremely slow and could not be fit to the Michaelis-Menten equation.

We also appended the PP-ssrA, AA-ssrA, and AAAPP-ssrA sequences to the C-terminus of ^{cp7}GFP (**Figure 4.2c**), a circularly permutated variant of superfolder GFP that ends with the seventh of the eleven β strands of GFP and follows a different and less stable unfolding pathway (Nager et al., 2011). Because degradation by AAA+ proteases initiates from the degron (Lee et al., 2001; Olivares et al., 2017), ClpAP encounters different local structures when initiating unfolding of the tagged variants of minGFP and ^{cp7}GFP. Nevertheless, degradation assays of the ^{cp7}GFP variants revealed the same trends observed using minGFP. Only ^{cp7}GFP-PP-ssrA resisted degradation by ClpAP, whereas and ^{cp7}GFP-AA-ssrA and ^{cp7}GFP-AAAPP-ssrA had similar V_{max} and K_M values. These results indicate that Pro-Pro abutting different regions of GFP structure is sufficient to inhibit ClpAP unfolding and degradation.

Pro-Pro acts in *cis* to protect ssrA-tagged GFP from ClpAP and ClpXP degradation

We mixed minGFP-AA-ssrA and minGFP-PP-ssrA in different ratios to determine if the PP-bearing substrate inhibited ClpAP and thus protected minGFP-AA-ssrA from degradation. ClpAP degradation rates of mixed minGFP proteins were proportional to the fraction of AA-ssrA and PP-ssrA tagged substrates present in the reaction (**Figure 4.3a**). Thus, minGFP-PP-ssrA appears to act as a competitive inhibitor of the degradation of minGFP-AA-ssrA, as expected if minGFP-PP-ssrA binds reversibly to ClpAP but is not degraded. We next measured the degradation rates of minGFP-AA-ssrA and minGFP-PP-ssrA by ClpXP, which followed the same pattern (**Figure 4.3b**). Both ClpAP and ClpXP presumably release bind and release minGFP-PP-ssrA, while minGFP-AA-ssrA is selectively degraded in the mixtures. Conversely, if minGFP-PP-ssrA was bound but not released, the protease would be stalled, such that degradation rates of mixed minGFP proteins would be the same as the reaction containing only minGFP-PP-ssrA. We conclude that the Pro-Pro motif is a *cis*-acting protective element against degradation by ClpAP and ClpXP.

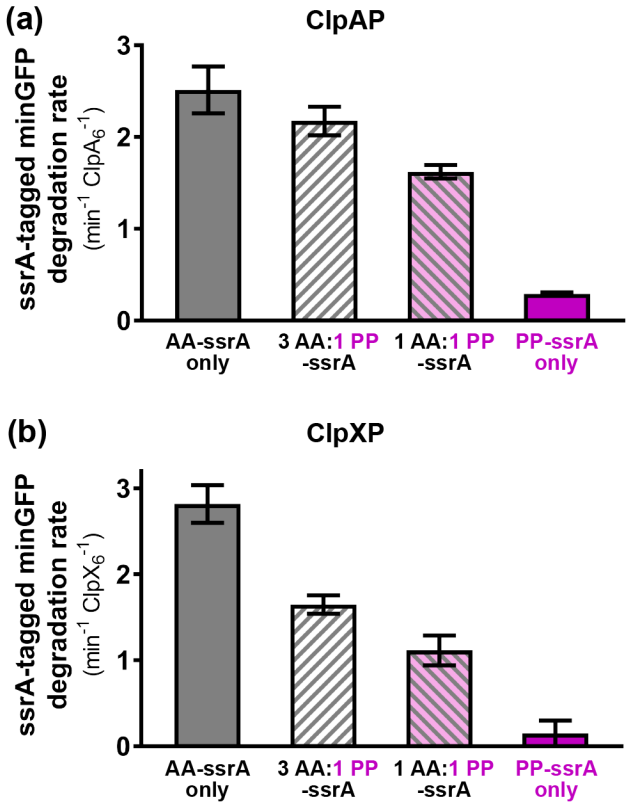


Figure 4.3 Pro-Pro protects ssrA-tagged minGFP from ClpAP and ClpXP degradation. (a) Degradation rates of 20 μM minGFP-AA-ssrA, 20 μM minGFP-PP-ssrA, or mixtures of AA-ssrA or PP-ssrA variants by ClpAP or (b) by ClpXP with the same total amount of minGFP protein (20 μM). For (a) and (b), values are mean degradation rates (min^{-1} $ClpA_6^{-1}$ or min^{-1} $ClpX_6^{-1}$) of triplicates with error bars of ± 1 SD.

Pro-Pro does not protect ssrA-tagged λ cI^N from degradation by ClpAP

We used the N-terminal domain of λ cI repressor (cI^N) as a metastable protein domain (Huang and Oas, 1995). We monitored degradation of cI^N-PP-ssrA, cI^N-AA-ssrA, and cI^N-AAAPP-ssrA fusions over time. λ cI^N-PP-ssrA did not resist ClpAP degradation, and less than 20% remained after 20 minutes. ClpAP degraded cI^N-AAAPP-ssrA at about the same rate as cI^N-PP-ssrA and degraded cI^N-AA-ssrA ~two-fold faster (**Figure 4.4**). Given the rapid equilibration between native and denatured λ cI^N in solution (Huang and Oas, 1995), we attribute degradation of λ cI^N-PP-ssrA to its metastable structure. λ cI^N dimerizes weakly in solution (K_D ~0.6 mM) (Weiss et al., 1987) and would be predominantly monomeric under our degradation conditions. Thus, the C-terminal

α-helix of λ cI^N (res. 79–92) may not adopt the folded conformation observed in λ cI^N dimers (Pabo and Lewis, 1982), placing the PP sequence farther from the native domain.

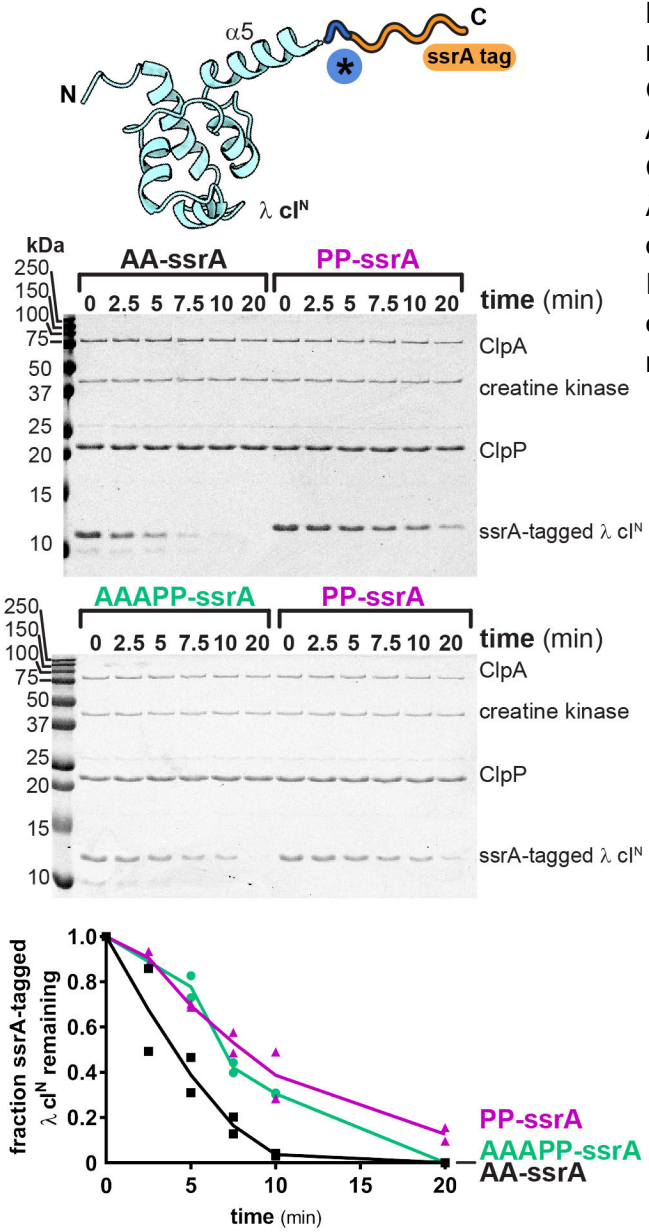


Figure 4.4 λ cI^N bearing PP-ssrA does not resist degradation by ClpAP.
Gel degradation assays of λ cI^N fused to AA/PP/AAAPP and the ssrA tag (15 μM) by ClpAP. Top panel shows design of ssrA-tagged λ cI^N (res. 1–92) fusion proteins based on the dimeric structure of λ cI^N (PDB 1LRP, chain A). In graph below, values are from individual experiments with the line representing the mean of duplicates.

Pro-Pro-tagged Rop, an α-helical protein, also resists ClpAP degradation

Both ClpS and our GFP constructs are predominantly formed from antiparallel β strands, raising the possibility that resistance by Pro-Pro requires a β-sheet structure. We generated fusion proteins using dimeric Rop, which lacks β strands but is both thermodynamically and kinetically stable (Banner et al., 1987; Munson et al., 1994, 1996, 1997; Dalal et al., 2008). Degradation of Rop-PP-ssrA over 30 minutes was significantly slower than either Rop-AA-ssrA (~10-fold) or Rop-AAAPP-ssrA (~3 fold) (**Figure 4.5**). Thus, we conclude the Pro-Pro motif abutting a stable folded domain, composed of α-helices or β strands, is necessary and sufficient to protect a degron-tagged protein from ClpAP proteolysis.

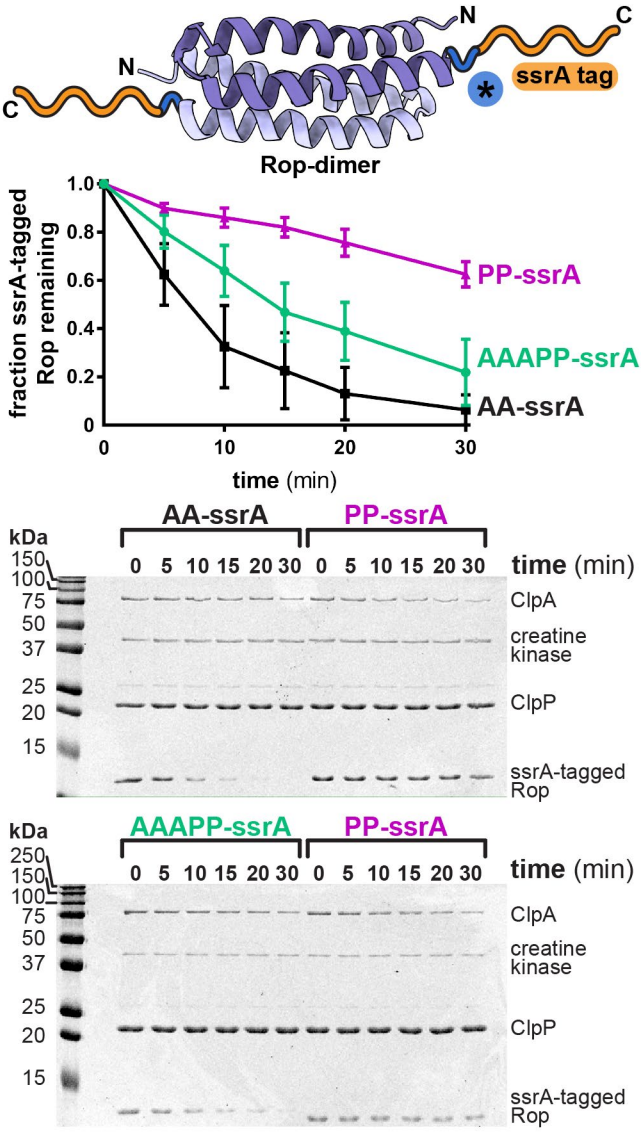


Figure 4.5 Pro-Pro protects ssrA-tagged Rop from ClpAP degradation.
ClpAP degradation of dimeric Rop fused to AA-ssrA, PP-ssrA, or AAAPP-ssrA (20 μM). The top panel shows the design of ssrA-tagged Rop (res. 1–56) fusion proteins based on the NMR structure (PDB 1RPR). In the graph below, values are mean fractions of Rop-ssrA variants, comparing to the Rop band density at 0 min, (n=3) with error bars of ± 1 SD. Representative gels are shown in the panels below.

In summary, we designed a panel of ssrA-tagged fusion proteins with PP-ssrA, AA-ssrA, or AAAPP-ssrA adjoining the folded domains of GFP, λ cI^N, or Rop. For GFP and Rop, the adjacent Pro-Pro motif antagonized ClpAP degradation and functions as an 'anti-degron'. For λ cI^N-PP-ssrA, however, the Pro-Pro sequence did not evidently protect against ClpAP degradation, either because the cI^N domain is metastable and/or because α-helix 5 of this domain unfolds partially in the monomeric protein, thus moving the Pro-Pro sequence too far from a folded structure. Our results show that a Pro-Pro dipeptide placed within a few residues of a stably folded structure is sufficient to protect a degron-bearing protein from both ClpAP and ClpXP. Pro^{24}-Pro^{25} in native ClpS neutralizes the degron-like properties of the rest of the NTE, allowing the ClpS adaptor to escape ClpAP degradation despite binding of the NTE in the ClpAP axial channel in the same manner as a substrate.

The Pro-Pro protective element contrasts with other mechanisms shielding adaptors and protease-associated proteins from degradation. Similar to ClpS, the prokaryotic ubiquitin-like protein, Pup, docks with the N-domains of the mycobacterial proteasome Mpa and is translocated by its cognate AAA+ protease (Sutter et al., 2009; Striebel et al., 2010; Kavalchuk et al., 2022). However, unlike ClpS, Pup can escape degradation via recycling, analogous to ubiquitin in the eukaryotic proteasome, or perhaps by passing through the peptidase due to a lack of favorable proteolytic cleavage sites, as proposed recently (Zerbib et al., 2021). Rad23, another adaptor refractory to degradation, binds to the eukaryotic proteasome, but lacks an unstructured initiation region that can be engaged effectively by the unfoldase machinery (Fishbain et al., 2011).

By contrast, ClpAP binds and engages the ClpS NTE, using the energy of ATP hydrolysis to conformationally remodel the ClpS adaptor during N-end-rule substrate delivery (Román-Hernández et al., 2011; Rivera-Rivera et al., 2014). Thus, ClpS is a 'pseudo-substrate' protected by the Pro-Pro motif abutting the ClpS core domain. Whether resistance to ClpAP degradation due to the Pro-Pro motif or by other means is conserved among all ClpS orthologs remains an

open question. We hypothesize that the Pro-Pro sequence may interact with AAA+ proteases to impair processivity, as observed for ‘slippery’ substrates (Lin and Ghosh, 1996; Levitskaya et al., 1997; Sharipo et al., 2001; Hoyt et al., 2006; Daskalogianni et al., 2008; Kraut et al., 2012; Kraut, 2013; Too et al., 2013; Vass and Chien, 2013; Bell et al., 2019). The conformational rigidity of proline may preclude efficient unfolding and/or translocation, as suggested by the slow translocation of poly-proline by ClpXP (Barkow et al., 2009). Biophysical studies using single-molecule optical trapping may help elucidate the mechanism of Pro-Pro protection against global unfolding by AAA+ enzymes and determine whether Pro-Pro enables protein remodeling rather than unfolding and degradation.

MATERIALS AND METHODS

Protein expression and purification

ClpA was expressed and purified by established protocols (Hou et al., 2008) and stored in HO buffer (50 mM HEPES-KOH [pH 7.5], 300 mM NaCl, 20 mM $MgCl_2$, 10% [w/v] glycerol, 0.5 mM DTT). ClpX was expressed and purified as described (Neher et al., 2003) and stored in 50 mM Tris-HCl [pH 8.2 at 4 °C], 100 mM KCl, 1 mM $MgCl_2$, 10% (w/v) glycerol, 5 mM DTT. ClpP was expressed in *E. coli* strain JK10 (*clpP::cat*, *Δlon*, *slyD::kan*, λDE3), purified as previously described, and stored in 50 mM Tris-HCl (pH 8) 150 mM KCl, 10% (w/v) glycerol, 0.5 mM EDTA, and 1 mM DTT (Kim et al., 2000; Kenniston et al., 2003). To generate minGFP variants, sequences encoding AA, PP, or AAAPP were introduced into pT7 GFP-AANDENYALAA (Fei et al., 2020), which encodes GFP carrying S2R, S65G, S72A, and R80Q mutations, using round-the-horn mutagenesis with T4 polynucleotide kinase and Q5 high-fidelity polymerase (New England Biolabs). minGFP fusion proteins were expressed in *E. coli* strain T7 Express (New England Biolabs) and purified by Ni-NTA chromatography, Superdex-75 size-exclusion chromatography, and Source-15Q anion-exchange chromatography. To generate ^{cp7}GFP variants, the sequence encoding ^{cp7}GFP-ssrA was ligated to the 3´-end of His_6-Sumo cloned into pET23b, followed by introduction of sequences encoding AA, PP, or AAAPP before the ssrA tag via round-the-horn mutagenesis (Nager et al., 2011). ^{cp7}GFP, λ cI^N, and Rop fusion proteins were expressed in *E. coli* strain T7 Express and purified as described for His_6-SUMO-fusion proteins (Hou et al., 2008; Román-Hernández et al., 2011). All fusion proteins were stored in 25 mM HEPES-KOH (pH 7.5), 150 mM KCl, 10% (w/v) glycerol, and 1 mM DTT.

Degradation of GFP-ssrA variants

Degradation of GFP-ssrA variants was performed in HO buffer for ClpAP or PD buffer (25 mM HEPES-KOH [pH 7.5], 100 mM KCl, 5 mM $MgCl_2$, 10% [w/v] glycerol, 0.032% Tween-20, and 0.2

mM DTT) for ClpXP with at least three technical replicates at 30 °C using a SpectraMax M5 Microplate Reader (Molecular Devices) to measure initial rates over 5 minutes. Degradation of ssrA-tagged minGFP or ^{cp7}GFP variants was monitored by loss of fluorescence (excitation 467 nm, emission 511 nm). Degradation rates were calculated by normalizing the slope values of relative fluorescence units (RFUs)/time by the fluorescence signal determined from a linear standard curve of RFUs versus increasing GFP concentration. Reactions contained varying GFP concentrations, $ClpA_6$ (0.2 μM), $ClpP_{14}$ (0.4 μM), and an ATP-regeneration system (5 mM ATP, 50 μg/mL creatine kinase, and 5 mM creatine phosphate) for Michaelis-Menten analyses. For Figure 4.3, mixing experiments with minGFP-AA-ssrA and minGFP-PP-ssrA, in the ratios indicated to a total concentration of 20 μM minGFP variant(s), were performed under similar conditions using $ClpA_6$ (0.25 μM) or $ClpX_6$ (0.25 μM), $ClpP_{14}$ (0.75 μM), and the ATP-regeneration system described above.

Degradation of ssrA-tagged λ cI^N and Rop variants

Degradation reactions containing λ cI^N (15 μM) or Rop (20 μM) fusion proteins, 0.1 μM $ClpA_6$, 0.2 μM $ClpP_{14}$, and an ATP-regeneration system (5 mM ATP, 50 μg/mL creatine kinase, and 5 mM creatine phosphate) was monitored by SDS-PAGE. Gel degradation of λ cI^N fusion proteins (15 μM) was performed in duplicate. Samples of each reaction at specific time points were stopped by addition of SDS-PAGE loading sample and boiled at 100 °C prior to loading on Tris-Glycine-SDS gels. Bands were visualized by colloidal Coomassie G-250, scanned on a Typhoon FLA 9500 instrument (GE Healthcare) and quantified using ImageQuant (GE Healthcare). The fraction of fusion protein remaining was calculated by dividing the density of the ssrA-tagged protein band at a given time point by the density at time zero, after normalizing loading in each lane using the ClpP band density.

Data visualization

Figures including structures were generated using ChimeraX (Pettersen et al., 2021). Biochemical assay data were plotted in GraphPad Prism.

ACKNOWLEDGMENTS

This work was supported by NIH grant AI-016892 (R.T.S., T.A.B.), the Howard Hughes Medical Institute (T.A.B.), and the National Science Foundation Graduate Research Fellowship grant 1745302 (S.K.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

Banner, D.W., Kokkinidis, M., and Tsernoglou, D. (1987). Structure of the ColE1 Rop protein at 1.7 Å resolution. J. Mol. Biol. *196*, 3, 657–675.

Barkow, S.R., Levchenko, I., Baker, T.A., and Sauer, R.T. (2009). Polypeptide translocation by the AAA+ ClpXP protease machine. Chem. Biol. *16*, 6, 605.

Bell, T.A., Baker, T.A., and Sauer, R.T. (2019). Interactions between a subset of substrate side chains and AAA+ motor pore loops determine grip during protein unfolding. eLife *8*, e46808.

Dalal, S., Canet, D., Kaiser, S.E., Dobson, C.M., and Regan, L. (2008). Conservation of mechanism, variation of rate: Folding kinetics of three homologous four-helix bundle proteins. Protein Eng. Des. Sel. *21*, 3, 197–206.

Daskalogianni, C., Apcher, S., Candeias, M.M., Naski, N., Calvo, F., and Fåhraeus, R. (2008). Gly-Ala repeats induce position- and substrate-specific regulation of 26S proteasome-dependent partial processing. J. Biol. Chem. *283*, 44, 30090–30100.

De Donatis, G.M., Singh, S.K., Viswanathan, S., and Maurizi, M.R. (2010). A single ClpS monomer is sufficient to direct the activity of the ClpA hexamer. J. Biol. Chem. *285*, 12, 8771–8781.

Dougan, D.A., Reid, B.G., Horwich, A.L., and Bukau, B. (2002). ClpS, a substrate modulator of the ClpAP machine. Mol. Cell *9*, 3, 673–683.

Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D.A., and Bukau, B. (2006). ClpS is an essential component of the N-end rule pathway in *Escherichia coli*. Nature *439*, 7077, 753–756.

Fei, X., Bell, T.A., Barkow, S.R., Baker, T.A., and Sauer, R.T. (2020). Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. eLife *9*, e61496, 1–39.

Fishbain, S., Prakash, S., Herrig, A., Elsasser, S., and Matouschek, A. (2011). Rad23 escapes degradation because it lacks a proteasome initiation region. Nat. Commun. *2*, 1, 1–9.

Gottesman, S., Roche, E., Zhou, Y., and Sauer, R.T. (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the ssrA-tagging system. Genes Dev. *12*, 9, 1338–1347.

Guo, F., Esser, L., Singh, S.K., Maurizi, M.R., and Xia, D. (2002). Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA. J. Biol. Chem. *277*, 48, 46753–46762.

Hou, J.Y., Sauer, R.T., and Baker, T.A. (2008). Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. Nat. Struct. Mol. Biol. *15*, 3, 288–294.

Hoyt, M.A., Zich, J., Takeuchi, J., Zhang, M., Govaerts, C., and Coffino, P. (2006). Glycine-alanine repeats impair proper substrate unfolding by the proteasome. EMBO J. *25*, 8, 1720–1729.

Huang, G.S., and Oas, T.G. (1995). Structure and stability of monomeric λ repressor: NMR evidence for two-state folding. Biochemistry *34*, 12, 3884–3892.

Kavalchuk, M., Jomaa, A., Müller, A.U., and Weber-Ban, E. (2022). Structural basis of prokaryotic ubiquitin-like protein engagement and translocation by the mycobacterial Mpa-proteasome complex. Nat. Commun. *13*, 1, 1–13.

Kenniston, J.A., Baker, T.A., Fernandez, J.M., and Sauer, R.T. (2003). Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. Cell *114*, 4, 511–520.

Kenniston, J.A., Burton, R.E., Siddiqui, S.M., Baker, T.A., and Sauer, R.T. (2004). Effects of local protein stability and the geometric position of the substrate degradation tag on the efficiency of ClpXP denaturation and degradation. J. Struct. Biol. *146*, pp. 130–140.

Kim, Y.I., Burton, R.E., Burton, B.M., Sauer, R.T., and Baker, T.A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. Mol. Cell *5*, 4, 639–648.

Koodithangal, P., Jaffe, N.E., Kraut, D.A., Fishbain, S., Herman, C., and Matouschek, A. (2009). ATP-dependent proteases differ substantially in their ability to unfold globular proteins. J. Biol. Chem. *284*, 28, 18674–18684.

Kraut, D.A. (2013). Slippery substrates impair ATP-dependent protease function by slowing unfolding. J. Biol. Chem. *288*, 48, 34729–34735.

Kraut, D.A., Israeli, E., Schrader, E.K., Patil, A., Nakai, K., Nanavati, D., Inobe, T., and Matouschek, A. (2012). Sequence- and species-dependence of proteasomal processivity. ACS Chem. Biol. *7*, 8, 1444–1453.

Lee, C., Schwartz, M.P., Prakash, S., Iwakura, M., and Matouschek, A. (2001). ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. Mol. Cell *7*, 3, 627–637.

Levitskaya, J., Sharipo, A., Leonchiks, A., Ciechanover, A., and Masucci, M.G. (1997). Inhibition of ubiquitin/proteasome-dependent protein degradation by the Gly-Ala repeat domain of the Epstein-Barr virus nuclear antigen 1. Proc. Natl. Acad. Sci. U. S. A. *94*, 23, 12616–12621.

Lin, L., and Ghosh, S. (1996). A glycine-rich region in NF-kappaB p105 functions as a processing signal for the generation of the p50 subunit. Mol. Cell. Biol. *16*, 5, 2248–2254.

Mahmoud, S.A., and Chien, P. (2018). Regulated proteolysis in bacteria. Annu. Rev. Biochem. *87*, 677–696.

Maillard, R.A., Chistol, G., Sen, M., Righini, M., Tan, J., Kaiser, C.M., Hodges, C., Martin, A., and Bustamante, C. (2011). ClpX(P) generates mechanical force to unfold and translocate its protein substrates. Cell *145*, 3, 459–469.

Martin, A., Baker, T.A., and Sauer, R.T. (2008). Protein unfolding by a AAA+ protease is dependent on ATP-hydrolysis rates and substrate energy landscapes. Nat. Struct. Mol. Biol. *15*, 2, 139–145.

Matouschek, A. (2003). Protein unfolding — an important process in vivo? Curr. Opin. Struct. Biol. *13*, 1, 98–109.

Matouschek, A., Huang, S., Ratliff, K.S., Schwartz, M.P., and Spenner, J.M. (1999). Mitochondria unfold precursor proteins by unraveling them from their N-termini. Nat. Struct. Biol. *6*, 12, 1132–1138.

Munson, M., Regan, L., O'Brien, R., and Sturtevant, J.M. (1994). Redesigning the hydrophobic core of a four-helix-bundle protein. Protein Sci. *3*, 11, 2015–2022.

Munson, M., Balasubramanian, S., Fleming, K.G., Nagi, A.D., O'Brien, R., Sturtevant, J.M., and Regan, L. (1996). What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. Protein Sci. *5*, 8, 1584–1593.

Munson, M., Anderson, K.S., and Regan, L. (1997). Speeding up protein folding: Mutations that increase the rate at which Rop folds and unfolds by over four orders of magnitude. Fold. Des. *2*, 1, 77–87.

Nager, A.R., Baker, T.A., and Sauer, R.T. (2011). Stepwise unfolding of a β barrel protein by the AAA+ ClpXP protease. J. Mol. Biol. *413*, 1, 4–16.

Neher, S.B., Sauer, R.T., and Baker, T.A. (2003). Distinct peptide signals in the UmuD and UmuD′ subunits of UmuD/D′ mediate tethering and substrate processing by the ClpXP protease. Proc. Natl. Acad. Sci. U. S. A. *100*, 23, 13219–13224.

Olivares, A.O., Nager, A.R., Iosefson, O., Sauer, R.T., and Baker, T.A. (2014). Mechanochemical basis of protein degradation by a double-ring AAA+ machine. Nat. Struct. Mol. Biol. *21*, 10, 871–875.

Olivares, A.O., Kotamarthi, H.C., Stein, B.J., Sauer, R.T., and Baker, T.A. (2017). Effect of directional pulling on mechanical protein degradation by ATP-dependent proteolytic machines. Proc. Natl. Acad. Sci. *114*, 31, E6306–E6313.

Ormö, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., and Remington, S.J. (1996). Crystal structure of the Aequorea victoria green fluorescent protein. Science (80-.). *273*, 5280, 1392–1395.

Pabo, C.O., and Lewis, M. (1982). The operator-binding domain of λ repressor: Structure and DNA recognition. Nature *298*, 5873, 443–447.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. *30*, 1, 70–82.

Rivera-Rivera, I. (2015). Mechanism of active substrate delivery by the AAA+ protease adaptor ClpS. Ph.D. thesis, Massachusetts Institute of Technology. Retrieved from https://dspace.mit.edu/handle/1721.1/101352

Rivera-Rivera, I., Román-Hernández, G., Sauer, R.T., and Baker, T.A. (2014). Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. Proc. Natl. Acad. Sci. U. S. A. *111*, 37, E3853-9.

Román-Hernández, G., Hou, J.Y., Grant, R.A., Sauer, R.T., and Baker, T.A. (2011). The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. Mol. Cell *43*, 2, 217–228.

Sauer, R.T., and Baker, T.A. (2011). AAA+ Proteases: ATP-fueled machines of protein destruction. Annu. Rev. Biochem. *80*, 1, 587–612.

Sen, M., Maillard, R.A., Nyquist, K., Rodriguez-Aliaga, P., Pressé, S., Martin, A., and Bustamante, C. (2013). The ClpXP protease unfolds substrates using a constant rate of pulling but different gears. Cell *155*, 3, 636.

Sharipo, A., Imreh, M., Leonchiks, A., Brändén, C.I., and Masucci, M.G. (2001). cis-Inhibition of proteasomal degradation by viral repeats: Impact of length and amino acid composition. FEBS Lett. *499*, 1–2, 137–142.

Singh, S.K., Grimaud, R., Hoskins, J.R., Wickner, S., and Maurizi, M.R. (2000). Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. Proc. Natl. Acad. Sci. U. S. A. *97*, 16, 8898–8903.

Striebel, F., Hunkeler, M., Summer, H., and Weber-Ban, E. (2010). The mycobacterial Mpa-proteasome unfolds and degrades pupylated substrates by engaging Pup's N-terminus. EMBO J. *29*, 7, 1262–1271.

Sutter, M., Striebel, F., Damberger, F.F., Allain, F.H.T., and Weber-Ban, E. (2009). A distinct structural region of the prokaryotic ubiquitin-like protein (Pup) is recognized by the N-terminal domain of the proteasomal ATPase Mpa. FEBS Lett. *583*, 19, 3151–3157.

Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991). The N-end rule in bacteria. Science *254*, 5036, 1374–1377.

Too, P.H.M., Erales, J., Simen, J.D., Marjanovic, A., and Coffino, P. (2013). Slippery substrates impair function of a bacterial protease ATPase by unbalancing translocation versus exit. J. Biol. Chem. *288*, 19, 13243–13257.

Torres-Delgado, A., Kotamarthi, H.C., Sauer, R.T., and Baker, T.A. (2020). The intrinsically disordered N-terminal extension of the ClpS adaptor reprograms its partner AAA + ClpAP protease. J. Mol. Biol. *432*, 17, 4908–4921.

Varshavsky, A. (2017). The ubiquitin system, autophagy, and regulated protein degradation. Annu. Rev. Biochem. *86*, 123–128.

Vass, R.H., and Chien, P. (2013). Critical clamp loader processing by an essential AAA+ protease in *Caulobacter crescentus*. Proc. Natl. Acad. Sci. U. S. A. *110*, 45, 18138–18143.

Wang, K.H., Sauer, R.T., and Baker, T.A. (2007). ClpS modulates but is not essential for bacterial N-end rule degradation. Genes Dev. *21*, 4, 403–408.

Wang, K.H., Roman-Hernandez, G., Grant, R.A., Sauer, R.T., and Baker, T.A. (2008). The molecular basis of N-end rule recognition. Mol. Cell *32*, 3, 406–414.

Weiss, M.A., Pabo, C.O., Karplus, M., and Sauer, R.T. (1987). Dimerization of the operator binding domain of phage lambda repressor. Biochemistry *26*, 3, 897–904.

Yang, F., Moss, L.G., and Phillips, G.N. (1996). The Molecular Structure of Green Fluorescent Protein. Nat. Biotechnol. *14*, 10, 1246–1251.

Zerbib, E., Schlussel, S., Hecht, N., Bagdadi, N., Eichler, J., and Gur, E. (2021). The prokaryotic ubiquitin-like protein presents poor cleavage sites for proteasomal degradation. Cell Rep. *36*, 4, 109428.

Zeth, K., Ravelli, R.B., Paal, K., Cusack, S., Bukau, B., and Dougan, D.A. (2002a). Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. Nat. Struct. Biol. *9*, 12, 906–911.

Zeth, K., Dougan, D.A., Cusack, S., Bukau, B., and Ravelli, R.B. (2002b). Crystallization and preliminary X-ray analysis of the *Escherichia coli* adaptor protein ClpS, free and in complex with the N-terminal domain of ClpA. Acta Crystallogr. Sect. D Biol. Crystallogr. *58*, 7, 1207–1210.

Zuromski, K.L., Kim, S., Sauer, R.T., and Baker, T.A. (2021). Division of labor between the pore-1 loops of the D1 and D2 AAA+ rings coordinates substrate selectivity of the ClpAP protease. J. Biol. Chem. *297*, 6, 101407.

CHAPTER 5

Exposing two protease recognition signals targets the FNR transcription factor for oxygen-regulated degradation

This chapter has been written as a manuscript for publication.

Sora Kim[1], Stephane Calmat[1,4], Lindsey R. F. Backman[2], Chi Nguyen[1], Robert T. Sauer[1], Catherine L. Drennan[1-3], and Tania A. Baker[1]

[1] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139
[2] Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139
[3] Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139
[4] Present address: SQZ Biotechnologies, Watertown, MA 02472

Author contributions:

S. Kim: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing – original draft, Writing – review and editing, Visualization, Funding acquisition

S. Calmat: Conceptualization, Methodology, Validation, Investigation, Resources, Writing – review and editing

L.R.F. Backman: Conceptualization, Methodology, Validation, Resources, Writing – review and editing, Visualization, Funding acquisition

C. Nguyen: Conceptualization, Methodology, Investigation, Resources, Writing – review and editing, Funding acquisition

R.T. Sauer, C.L. Drennan, and T.A. Baker: Conceptualization, Methodology, Validation, Formal analysis, Writing – review & editing, Supervision, Funding acquisition

ABSTRACT

The global bacterial transcriptional regulator Fumarate and Nitrate Reductase (FNR) senses environmental O_2, which destabilizes its [4Fe-4S] cluster and converts the active dimer (*holo*FNR) into an inactive monomer (*apo*FNR). The monomer but not the dimer is selectively degraded by the ClpXP protease. We find that the N-terminal domain of ClpX is required for efficient FNR degradation and binds to a short peptide near the N-terminus of FNR. By engineering GFP-FNRtag chimeras, we demonstrate that a second short C-terminal FNR sequence binds weakly to the axial pore or channel of ClpX. The N-terminal FNR tag is not an autonomous degron. However, the presence of this signal with the C-terminal tag tightens the K_M for degradation five-fold. These results indicate that the N-terminal FNR sequence functions as an 'enhancement tag' (e-tag), allowing monomeric FNR to tether itself to ClpX and providing a simple example of coupling an e-tag and weak pore-binding tag within the same polypeptide for multicomponent protease recognition. Crystallographic analysis of *apo* and *holo*FNR reveals that both the e-tag and pore tag are more solvent-exposed in the *apo* form, likely rendering them more accessible to ClpXP. Using FNR-tag variants that are dimeric in the presence of O_2 but lack [4Fe-4S] clusters, we show that dimerization is sufficient to inhibit ClpXP proteolysis by masking the pore tag. Thus, oxygen-sensitive exposure of the dual e-tag and pore-binding signals controls degradation of an important cellular regulator specifically during aerobic growth.

Keywords

ATPases associated with diverse cellular activities; protein degradation; redox signaling; enhancement tag

INTRODUCTION

Molecular oxygen, the terminal electron acceptor during cellular respiration, is required for aerobic life but can also become toxic and give rise to reactive oxygen species. Oxygen deprivation or hypoxia results in reduced ATP production by oxidative phosphorylation, leading to metabolic adaptation or cell death (Harris, 2002; Boshoff and Barry, 2005). Hyperoxia, occurring when oxygen levels exceed atmospheric pressure, promotes the formation of reactive oxygen species that damage nucleic acids, proteins, and lipids, thereby inhibiting growth and threatening cell viability (Baez and Shiloach, 2014). To maintain oxygen homeostasis, various oxygen-sensing systems have evolved in all kingdoms of life to protect organisms from changes in oxygen concentration. Facultative anaerobes, such as *Escherichia coli*, use redox-sensitive mechanisms to shift between aerobic and anaerobic metabolism in response to oxygen availability. This ability to physiologically adapt between high and low oxygen environments is critical during pathogenesis and contributes to biofilm formation (Marteyn et al., 2011). In *E. coli*, for example, the transcription factor Fumarate and Nitrate Reductase (FNR) directly senses O_2 and controls the transcription of hundreds of genes, including those required for anaerobic respiration (Salmon et al., 2003; Kang et al., 2005). FNR deletion strains exhibit attenuated virulence, such as defects in mouse intestinal colonization, demonstrating the importance of oxygen sensing for microbial adaptation during the infection of a eukaryotic host (Jones et al., 2007; Barbieri et al., 2014, 2017).

FNR is a member of the cAMP receptor protein CRP/FNR superfamily of transcriptional regulators. These usually dimeric proteins contain an N-terminal sensory domain, which responds to environmental conditions, and a C-terminal DNA-binding domain, which controls transcription of its regulon by binding specific DNA sequences (Mettert and Kiley, 2018). Under anaerobic conditions, the FNR sensory domain binds an oxygen-labile [4Fe-4S] cluster that promotes and stabilizes FNR dimers. *Holo*FNR is the active form that carries out site-specific DNA binding and interacts with RNA polymerase (Lazazzera et al., 1993, 2002; Moore and Kiley, 2001; Moore et

al., 2006; Jervis and Green, 2007). Upon oxygen exposure, destruction of the [4Fe-4S] cluster inactivates the tight DNA binding activity of the FNR dimer by promoting dissociation to the monomeric *apo*FNR form, which is then degraded (Sutton et al., 2004; Mettert and Kiley, 2005).

Degradation of *apo*FNR depends on ClpXP, a AAA+ (ATPases associated with diverse cellular activities) protease, which selectively recognizes *apo*FNR monomers via sequences located near the protein's N- and C-termini (Flynn et al., 2003; Mettert and Kiley, 2005; Pan et al., 2012). ClpXP consists of the AAA+ $ClpX_6$ unfoldase/translocase and the self-compartmentalized $ClpP_{14}$ peptidase. AAA+ proteases perform energy-dependent proteolysis of damaged, deleterious, or unneeded proteins to support protein homeostasis in all kingdoms of life (Sauer and Baker, 2011; Baker and Sauer, 2012; Olivares et al., 2015; Mahmoud and Chien, 2018). To perform targeted protein degradation, the AAA+ unfoldase/translocase binds and engages an unstructured peptide sequence on the substrate protein called a degron, which is often a pore-binding tag. Pore loops lining the axial channel of the AAA+ protease interact with the degron, and ATP hydrolysis drives conformational changes in the AAA+ enzyme that power unfolding of the substrate and translocation of the denatured polypeptide into the peptidase for destruction. Regulation of protein degradation by AAA+ proteases occurs in many ways: (i) altering degron accessibility; (ii) co- or post-translational modifications of the substrate; and (iii) by using secondary recognition signals and/or adaptor proteins that interact with auxiliary domains of the AAA+ enzyme to enhance substrate affinity (**Figure 5.1**) (Sauer and Baker, 2011; Olivares et al., 2015; Mahmoud and Chien, 2018).

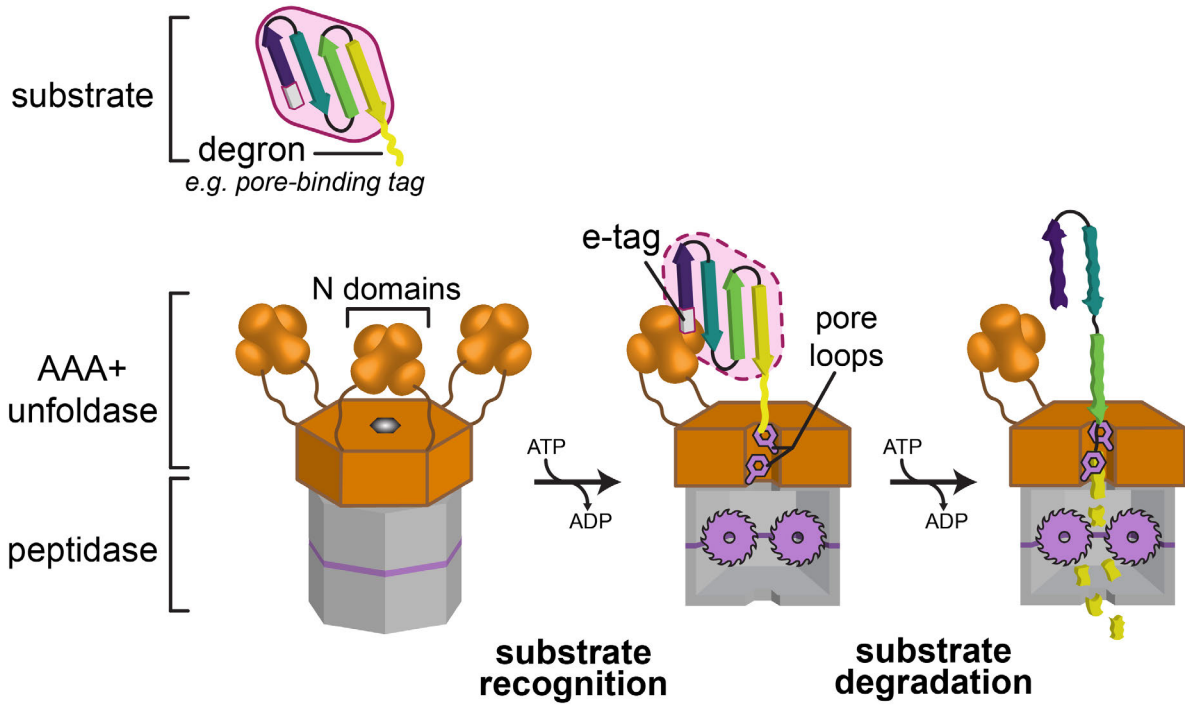


Figure 5.1 Mechanism of substrate recognition and degradation by AAA+ proteases. AAA+ proteases consist of a AAA+ unfoldase and partner peptidase. Recognition of a protein substrate is mediated through binding of a degron (usually a pore-binding tag) within the substrate to the axial pore of the AAA+ unfoldase and can involve auxiliary domains on the AAA+ enzyme. The ClpX N domains, which form dimers, are represented here interacting with the enhancement tag (e-tag) in another region of the substrate. Pore loops within the unfoldase channel bind and engage the degron, allowing for processive ATP-powered unfolding and translocation into the peptidase chamber for degradation (rightmost panel). Only one set of N domain dimers is shown for clarity during the substrate recognition and degradation steps.

Here, we investigate how ClpXP recognizes the two recognition signals of FNR under aerobic conditions and how dimerization protects FNR from ClpXP proteolysis. We report that the N-terminal signal of FNR tethers the substrate to the N domain of ClpX and serves as an enhancement tag (e-tag) to stabilize the enzyme•substrate complex and enhance recognition of FNR's weak C-terminal pore-binding tag. Using dimeric *apo*FNR mutant proteins and comparisons of *apo*FNR and *holo*FNR crystal structures, we find that dimerization masks these recognition signals to protect FNR dimers from ClpXP proteolysis. Our studies of *apo*FNR recognition by ClpXP demonstrate how environmentally-sensitive conformational changes in a master regulatory protein control its homeostasis by coupling protein function to protein turnover.

RESULTS

FNR degradation depends on the ClpX N domain

The N domain of ClpX binds peptide sequences in adaptor proteins and some substrates, often serving as a tethering site to enhance ClpXP substrate recognition (Sauer et al., 2021). To investigate whether direct recognition of *apo*FNR requires the N domain of ClpX, we compared *apo*FNR degradation by ClpXP and ClpX$^{\Delta N}$P, where ClpX$^{\Delta N}$ is a variant lacking the N domain. ClpX$^{\Delta N}$P degrades substrates that do not require binding to the N domain, such as Arc-ssrA, with an affinity and rate indistinguishable from that of wild-type ClpXP (Neher et al., 2003). We found, however, that ClpX$^{\Delta N}$P degraded *apo*FNR poorly compared to ClpXP, suggesting that the N terminal domain of ClpX assists recognition and/or processing of this substrate (**Figure 5.2A**).

As another method of testing the involvement of the ClpX N domain in *apo*FNR recognition, we used a synthetic peptide from the SspB adaptor containing the XB sequence that directly binds to the N domain of ClpX. Addition of the XB peptide (Y-RGGRPALRVVK, residues 155-165 of native SspB) inhibited FNR degradation (**Figure 5.2B**). Interestingly, FNR degradation was not completely blocked by removing the ClpX N domain or by the addition of SspB peptide, indicating the presence of two ClpX-recognition signals in *apo*FNR, as established previously (Flynn et al., 2003; Mettert and Kiley, 2005; Pan et al., 2012). Nonetheless, the substantial negative effect of blocking *apo*FNR contacts with the ClpX N domain demonstrates the importance of FNR interactions with this auxiliary domain. These results thus support a model in which *apo*FNR tethers itself to the ClpX N domain to boost the efficiency of degradation using one signal, while another interacts with a distinct region of ClpX.

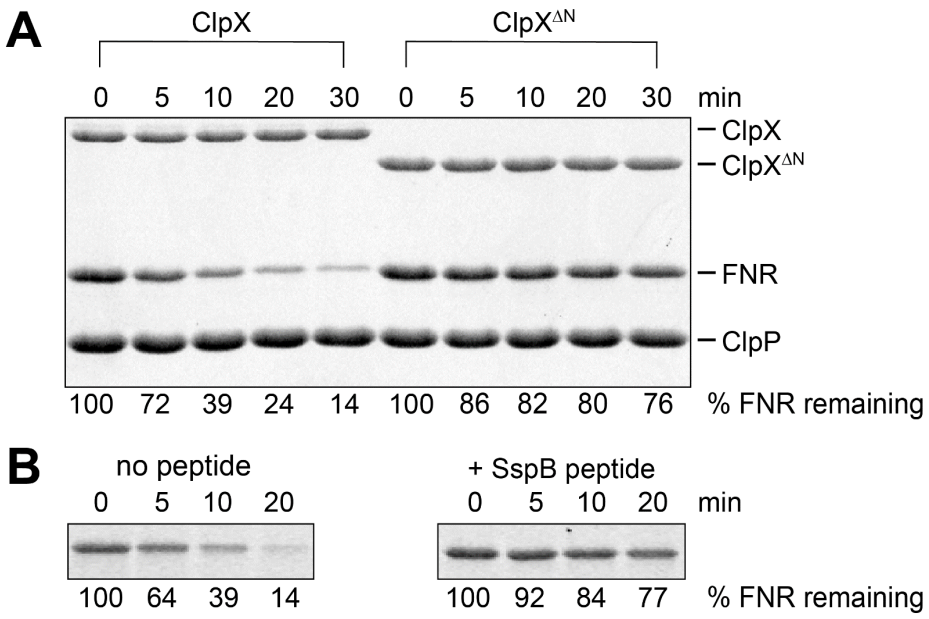


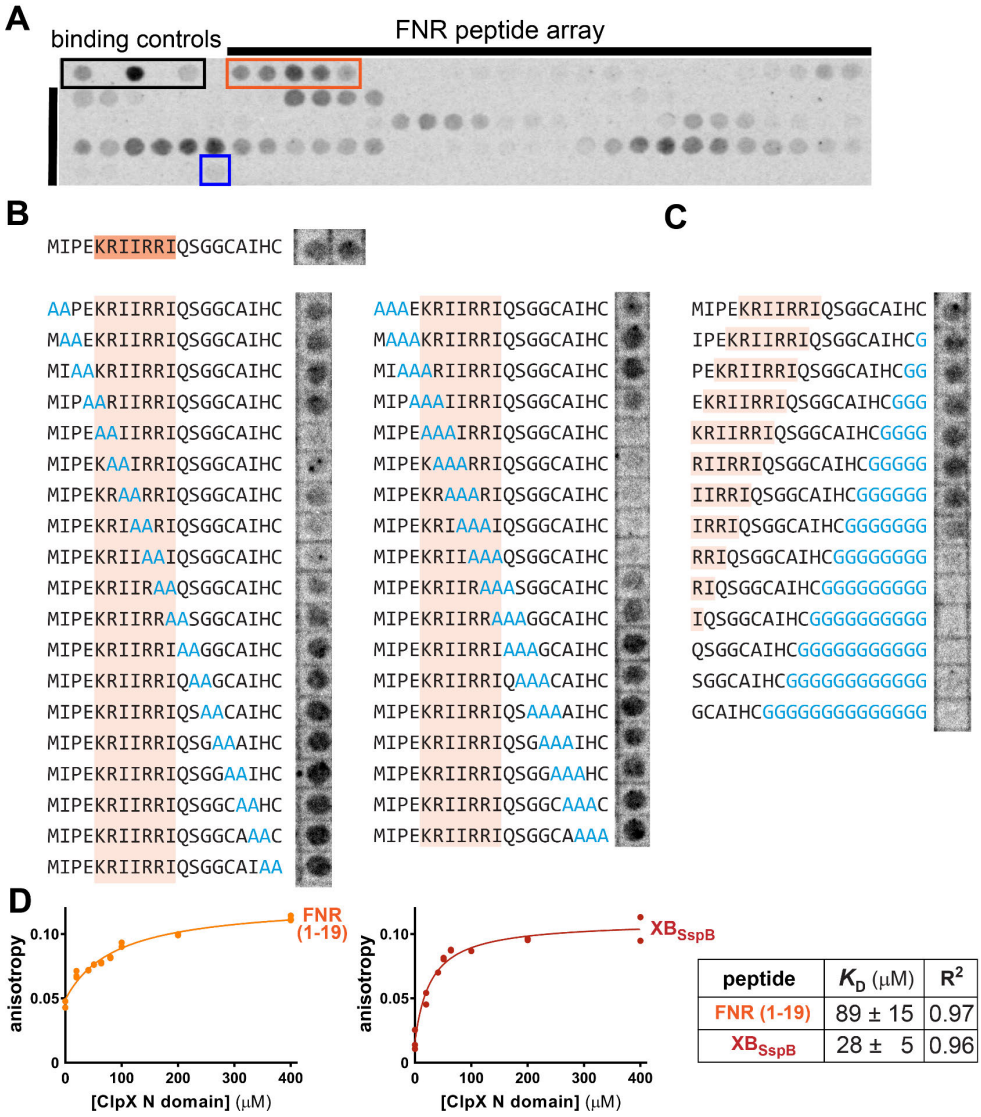
Figure 5.2 FNR requires the ClpX N domain for efficient degradation.
(A) Kinetics of *apo*FNR degradation by ClpXP or ClpX$^{\Delta N}$P monitored by SDS-PAGE and colloidal Coomassie staining. Gel degradation reactions contained *apo*FNR (5 μM), ClpX or ClpX$^{\Delta N}$ (0.3 μM hexamer), $ClpP_{14}$ (0.8 μM), and 5 mM ATP. **(B)** Kinetics of *apo*FNR degradation showing inhibition by the presence of equimolar SspB XB peptide (YRGGRPALRVVK). Assay conditions were identical to those in (A).

The ClpX N domain binds to a short sequence near the N-terminus of FNR

To identify sequences in FNR bound by the ClpX N domain, we incubated purified ^{35}S-labeled ClpX N domain with an array of overlapping 12-residue peptides corresponding to the full sequence of FNR (**Figure 5.3A**). The spot-blot pattern revealed N domain-binding to an FNR sequence of seven amino acids (KRIIRRI, residues 5–11), with the underlined resides sharing some chemical similarity with the LRVVK sequence in the SspB XB motif that is an important determinant of substrate delivery to ClpXP (Chowdhury et al., 2010). The FNR C-terminal recognition signal (NVA, residues 248–250) did not bind the ClpX N domain, consistent with the hypothesis that it interacts with a distinct region of ClpX, most likely the axial pore (see next section). Other peptides showing binding to ClpX N domain correspond to regions in FNR that are buried in the native protein or are located within the DNA-binding domain based on structural analysis (C. Nguyen *et al.*, personal communication). These internal sequences would not be

freely accessible to ClpXP in native structures and thus are unlikely to be important for specific ClpX-FNR recognition. To test this hypothesis, we focused on Arg^{179}-Arg^{180}, which are solvent inaccessible in the folded protein, but peptides containing this sequence bound most tightly to the ClpX N domain in the blot (see spots in row above the blue box in **Figure 5.3A**). Degradation rates of an Ala^{179}-Ala^{180} FNR variant was similar to the wild-type (less than 10% different, data not shown), establishing that the identity of these internal residues is not critical for robust recognition or degradation by ClpXP.

We further dissected the contribution to ClpX N domain-binding of individual residues in the N-terminal 20 residues of FNR using additional spot-blots. In this experiment, the FNR peptides contained two or three substitutions of consecutive wild-type residues with alanine (**Figure 5.3B**). Only the ^{5}KRIIRRI^{11} region was critical for binding. An orthogonal spot-blot experiment in which glycine residues systematically replace the native sequence of the FNR N-terminal peptide also indicated that the IRRI sequence was a minimal determinant for ClpX N domain binding (**Figure 5.3C**). Finally, we validated solution-based binding of these residues near the N-terminus of FNR to the ClpX N domain using fluorescence anisotropy (**Figure 5.3D**). The binding constant (K_D) of a fluorescently-labeled FNR peptide (residues 1–19) to purified ClpX N domain was ~three-fold weaker than the K_D of the SspB XB peptide, indicating that the ^{5}KRIIRRI^{11} region binds the ClpX N domain less tightly (**Figure 5.2B**).



peptide	K_D (μM)	R^2
FNR (1-19)	89 ± 15	0.97
XB_{SspB}	28 ± 5	0.96

Figure 5.3 FNR contains a ClpX N domain-binding sequence.
(A) Spot-blot of FNR peptide library incubated with ^{35}S-labeled ClpX N domain. The peptide library consisted of 12-residue peptides corresponding to overlapping regions of the FNR protein sequence with an offset of two residues between neighboring peptides. The ClpX N domain-binding sequence near the FNR N-terminus (^{5}KRIIRRI^{11}) is outlined in orange; the C-terminal pore-binding sequence (NVA) is outlined in blue; and binding controls (peptides that bind the ClpX N domain with medium, strong, and weak strength) are outlined in black. **(B)** Substitution analysis of the FNR N-terminal region monitoring ClpX N domain binding by spot-blot. Double and triple alanine substitutions in the first 12 residues of FNR are shown in light blue text. **(C)** Sequential elimination of the first 20 residues of FNR, analyzed by spot-blot. To scan the FNR N-terminal region (res. 1–20), 20-residue peptides containing portions of the ClpX N-domain binding sequence or glycines (indicated in light blue text). **(D)** Fluorescence anisotropy of ClpX N domain titrated against a fixed concentration (1 μM) of dansyl-labeled FNR peptide (res. 1–19) or fluorescein-labeled SspB XB peptide. Values are from individual experiments (n ≥ 2) and fit using a hyperbolic isotherm (see *Methods*). The K_D and R^2 values are reported on the right with (±) the standard error of non-linear least-squares.

The N-domain-binding region in FNR is critical for efficient degradation

To assess the contribution of the FNR residues that bind the ClpX N domain to degradation by ClpXP, we introduced alanine or aspartate mutations into full-length FNR at residues 5-11 (**Figure 5.4A–B**). Point mutations at any position within this region reduced the rate of ClpXP degradation four- to five-fold relative to wild-type *apo*FNR. Triple alanine substitutions at residues 5–7 or 9–11 similarly slowed degradation rates. These results establish that within the context of native *apo*FNR, this N-domain-binding sequence is required for efficient degradation by ClpXP. However, each variant also degraded at ~20-30% of the wild-type rate, supporting earlier work establishing that FNR also contains a second recognition signal located in its C-terminal region (Flynn et al., 2003; Mettert and Kiley, 2005).

The C-terminal NVA-COO^- tripeptide of FNR (residues 248–250) shares sequences similarity with LAA-COO^- of the ssrA degron, which is known to bind in the axial pore of the ClpX ring, and to other C-motif-1 ClpX-recognition signals (Flynn et al., 2003; Fei et al., 2020). ClpXP degraded FNR variants with C-terminal DVA and NVS tripeptides poorly (**Figure 5.4C**), consistent with previous results using an FNR variant with a deletion of the C-terminal VA sequence (Mettert and Kiley, 2005). Hereafter, we will use $Ntag_{FNR}$ to refer to FNR residues 1–15 and $Ctag_{FNR}$ to specify FNR residues 240–250. We also assume that $Ctag_{FNR}$ is engaged by the axial channel of ClpX.

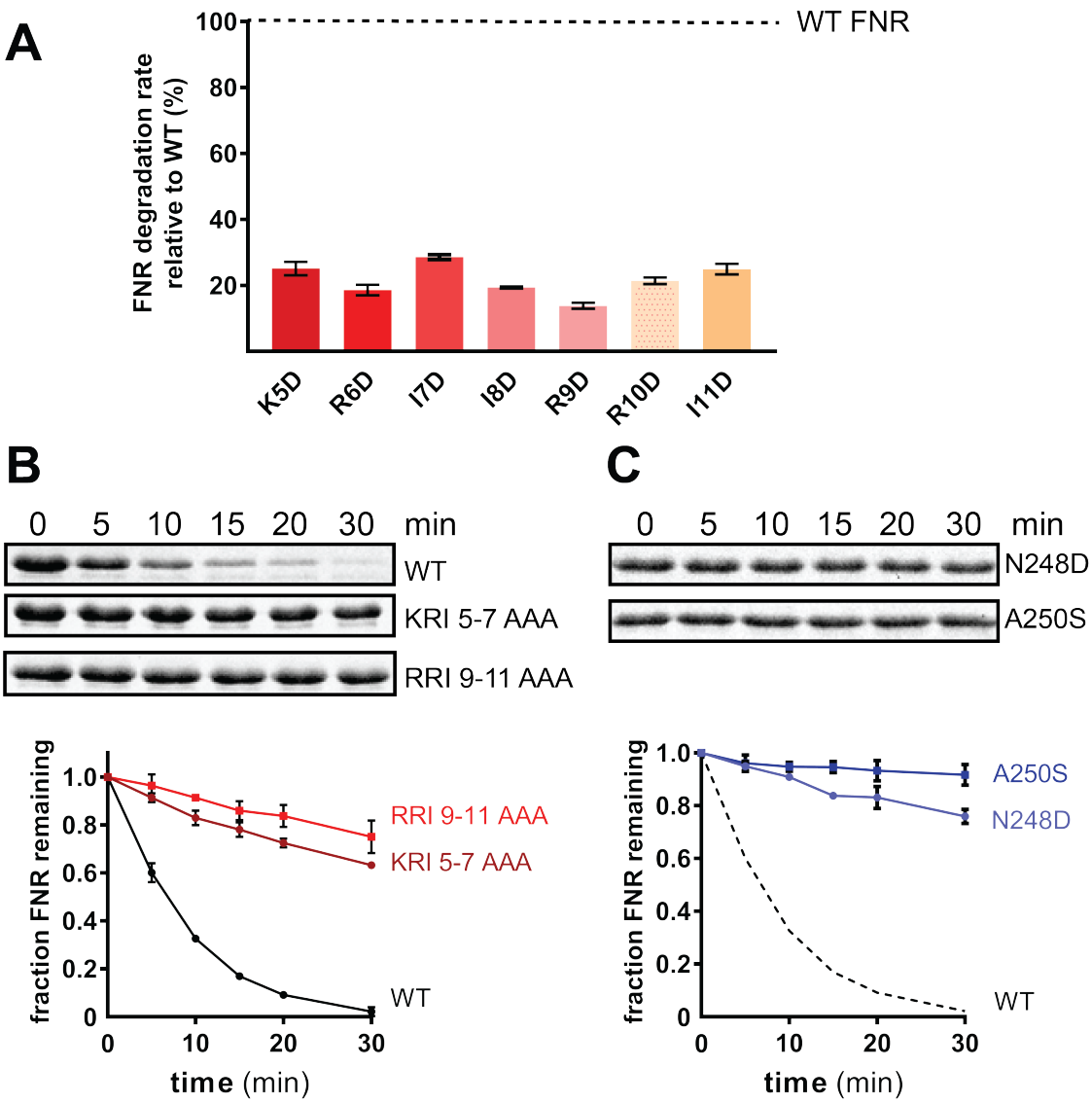


Figure 5.4 FNR residues 5–11 are critical for efficient ClpXP degradation.
(A) Relative degradation rates of FNR protein variants compared to wild-type (WT) *apo*FNR using SDS-PAGE and colloidal Coomassie staining. Bars are mean percentages (calculated from initial degradation rates of each variant divided by the mean degradation of wild-type *apo*FNR) from three independent experiments with error bars representing ± 1 SD. The dashed line represents the degradation rate of wild-type *apo*FNR (1.8 ± 0.01 FNR degraded $ClpX_6^{-1}$ min^{-1}).
(B) Degradation kinetics of FNR triple alanine variants in residues 5–11 or **(C)** DVA (N248D) or NVS (A250S) variants, compared to wild-type *apo*FNR using the assay as described in panel (A). Each point represents the mean fraction FNR remaining (calculated from the density of the FNR band at each time point compared to the beginning of the reaction at zero). Error bars are ± 1 SD from triplicates. The dashed line shows mean values of wild-type *apo*FNR degradation kinetics from the left panel for comparison.
Gel degradation reactions for (A-C) contained *apo*FNR or variants (10 μM), ClpX (0.3 μM hexamer), $ClpP_{14}$ (0.8 μM), 4 mM ATP, and an ATP-regeneration system.

The N- and C-terminal recognition signals work together during FNR degradation

We designed GFP fusion proteins bearing $Ntag_{FNR}$, $Ctag_{FNR}$, or both signals and tested for potential synergy between the tags by performing Michaelis-Menten analysis of ClpXP degradation (**Figure 5.5A; Table 5.1**). GFP itself is not recognized by ClpX (Kim et al., 2000). ClpXP degraded GFP-$Ctag_{FNR}$ with a K_M of ~80 μM, and $Ntag_{FNR}$-GFP interacted with a K_M of ~180 μM. Notably, ClpXP degraded GFP carrying both the $Ntag_{FNR}$ and $Ctag_{FNR}$ signals with a K_M of ~15 μM. We observed similar synergistic degradation enhancement by combining the $Ntag_{FNR}$ and $Ctag_{FNR}$ signals on dihydrofolate reductase (DHFR), another model protein (**Figure 5.5B**). DHFR is frequently used in protein degradation studies because the binding of its ligand methotrexate stabilizes its mechanical stability, which can be used to assay processive degradation by AAA+ proteases (see next section) (Lee et al., 2001; Ainavarapu et al., 2005; Koodithangal et al., 2009; Gur et al., 2012).

Table 5.1 Steady-state kinetic parameters for tagged GFP substrate degradation by ClpXP. K_M and V_{max} values ± errors were obtained by non-linear least squares fitting to the Michaelis-Menten equation from triplicates.

Substrate	K_M (μM)	V_{max} (min^{-1} $ClpX_6^{-1}$)	R^2
$Ntag_{FNR}$-GFP	180 ± 130	2.6 ± 1.4	0.88
GFP-$Ctag_{FNR}$	78 ± 16	2.9 ± 0.4	0.97
$Ntag_{FNR}$-GFP-$Ctag_{FNR}$	15 ± 1.0	1.6 ± 0.1	0.97
XB_{SspB}-GFP-$Ctag_{FNR}$	5.4 ± 0.1	1.2 ± 0.1	0.86

We hypothesized that if $Ntag_{FNR}$ functions as an e-tag by tethering the substrate to the ClpX N domain, then the SspB XB peptide (LRVVK, residues 160–164 from *E. coli* SspB) should be able to substitute for $Ntag_{FNR}$ to enhance degradation of GFP-$Ctag_{FNR}$ (**Figure 5.5C**). K_M for ClpXP degradation of XB_{SspB}-GFP-$Ctag_{FNR}$ was ~5 μM or ~three-fold tighter than the K_M for $Ntag_{FNR}$-GFP-$Ctag_{FNR}$ (**Table 5.1**). This tighter K_M further supports an e-tag tethering mechanism, as the N

domain of ClpX binds the XB_{SspB} peptide (K_D ~20-30 μM; Wah et al., 2003) more tightly than a $Ntag_{FNR}$ peptide (K_D ~89 μM). We conclude that the N- and C-terminal tags of FNR are modular degradation signals and play distinct but markedly synergistic roles in targeting substrates for ClpXP proteolysis.

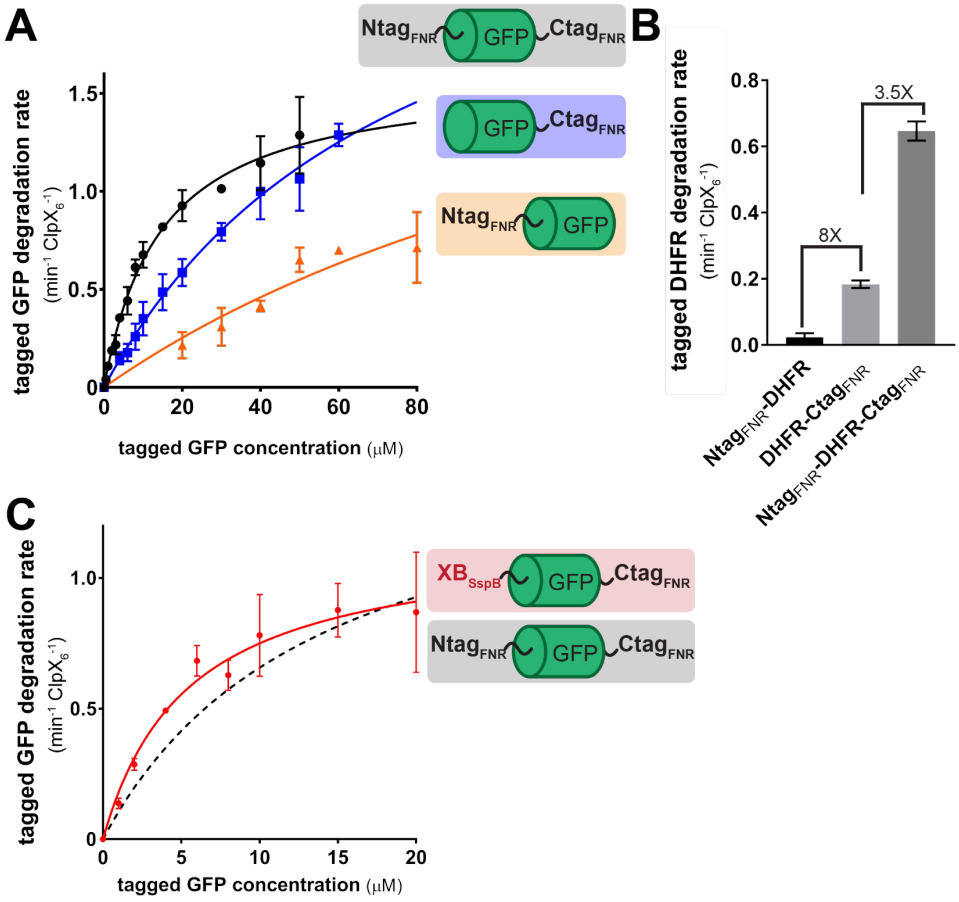


Figure 5.5 FNR N- and C-tags can be transferred to other proteins and have distinct functions. (A) Michaelis-Menten analysis of GFP-fusion protein degradation. Reactions contained ClpX (0.3 μM hexamer), $ClpP_{14}$ (0.8 μM), 5 mM ATP, and an ATP-regeneration system. Degradation was monitored by loss of fluorescence at 485 nm over time. Points are mean degradation rates (n=3); error bars ± 1 SD. **(B)** DHFR-fusion protein degradation rates in the presence of ClpX (0.3 μM hexamer), $ClpP_{14}$ (0.8 μM), DHFR-fusion proteins (5 μM), 5 mM ATP, and an ATP-regeneration system. Degradation was monitored by SDS-PAGE and colloidal Coomassie staining. Values correspond to mean degradation rates (n=3); error bars ± 1 SD. **(C)** Michaelis-Menten analysis of XB_{SspB}-GFP-$Ctag_{FNR}$ degradation. Conditions as in panel (A). The Michaelis-Menten fit for $Ntag_{FNR}$-GFP-$Ctag_{FNR}$ degradation from panel (A) is shown by the black dashed line for comparison. Degradation was monitored by loss of fluorescence at 485 nm over time. Points are mean degradation rates (n=3); error bars ± 1 SD.

FNR is degraded from the C-terminus

To determine whether ClpXP initiates degradation of FNR from its N- or C- terminus, we designed chimeric proteins containing full-length FNR (FNR^{FL}) and an $Ntag_{FNR}$ or $Ctag_{FNR}$ signal fused to DHFR. ClpXP only degrades methotrexate-bound DHFR slowly, and truncated degradation products are typically produced when ClpXP reaches the ligand-stabilized DHFR domain, because the partially degraded product is released with no degradation tag and is resistant to further degradation (Lee et al., 2001; Kenniston et al., 2005; Koodithangal et al., 2009). We generated $Ntag_{FNR}$-DHFR-FNR^{FL} to detect degradation initiated at the FNR C-terminus (**Figure 5.S1**) and FNR^{FL}-DHFR-$Ctag_{FNR}$ to assay for degradation from the FNR N-terminus (**Figure 5.S2**). In the absence of methotrexate, ClpXP degraded the FNR^{FL} and DHFR domains in both chimeric substrates, indicating that the $Ntag_{FNR}$ and $Ctag_{FNR}$ signals provided intrinsically in FNR^{FL} or appended to DHFR are functional (**Figure 5.S1–3**). In the presence of methotrexate, we observed a truncated protein at the approximate position expected of the DHFR domain accumulating during degradation of the $Ntag_{FNR}$-DHFR-FNR^{FL} chimera (**Figure 5.S1A**). We confirmed that the truncated protein fragment contained the DHFR domain by Western blot using an anti-DHFR antibody (**Figure 5.S1B**), indicating that ClpXP initiates processive unfolding and degradation of the FNR^{FL} portion from the C-terminus. Formation of the DHFR fragment appeared sub-stoichiometric compared to the amount of substrate degraded (**Figure 5.S3**), which is likely explained by slow ClpXP degradation methotrexate-bound mouse DHFR. By contrast, no truncation product was observed during ClpXP degradation of FNR^{FL}-DHFR-$Ctag_{FNR}$ (**Figure 5.S2**), indicating that the FNR portion of this chimera is not degraded from the N-terminal end. Therefore, these results reveal that ClpXP initiates processive unfolding and degradation from the C-terminal end of *apo*FNR, consistent with the ssrA tag-like function of the $Ctag_{FNR}$ signal.

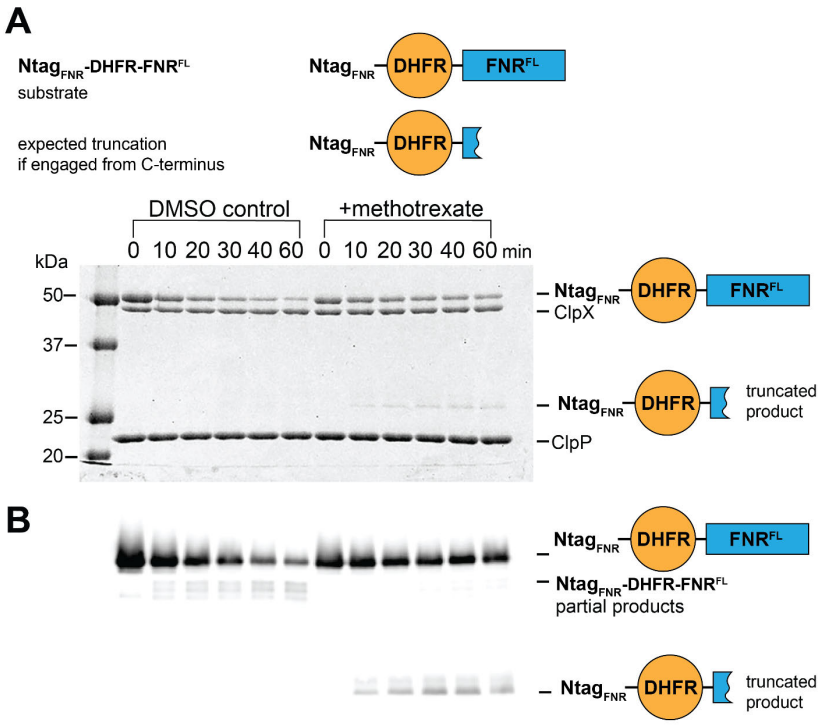


Figure 5.S1 Processive degradation by ClpXP proceeds from the C-terminus of FNR.
(A) Gel degradation of Ntag$_{FNR}$-DHFR-FNR^{FL} in the presence and absence of methotrexate (10 μM in 0.2% DMSO) detected by colloidal Coomassie staining. The control reactions contained 0.2% DMSO. Degradation reactions contained 5 μM Ntag$_{FNR}$-DHFR-FNR^{FL}, ClpX$_6$ (0.3 μM), ClpP$_{14}$ (0.8 μM), and 5 mM ATP. The molecular weight of FNR is ~28 kDa, and the molecular weight of DHFR is ~22 kDa. **(B)** Western blot of degradation reactions in (A) probed with mouse anti-DHFR antibody. The migration of the truncated DHFR-containing product is indicated on the right.

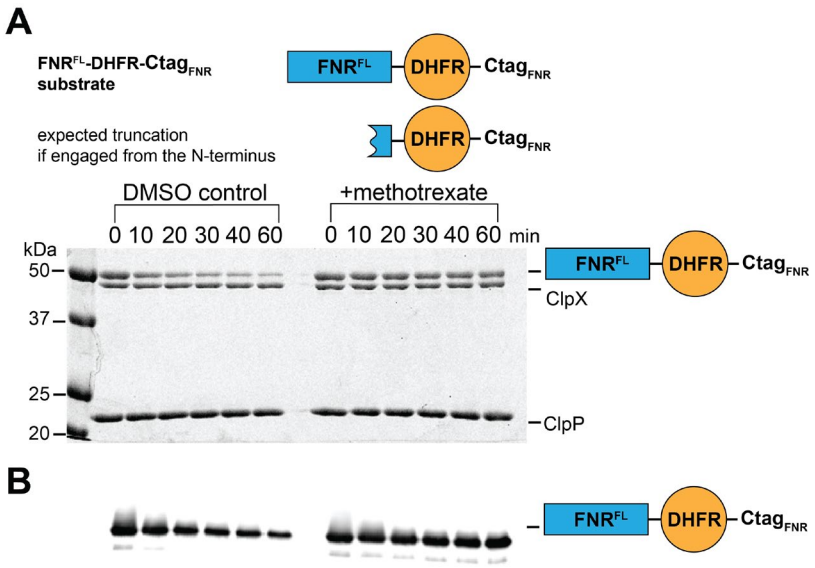


Figure 5.S2 The N-terminus of FNR is not the principal ClpXP initiation site.
(A) Gel degradation of FNR^{FL}-DHFR-Ctag$_{FNR}$ in the presence and absence of methotrexate (10 μM in 0.2% DMSO) detected by colloidal Coomassie staining. Degradation reactions contained 5 μM FNR^{FL}-DHFR-Ctag$_{FNR}$, ClpX$_6$ (0.3 μM), ClpP$_{14}$ (0.8 μM), and 5 mM ATP. The control reactions contained 0.2% DMSO. **(B)** Western blot of degradation reactions in (A) probed with mouse anti-DHFR antibody. No truncation products containing DHFR were detected.

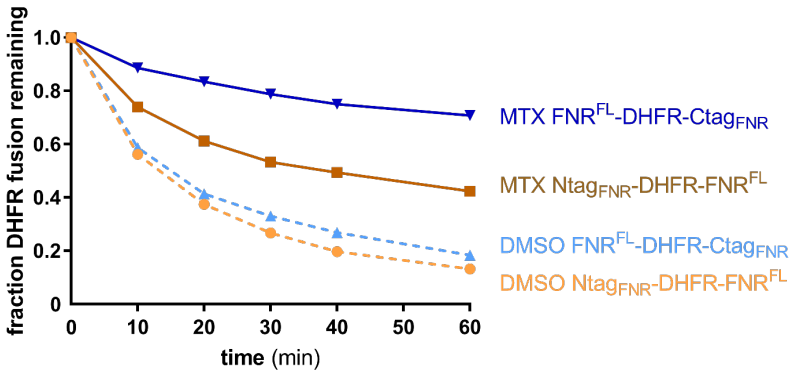


Figure 5.S3 Kinetics of degradation of $Ntag_{FNR}$-DHFR-FNR^{FL} and FNR^{FL}-DHFR-$Ctag_{FNR}$ fusion proteins in the presence (solid lines) and absence (dashed lines) of methotrexate (MTX; 10 μM, 0.2% DMSO) monitored by SDS-PAGE and colloidal Coomassie staining, quantified from Figure 5.S1–2.

Dimerization can protect *apo*FNR from ClpXP proteolysis by masking recognition signals

Previous work established that dimerization protects FNR from proteolysis (Mettert and Kiley 2005). To gain structural insights into the conformational changes associated with dimerization underlying differences in ClpXP degradation, we crystallized *apo*FNR and *holo*FNR variants, both bearing a deletion of the N-terminal 14 residues (Δ14) and a L28H substitution, and solved structures to ~2 Å resolution (**Table 5.2**). Residues 17–35 in the N-terminal region of the *holo*FNR variant form a loop that wraps around the [4Fe-4S] cluster, generating a compact folded motif (**Figure 5.6A**). In the crystal structure of the *apo*FNR variant, however, the corresponding residues are flexibly extended (**Figure 5.6B**). This conformational change is likely to expose residues near the N-terminus of FNR, making the e-tag accessible to ClpXP, although residues 1–14 were not present in the electron-density map for either *apo* or *holo*FNR. Attempts to crystallize full-length FNR were unsuccessful, suggesting that residues 1–14 may be flexible in both states, whereas residues 17–35 were accessible only in the *apo* structure. By contrast, the C-terminal segment of FNR was visible in both the *apo* and *holo* variant structures. $Ctag_{FNR}$ in *holo*FNR packed against the DNA-binding domain near the dimerization helix in *holo*FNR but was rotated by ~90° and was more accessible in *apo*FNR (**Figure 5.6C**). Because dimerization normally occurs only in the *holo* state, these structures suggest dimer formation sequesters both

FNR recognition signals, preventing binding to ClpXP via distinct mechanisms. Specifically, coordination to the [Fe-S] cluster in residues 17–35 likely hinders access to $Ntag_{FNR}$, while ligand-induced conformational changes stabilize packing of $Ctag_{FNR}$ near the dimerization helix.

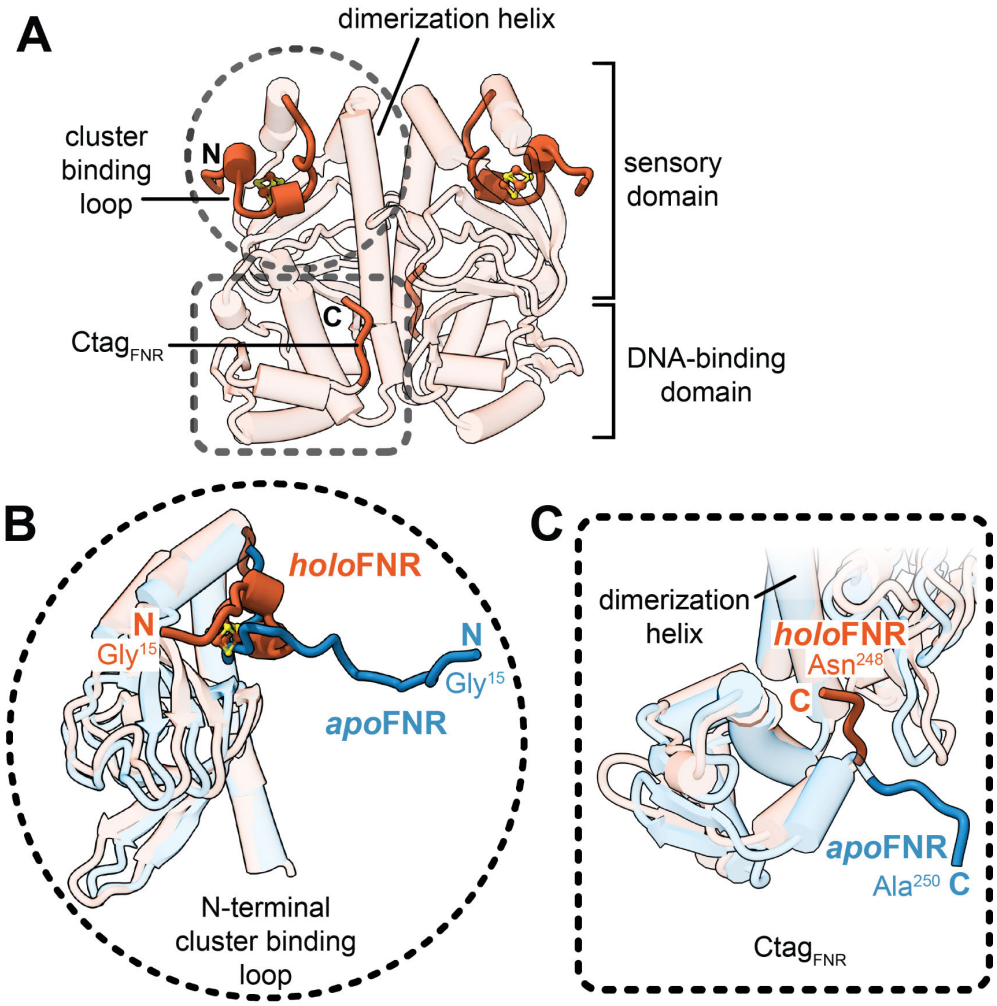


Figure 5.6 N- and C-terminal regions of *apo*FNR are more exposed than in *holo*FNR structures. (A) *holo*FNR dimer structure (colored dark orange) showing the dimerization helix, N-terminal cluster-binding loop (dotted circle; residues 17–35), and the C-terminal $Ctag_{FNR}$ signal (dotted rectangle). The overall domain organization of each FNR monomer, which consists of the N-terminal sensory domain and the C-terminal DNA binding domain, is indicated on the right. **(B)** Conformational differences in the N-terminal region of *holo*FNR (colored dark orange) and *apo*FNR (colored light blue) (residues 15–35) from the region indicated in the dotted circle in panel (A). The first residue (Gly^{15}) resolved in the *apo* crystal structure is labeled. **(C)** Comparison of the orientation of the C-terminal region of *holo* and *apo*FNR (residues 245-250), from the region indicated by the dotted rectangle in panel (A). In *holo*FNR, the last modeled residue in the structure is Asn^{248}, whereas in *apo*FNR, it is Ala^{250}.

Table 5.2. Data collection and refinement statistics for FNR structures.

	***Apo*FNRΔ14/L28H**	***Holo*FNRΔ14/L28H**
Wavelength (Å)	0.979145	0.979145
Resolution range	45.81 – 1.43 (1.47 – 1.43)	50.00 – 1.64 (1.68 – 1.64)
Space group	P 2_1 2_1 2_1	P 3_2 2 1
Unit cell dimensions (Å) α, β, γ (°)	51.27 81.01 101.80	112.62 112.62 124.26 γ = 120
Total reflections	698651 (99045)	1121875 (181205)
Unique reflections	78947 (12563)	111317 (17795)
Multiplicity	8.85 (8.88)	10.08 (10.18)
Completeness (%)	99.9 (99.6)	99.9 (99.7)
Mean I/σ(I)	22.35 (1.99)	22.83 (1.95)
R_{sym} (%)	4.2 (94.5)	5.2 (99.8)
CC-1/2 (%)	99.9 (84.2)	100.0 (74.2)
Reflections used in refinement	78913 (7736)	111309 (10993)
Reflections used for R_{free}	3946 (387)	5565 (550)
R_{work}	17.20 (28.24)	19.17 (34.85)
R_{free}	20.59 (33.31)	20.74 (36.35)
Number chains per ASU protein	2	2
Number of non-hydrogen atoms	4069	4133
protein	3694	3705
metal-cluster	–	16
solvent	375	412
RMS (bonds)	0.008	0.006
RMS (angles)	1.00	0.88
Ramachandran favored (%)	98.28	96.63
Ramachandran allowed (%)	1.29	3.16
Ramachandran outliers (%)	0.43	0.22
Rotamer outliers (%)	1.26	3.75
Clashscore	2.84	3.64
Average B-factor (Å^2)	30.12	36.16
protein	29.08	35.36
metal-cluster (average)	-	29.12
solvent	40.34	43.64

Statistics for the highest-resolution shells are shown in parentheses.

$$R_{\text{sym}} = \frac{\sum_{hkl} \sum_j \left| I_{hkl,j} - \langle I_{hkl} \rangle \right|}{\sum_{hkl} \sum_j I_{hkl,j}}$$

We tested if dimerization could protect FNR from ClpXP degradation by masking FNR's recognition signals, as suggested by the relative exposure of the $Ntag_{FNR}$ and $Ctag_{FNR}$ sequences deduced from the *apo* and *holo* structures. To separate the possible effects of the presence of the [4Fe-4S] cluster in *holo*FNR from those of dimerization, we took advantage of a previously characterized FNR variant bearing the D154K mutation which stabilizes the FNR dimer interface in the absence of the cluster (Moore et al., 2006) (**Figure 5.7A**). If dimerization-mediated burial of the recognition signals is sufficient for resistance to proteolysis, then *apo*FNR^{D154K}-dimers should be resistant to ClpXP degradation. Consistent with this hypothesis and previous studies using the D154A mutant (Mettert and Kiley, 2005), we observed a reduction in the degradation rate of the *apo*FNR^{D154K}-dimer compared to *apo*FNR-monomer under aerobic conditions where neither protein should contain an intact cluster. Thus, at the very least, FNR dimerization slows ClpXP degradation rates.

Next, we used *apo*FNR^{D154K} to test whether occlusion of one or both of the ClpX recognition signals upon dimerization is sufficient to explain the resistance of the dimer to ClpXP degradation. Inspection of the *holo*FNR structure suggests that the $Ctag_{FNR}$ signal is less accessible in an FNR dimer, whereas the region containing $Ntag_{FNR}$ would likely only be protected via coordination to the [4Fe-4S] cluster. Therefore, we predicted that appending a longer $Ntag_{FNR}$ sequence should not greatly affect degradation of *apo*FNR^{D154K} dimers, which already have an exposed $Ntag_{FNR}$. By contrast, appending a longer $Ctag_{FNR}$ sequence would expose this signal and allow ClpXP to access these sites that would otherwise be hidden within an FNR dimer. Indeed, when the C-terminus of the *apo*FNR^{D154K}-dimer was extended by the addition of $Ctag_{FNR}$, the dimeric protein was rapidly degraded by ClpXP. Thus, restoring exposure of the ClpX pore-binding tag was sufficient to confer rapid degradation to a dimeric variant of FNR (**Figure 5.7B**). Duplication of the $Ntag_{FNR}$ sequence enhanced degradation of the *apo*FNR^{D154K}-dimer, by ~3-fold, compared to the 14-fold enhancement of the degradation rate after $Ctag_{FNR}$ was appended (both rates are

compared to the 'untagged' *apo*FNR^{D154K}-dimer). We attribute the increased degradation rate of the Ntag$_{FNR}$-*apo*FNR^{D154K}-dimer to the ability of Ntag$_{FNR}$ to serve as a very poor degron (see **Figure 5.5A**), which may be improved when fused to the existing Ntag$_{FNR}$ sequence (generating two tandem Ntag$_{FNR}$ sequences) in the *apo*FNR^{D154K} protein. These results, in conjunction with the structures of *apo* and *holo*FNR, establish that oxygen-regulated dimerization controls exposure of two synergistically acting ClpX recognition signals in FNR and therefore provides a means to couple degradation to oxygen levels, ensuring robust, environmentally sensitive regulation of FNR-controlled gene expression *in vivo*.

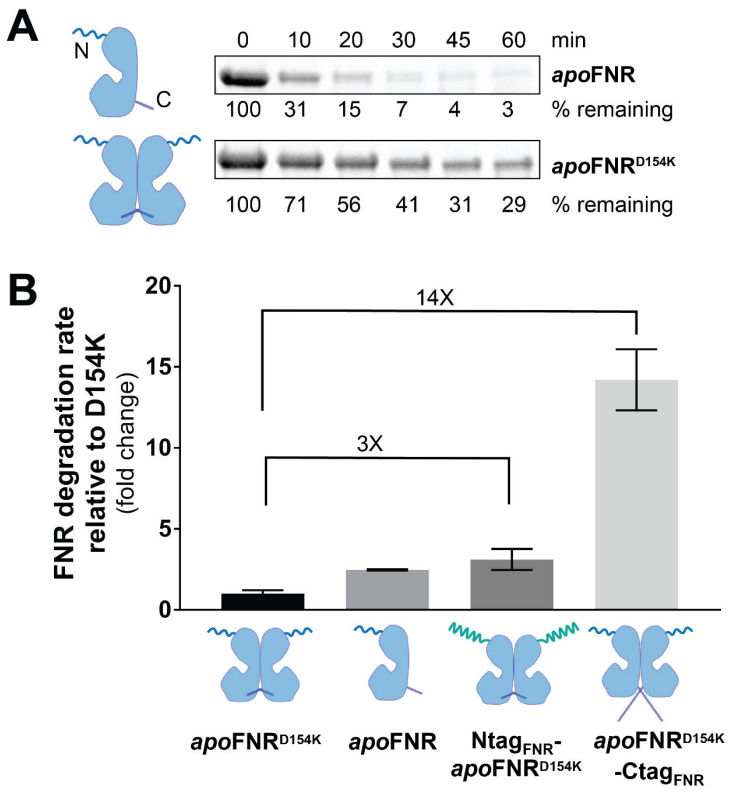


Figure 5.7 Dimerization masks FNR recognition signals.
(A) Kinetics of *apo*FNR and *apo*FNR^{D154K} degradation. Gel degradation reactions contained 20 μM *apo*FNR or *apo*FNR^{D154K}, $ClpX_6$ (1 μM), $ClpP_{14}$ (2 μM), 5 mM ATP, and an ATP-regeneration system. The degradation rate of *apo*FNR is 1.4 ± 0.1 monomers degraded min^{-1} $ClpX_6^{-1}$ and that of *apo*FNR^{D154K} is 0.4 ± 0.1 monomers degraded min^{-1} $ClpX_6^{-1}$; error reported is the SD with n=4.
(B) Kinetics of *apo*FNR, *apo*FNR^{D154K}, Ntag$_{FNR}$-*apo*FNR^{D154K}, and *apo*FNR^{D154K}-Ctag$_{FNR}$ degradation. Degradation reactions contained 20 μM *apo*FNR or variants, $ClpX_6$ (0.3 μM), $ClpP_{14}$ (0.8 μM), 5 mM ATP, and an ATP-regeneration system. Bars are mean degradation rate ratios (n=3) and error bars represent ± 1 SD. The Ntag$_{FNR}$ recognition signal is depicted as a wavy line, and the Ctag$_{FNR}$ signal is depicted as a straight line, which is usually inaccessible in the D154K dimeric protein. An additional 80 mM KCl was used to help stabilize the FNR dimer in solution for (A) and (B).

DISCUSSION

FNR's two recognition signals have distinct functions in ClpX proteolysis

Our results support a model in which ClpXP degradation of *apo*FNR involves an N-terminal e-tag (^{5}KRIIRRI^{11}) that binds to the ClpX N domain and a C-terminal pore tag (^{248}NVA^{250}) that binds in the axial pore of the ClpX ring (**Figure 5.8**). Similar to the well-studied ssrA degron, which also binds to the pore of ClpX (Martin et al., 2008; Fei et al., 2020), the FNR pore tag can function as an autonomous degron (**Figure 5.5A–B**) and appears to be the site at which ClpXP initiates degradation (**Figure 5.S1–2**). Enhancement tags or e-tags strengthen the interaction between the AAA+ proteolytic enzyme and its substrate to improve recognition of a principal recognition signal or degron. E-tag sequences can be located in any region within the substrate polypeptide and have been identified near the termini or in the interior of various substrates (Abdelhakim et al., 2008; Camberg et al., 2014; Ling et al., 2015). For *apo*FNR, the position of the e-tag near the N-terminus would enable the C-terminal tag to bind the ClpX pore and to tether the N-terminal e-tag to one of ClpX's N domains at the same time, as the N- and C-termini of FNR are ~ 38 Å apart from each other (Gly^{15} to Arg^{247} of chain A in the structure). Conversely, if the e-tag were located too close to the pore tag, then only one signal could bind ClpX at a time.

In our model, binding of the e-tag to the ClpX N domain or the pore tag to the ClpX channel promotes binding of the other tag. For bimolecular binding, effective concentration or C_{eff} can be defined as: $C_{eff} = K_1/K_4 = K_3/K_2 = K_1 \bullet K_3/K_{link}$. Using the K_M for GFP-fusion-protein degradation as a proxy for K_D (Table 5.1), the bivalent e-tag/pore tag fusion protein has an estimated C_{eff} value of ~0.9 mM, calculated as $K_1 \bullet K_3/K_{link}$. We conclude, therefore, that the FNR e-tag functions as a tethering motif to increase the effective concentration of the FNR N-terminal tag with respect to the ClpX pore (**Figure 5.5C**).

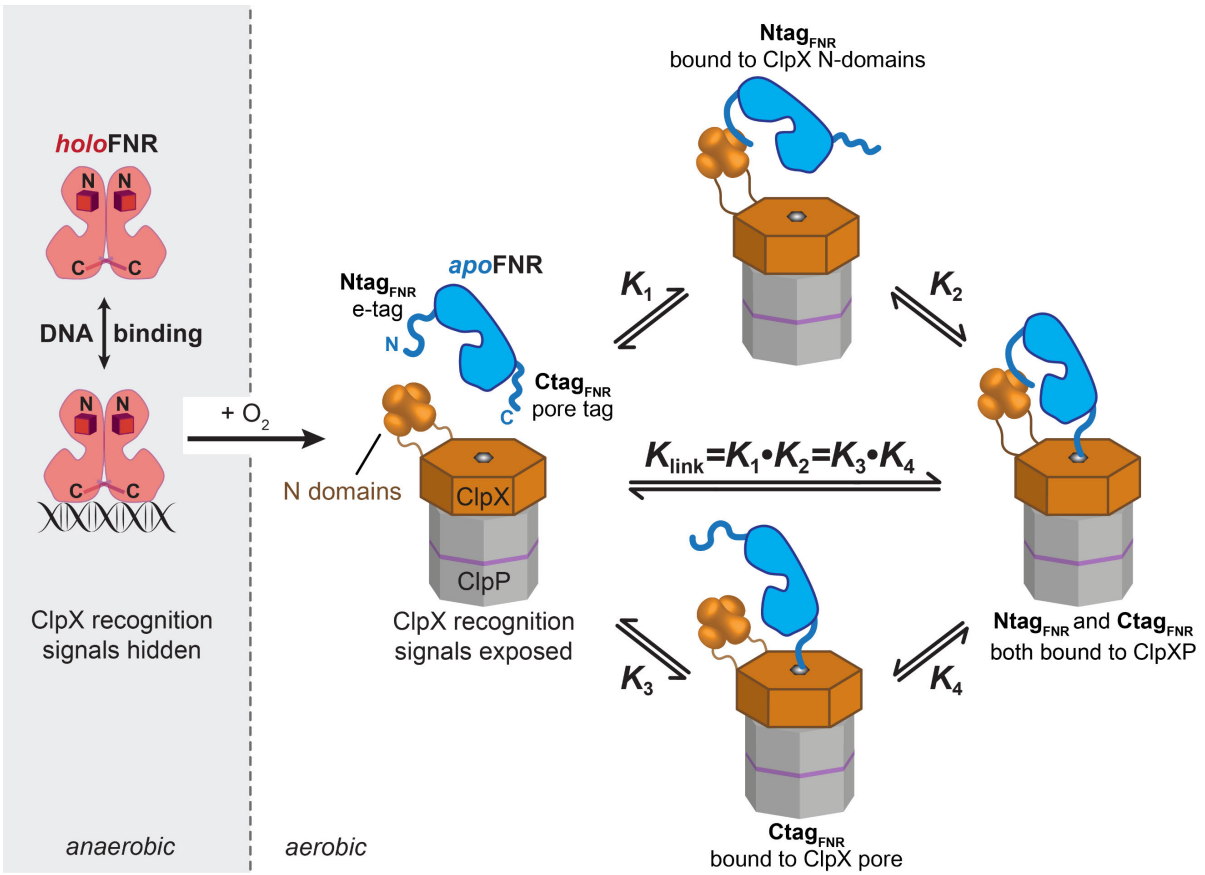


Figure 5.8 Proposed model of *apo*FNR recognition by ClpXP. Oxygen exposure regulates accessibility of $Ntag_{FNR}$ e-tag (residues 5-11) and $Ctag_{FNR}$ pore-binding tag (residues 248-250). Under anaerobic conditions, [4Fe-4S]-cluster binding stabilizes *holo*FNR dimers that are active in DNA binding, and both recognition signals are protected from ClpXP. The N-terminal region of FNR coordinates the [4Fe-4S] cluster, whereas access to the C-terminal region of FNR is hindered by dimerization. Only one set of N-domain dimers is shown for clarity. Recognition of *apo*FNR occurs in a bimolecular binding reaction with the e-tag and pore tag, depicted by the equilibrium arrows.

Both signals are important for efficient degradation by ClpXP, as mutation of either the e-tag or pore-binding tag impairs FNR degradation (**Figure 5.4**). A previous study reported that deletion of the e-tag or alanine substitution of residues 5–7 in the e-tag did not increase steady-state FNR protein levels as quantified by Western blot (Pan et al., 2012), but our direct protein degradation assays *in vitro* and those from prior degradation analyses *in vivo* and *in vitro* (Mettert and Kiley, 2005) demonstrate the importance of the e-tag region. We also find that the e-tag and pore tag signals are sufficient to specify ClpXP proteolysis by transplanting them into GFP or DHFR.

Combining the $Ntag_{FNR}$ signal with the $Ctag_{FNR}$ signal enhanced K_M for degradation of the GFP substrate ~five-fold and degradation rate of the DHFR substrate ~four-fold. Together, these data demonstrate how two distinct peptide signals cooperate in the same polypeptide to enhance regulated proteolysis.

apoFNR recognition by ClpXP requires the ClpX N domain but not an adaptor

We found that *apo*FNR degradation by ClpXP requires the ClpX N domain (**Figure 5.2A**) and that a peptide containing the e-tag binds directly to the ClpX N domain. Furthermore, our experiments replacing the ClpX N domain binding sequence in the e-tag signal with another ClpX N domain tethering motif (XB_{SspB}) and the inhibition of *apo*FNR degradation by a competing XB peptide suggest that FNR and SspB may bind to a common site on the ClpX N domain (**Figure 5.5C and 5.2B**). Similar to FNR, FtsZ is another ClpXP substrate that features N-domain-dependent recognition of two distinct recognition signals (Camberg et al., 2009, 2014). However, unlike FNR which has only one e-tag tethering motif, both FtsZ signals may bind the N domain (Camberg et al., 2014). At least one of these sequences binds the ClpX pore to initiate substrate degradation, which suggests that the ClpX N domain or the pore may compete for binding. Structurally elucidating how ClpXP interacts with *apo*FNR will confirm whether the ClpX N domain binds the e-tag and the axial channel engages the pore tag as we predict (**Figure 5.8**). These results also support the possibility that FNR may compete *in vivo* with other adaptor/delivery factor-mediated substrates that also bind the ClpX N domain, as proposed previously for UmuD' and CtrA (Neher et al., 2003; Chien et al., 2007).

The successful modular replacement of the FNR e-tag with the XB_{SspB} motif suggests that any N-terminal tethering motif will be sufficient to enhance recognition of the FNR pore tag, as long as the distance between both signals is sufficiently long and flexible. Importantly, ClpXP degradation of *apo*FNR does not require an adaptor protein. By tethering itself to ClpX through its e-tag

directly, FNR circumvents the need for a *trans*-targeting delivery factor, which is required for UmuD' and σ^S (Zhou et al., 2001; Neher et al., 2003; Studemann et al., 2003). Additionally, because these signals function within the same polypeptide, the placement of the e-tag and pore tags enables degradation of the *apo*FNR monomer. The e-tag/pore-tag *apo*FNR recognition mode provides a minimal model of targeted protein degradation using multiple recognition signals. Proteomic profiling of *E. coli* ClpXP substrates revealed that ~25% of all identified substrates had recognition signals located near their N- or C- termini. Whether or not these and other AAA+ protease substrates carrying recognition signals at both termini also follow the example of FNR or require delivery factors and/or adaptor proteins will help determine whether the dual-recognition strategy presented here is more universal or limited in scope.

Conformational changes modulating the FNR oligomeric state drive *apoFNR* degradation

The autodelivery system used by *apo*FNR is unusual as efficient recognition by ClpXP occurs when FNR is monomeric (but not in the *holo* dimer). Other ClpXP substrates that require a similar e-tag/pore-tag strategy for efficient degradation, such as UmuD/UmuD' and the MuA tetramer, are recognized as oligomers but not monomers (Neher et al., 2003; Ling et al., 2015). By contrast, *holo*FNR dimers are not recognized by ClpXP despite containing the necessary recognition signals. Hence, a mechanism must exist that protects active *holo*FNR dimers from ClpXP degradation, while allowing proteolysis of the unneeded and potentially deleterious oxygen-damaged *apo*FNR monomers. Our structures suggest that this mechanism involves masking recognition signals in the dimer, with the e-tag being at least partially hidden by coordinating the iron-sulfur cluster and access to the pore-binding tag being obstructed by the dimer interface (**Figure 5.6**). These terminal regions are exposed as flexible peptides in the monomer, which lacks the [4Fe-4S] cluster. The linked equilibria driven by [4Fe-4S] binding, FNR dimerization, and DNA binding serve both to limit exposure of the ClpXP recognition signals under anaerobic conditions and to ensure their accessibility in the presence of oxygen.

Effect of ClpXP-mediated degradation of FNR on O_2 sensing

Turnover of FNR by ClpXP-dependent degradation is proposed to allow dynamic transitions from aerobic-to-anaerobic conditions (Partridge et al., 2007; Tolla and Savageau, 2010; Tolla et al., 2015). Computational modeling predicts that FNR is a fast-reacting O_2 sensor, with inactive *apo*FNR being quickly depleted by ClpXP degradation under aerobic conditions, but steady-state levels of [4Fe-4S]-containing *holo*FNR accumulating from ~0.15 μM to 5 μM within four minutes of the transition to anaerobic growth (Tolla and Savageau, 2010; Tolla et al., 2015). This rapid increase in active FNR helps *E. coli* quickly adapt to low oxygen. Furthermore, without the targeted proteolysis of *apo*FNR, growth-dependent dilution would be insufficient to reduce FNR levels and thus, to tightly control induction of the *fnr* regulon under aerobic growth (Jervis et al., 2009).

The switch-like post-translational regulation of FNR levels is similar to hypoxia-inducible factor 1-alpha (HIF-1α), an important regulator of oxygen homeostasis in mammals. Analogous to FNR, cells express HIF-1α continuously but maintain low protein levels via constitutive degradation by the 26S proteasome, the major eukaryotic AAA+ protease (Salceda and Caro, 1997; Semenza, 2004). Under oxygen-replete conditions, prolines in HIF-1α are hydroxylated, allowing specific recognition by an E3 ubiquitin ligase (Ohh et al., 2000; Semenza, 2004; Marxsen et al., 2004). Under hypoxia, HIF-1α is not targeted for degradation, as hydroxylation requires oxygen, resulting in rapid accumulation of HIF-1α under low-oxygen levels (Salceda and Caro, 1997; Kohn et al., 2004; Qutub and Popel, 2006; Bagnall et al., 2014). Similarly, the oxidized form of the hydrogen peroxide sensor PerR is selectively degraded by the Lon AAA+ protease (Ahn and Baker, 2016). Like the loss of the [4Fe-4S] cluster in the conversion from *holo* to *apo*FNR, peroxide-catalyzed oxidation results in loss of PerR's metal ligand, inducing an 'open' conformation that is more accessible for protease recognition (Lee and Helmann, 2006; Traoré et al., 2009; Jacquamet et al., 2009). Thus, in FNR, HIF-1α, and PerR, selective protease recognition of the oxidized state

regulates levels of important transcriptional regulators involved in redox-sensing and oxygen homeostasis.

In addition to FNR, five other redox-sensitive proteins – IscR (iron-sulfur cluster regulator), IscU (iron-sulfur cluster assembly enzyme U), AcnB (aconitate hydratase B), MoaA (molybdenum cofactor biosynthesis protein), and LipA (lipoyl synthase) – are substrates of ClpXP (Flynn et al., 2003). If these substrates are also selectively degraded under aerobic conditions, ClpXP may generally recognize oxygen-sensitive proteins in a specific redox state. For example, Fe-S cluster biogenesis proteins, such as IscU, may interact with FNR (near the N-terminal cluster-binding region) to occlude binding of the e-tag to ClpX. During the transition to aerobic growth, Fe-S assembly factors could promote regeneration of functional *holo*FNR both by catalyzing cluster assembly in *apo*FNR and by inhibiting ClpXP degradation as proposed (Mettert and Kiley 2005).

Regulation of the global transcription factor FNR encompasses both transcription and post-translational control (Spiro and Guest, 1987; Guest et al., 1996; Flynn et al., 2003; Mettert et al., 2008; Green et al., 2009). The oxygen-sensitive [4Fe-4S] cluster ligand controls both DNA-binding activity of the *holo*FNR dimer and accessibility of two AAA+ protease recognition signals in *apo*FNR monomers, preventing constitutive degradation of active *holo* dimers. The *apo*FNR recognition system consists of an e-tag that tethers the substrate to the ClpX N domain and a pore tag that is engaged by the ClpX axial channel. Despite their different functions, these signals cooperate to support efficient degradation within the same polypeptide and do not require the assistance of another protein in the form of a delivery factor or adaptor protein. This dual-tag adaptor-independent recognition mechanism of *apo*FNR presents a new example of combinatorial control of proteolysis targeting the specific redox and assembly state of a substrate protein.

MATERIALS AND METHODS

Proteins and peptides

*Apo*FNR and variants were purified as described (Lazazzera et al., 1993; Yan and Kiley, 2009). *E. coli* PK22 cells, a gift from P. Kiley (Lazazzera et al., 1993), used for FNR overexpression were grown in M9 minimal medium supplemented with 0.2% (w/v) glucose, 0.2% (w/v) casamino acids, 1 mM $MgCl_2$, 100 µM $CaCl_2$, 2 µg/mL thiamine hydrochloride, 20 µM ferric ammonium citrate, and 50 µg/mL ampicillin (pET11a derivatives) or kanamycin (pET28b derivatives) at 37 °C with shaking at 200 rpm. After induction with 0.4 mM IPTG at an OD_{600} of 0.4-0.6, cells were harvested after growth for 1 h. Cells were resuspended in lysis buffer (50 mM potassium phosphate (pH 6.8), 100 mM KCl, 10% (w/v) glycerol, and 1 mM DTT) and lysed by french press at 28 kpsi, followed by treatment with protease inhibitor cocktail set III, EDTA-free (Calbiochem) and benzonase. After centrifugation at 30,000 x *g* for 30 min at 4 °C, the supernatant was applied to a 5 mL HiTrap Heparin HP column (GE Healthcare). Following washing with lysis buffer, an eight column-volume (CV) gradient of 0-100% FB (lysis buffer plus 0.9 M KCl) was applied at a 2 mL/min flow rate. Fractions containing *apo*FNR were identified by SDS-PAGE, pooled, and loaded on a HiLoad 16/600 Superdex-75 pg column at 0.8 mL/min in lysis buffer. Fractions containing *apo*FNR following size-exclusion chromatography (SEC) were pooled and run on a 1 mL HiTrap Heparin HP column (GE Healthcare) with a 0-80% FB gradient over 5 CV at 1 mL/min. Fractions containing purified *apo*FNR were identified by SDS-PAGE and pooled for storage. Plasmids expressing FNR variants (K5D, R6D, I7D, I8D, R9D, R10D, I11D, N248D, A250S) were generated by site-directed mutagenesis using pPK823 (Lazazzera et al., 1993), which carries the coding sequence of *fnr* cloned into pET11a (Novagen). Plasmids for FNR proteins (*apo*FNR^{D154K}, $Ntag_{FNR}$-*apo*FNR^{D154K}, *apo*FNR^{D154K}-$Ctag_{FNR}$) were generated by site-directed mutagenesis using pET28b_FNR, which carries *fnr* cloned into pET28b (Novagen).

For FNR used in crystallography, the plasmid expressing FNRΔ1–14 (pCNFNRΔ14) bearing the L28H mutation was generated by site-directed mutagenesis and Gibson Assembly (NEB) of pPK823 into pET28b (Novagen). To express *holo*FNRΔ14/L28H, pCNFNRΔ14 was co-transformed with pDB1282, a gift from D. Dean (Zheng et al., 1998) that encodes the Isc operon, into *E. coli* PK22 cells using selection with 50 µg/mL kanamycin and 100 µg/mL ampicillin. *E. coli* PK22 cells expressing both plasmids were grown in M9 minimal media with supplemented as described above, and first induced at OD_{600} 0.4–0.6 with 0.5 g/L arabinose for 1–2 h, followed by addition of 0.4 mM IPTG at OD_{600} 0.8 and 30 µM ferric ammonium citrate for 30 min. Finally, 50 mg L-cysteine was added, and cells were grown for another 1.5–2 h, then transferred to Pyrex bottles for argon gas exchange at room temperature for 2 h. All buffers used for *holo*FNR purification were prepared aerobically, gas exchanged with argon for at least 2 h. Following oxygen removal, cells were transferred to a 4 °C MBraun anaerobic chamber for purification (kept under a nitrogen environment, O_2 < 0.1 ppm) and resuspended in 50 mM potassium phosphate (pH 6.8), 100 mM KCl, 20% (w/v) glycerol, 2 mM DTT, and cOmplete EDTA-free protease inhibitor cocktail (Roche). Cells were lysed by sonication, followed by centrifugation at 14,000 x *g* for 45 min at 4 °C and syringe-filtration (0.45 µm). The supernatant was then applied to a 5 mL HiTrap Heparin HP column (GE Healthcare), which was washed with the resuspension buffer for 20 CV. *Holo*FNR was eluted in a 10% step gradient of lysis buffer to 100% elution buffer (50 mM potassium phosphate (pH 6.8), 1 M KCl, and 20% (w/v) glycerol). Fractions containing *holo*FNR were analyzed by SDS-PAGE, pooled, and concentrated using 50 kDa Amicon filters (Millipore-Sigma), prior to SEC using a HiLoad 16/600 Superdex-75 pg column in 25 mM HEPES (pH 6.8), 150 mM KCl, and 2 mM DTT. Fractions containing *holo*FNR from SEC were analyzed by SDS-PAGE, pooled, concentrated with 50 kDa Amicon filters, and kept at 4 °C in the MBraun anaerobic chamber.

ClpX, ClpX$^{\Delta N}$, the ClpX N domain, and ClpP were purified using established methods (Kim et al., 2000; Neher et al., 2003; Abdelhakim et al., 2008; Chowdhury et al., 2010). To produce ^{35}S-labeled ClpX N domain, cells were grown in M9 minimal media lacking methionine to an OD_{600} of 0.6 with 0.5 mM IPTG for 2 h. EasyTag L-[^{35}S]-methionine was added to 20 µCi/mL of culture, and cells were grown for an addition 2 h before harvesting. Cells were resuspended in 50 mM sodium phosphate (pH 8), 300 mM NaCl, 5 mM imidazole, 10% (w/v) glycerol and lysed by the addition of lysozyme and protease inhibitor cocktail set III, EDTA-free (Calbiochem) and three freeze-thaw cycles. After centrifugation at 20,000 x *g* for 15 min at 4 °C, the supernatant was incubated with Ni-NTA resin for 2 h at 4 °C. The resin was then washed with 15 CV lysis buffer and 30 CV lysis buffer + 20 mM imidazole and resuspended in 1 CV lysis buffer prior to incubation with biotinylated thrombin (Novagen) overnight. Cleaved protein was recovered following incubation with streptavidin agarose resin and concentrated with the spin filters provided in the Thrombin Cleavage Capture Kit (Novagen).

Plasmids for GFP variants were generated by site-directed mutagenesis of GFPmut3 to introduce F64L, R80Q, and A206K mutations, with a C-terminal His_6-Arg-Pro-ssrA sequence cloned into pET23b (Novagen). GFP proteins were overexpressed in JK10 cells (Kenniston et al., 2005) and purified by Ni-NTA chromatography and Superdex-75 gel filtration. Plasmids for DHFR variants were generated by site-directed mutagenesis of mouse DHFR fused to the 3´ end of His_6-Sumo cloned in pET23b. DHFR proteins were overexpressed in JK10 cells and purified by Ni-NTA chromatography, followed by ULP1 cleavage, Ni-NTA, and size-exclusion chromatography as described (Rivera-Rivera et al., 2014). GFP and DHFR proteins were stored in 20 mM HEPES-KOH (pH 7.5), 150 mM KCl, 10% (w/v) glycerol, and 1 mM DTT.

FNR peptides were generated using standard Fmoc techniques using an Apex 396 solid-phase instrument.

Peptide-blot screening for ClpX N-domain binding

The FNR peptide array was generated by synthesizing overlapping 12 residue peptides and covalent attachment to a cellulose membrane. Each adjacent spot corresponded to a peptide initiated two residues C-terminal to the preceding peptide in the FNR sequence. The blot was incubated with ^{35}S-ClpX N-domain in 25 mM HEPES-KOH (pH 7.6), 20 mM KCl, 5 mM $MgCl_2$, 0.032% NP-40, 10% (w/v) glycerol and washed with PBS. Following overnight exposure, the radioactive signals were detected by Molecular Dynamics PhosphorImager scanning and imaged using a Typhoon FLA 9500 instrument (GE Healthcare).

Fluorescence anisotropy

Solution binding of the FNR peptide, MIPEKRIIRRIQSGGCAIH-dansyl, was performed at 30 °C in duplicate or triplicate by measuring changes in fluorescence anisotropy (327 nm excitation; 550 nm emission) using a Photon Technology International fluorimeter. Solution binding of the SspB XB peptide, RGGRPALRVVK-fluorescein, was performed under similar conditions and optimized for the fluorescein moiety (495 nm excitation; 520 nm emission) Each fluorescent peptide (1 μM) and varying concentrations of the ClpX N domain in 20 mM HEPES-KOH (pH 8.0), 50 mM KCl, 10% (w/v) glycerol were incubated until reaching equilibrium. Data were fit to a hyperbolic isotherm using non-linear least-squares regression analysis:

$$\text{fluorescence anisotropy} = f_{min} + \left(\frac{f_{max} * K_{\mathrm{D}}}{K_D + X}\right)$$

where f_{min} is the background anisotropy value, f_{max} is the maximum anisotropy value at saturated binding, K_{D} is the dissociation equilibrium constant (in μM), and X is the concentration of ClpX N domain (in μM).

Gel degradation assays

Degradation assays (for Figures 5.2, 5.S1–3) were performed at 30 °C by preincubating $ClpX_6$ or $ClpX^{\Delta N}{}_6$ and $ClpP_{14}$ with 5 mM ATP in 25 mM HEPES-KOH (pH 7.5), 20 mM KCl, 5 mM $MgCl_2$,

10% (w/v) glycerol, 0.032% NP-40, and 0.2 mM DTT and adding substrate to initiate the reactions. For Figure 5.4, degradation assays were performed under similar conditions, except with an ATP-regeneration system (4 mM ATP, 50 μg/mL creatine kinase, 5 mM creatine phosphate). *apo*FNR^{D154K} variant degradation assays were performed in 25 mM HEPES-KOH (pH 7.5), 100 mM KCl, 5 mM $MgCl_2$, 10% (w/v) glycerol, 0.032% NP-40, and 0.2 mM DTT, with an ATP-regeneration system (4 mM ATP, 50 μg/mL creatine kinase, 5 mM creatine phosphate). Samples of each reaction were taken at specific time points and stopped by addition of SDS-PAGE loading dye and boiling at 100 °C before loading on Tris-Glycine-SDS gels. Bands were visualized by staining with colloidal Coomassie G-250 and quantified by ImageQuant (GE Healthcare) after scanning using a Typhoon FLA 9500 instrument (GE Healthcare). The fraction of FNR remaining was calculated by dividing the FNR density at a given time point by the density at time zero and was normalized by the ClpP density.

GFP degradation assays

GFP degradation was monitored by loss of fluorescence (485 nm excitation; 511 nm emission) over a 5 min time course. Prior to initiation with the ATP-regeneration system (4 mM ATP, 50 μg/mL creatine kinase, 5 mM creatine phosphate), $ClpX_6$, $ClpP_{14}$, and varying amounts of GFP substrate were incubated for 3 min at 30 °C. Average initial rates of degradation were fit by non-linear least-squares regression to the Michaelis-Menten equation to obtain K_M and V_{max} values.

Western blotting

Formation of DHFR-FNR fusion protein truncation products during degradation reactions were monitored by SDS-PAGE and Western blotting. Proteins were wet-transferred to Immobilin-P membranes (EMD Millipore). Membranes were probed with a 1:10,000 dilution of anti-DHFR A-9 (Santa-Cruz) in 1X TBST with 5% milk overnight at 4 °C, followed by incubating with goat anti-mouse immunoglobulin G-alkaline phosphatase conjugate (Bio-Rad; 1:15,000 dilution) for 1 h at

room temperature, washing, and exposure with ECF substrate (GE Healthcare). Blots were detected by scanning using a Typhoon FLA 9500 instrument (GE Healthcare).

*Apo*FNR and *holo*FNR crystallization

Crystals of *holo*FNRΔ14/L28H were grown at room temperature in an MBraun anaerobic chamber (kept under a nitrogen environment, O_2 < 0.1 ppm) by the sitting-drop method after mixing 1 µL of *holo*FNRΔ14/L28H and an equal volume of well solution (0.1 M HEPES (pH 7.5), 10% w/v PEG 8,000, and 10% (w/v) ethylene glycol). After *holo*FNR∆14/L28H crystals were fully formed (~2 days), crystals were transferred to a COY anaerobic chamber (under an argon-hydrogen environment, O_2 = 1–20 ppm), cryo-protected in 100% paraffin oil, and cryo-cooled by rapid submersion in liquid nitrogen. After several unsuccessful attempts to crystallize *apo*FNR, which is prone to precipitation at high concentrations, *holo*FNRΔ14/L28H crystals were grown by the sitting-drop method and exposed to acidic conditions. 1 µL of 7.5 mg/mL *holo*FNRΔ14/L28H was mixed with an equal volume of well solution (2% (w/v) tacsimate (pH 5.0), 0.1 M sodium citrate tribasic dihydrate (pH 5.6), and 16% (w/v) polyethylene glycol (PEG) 3,350) to generate *apo* crystals. Loss of the [4Fe-4S] cluster in the crystals was monitored over 5 days, resulting in a colorless appearance. *apo*FNRΔ14/L28H crystals were removed from the MBraun anaerobic chamber and harvested under atmospheric conditions, cryo-protected in its well solution with 25% (w/v) glycerol, and cryo-cooled by rapid submersion in liquid nitrogen. All solutions used for crystal optimization and cryo-protection were purchased as prepared solutions from Hampton Research.

Data collection and structure determination

Apo and *holo*FNR crystals were screened in-house using a rotating copper anode diffractometer. Data collection was performed at the Northeastern Collaborative Access Team, Advance Photon Source (APS) on beam-lines IDE (*apo*FNRΔ14/L28H) and IDC (*holo*FNRΔ14/L28H). X-ray data were processed using XDS (Kabsch, 2010a, 2010b), and structures were determined by

molecular replacement using PHENIX Phaser-MR (McCoy et al., 2007). Both structures were solved using molecular replacement using a poly-alanine chain of monomeric *E. coli apo*FNRΔ4 (C. Nguyen *et al.*, personal communication). The initial refinement in PHENIX (Adams et al., 2010) of the molecular replacement solution of *apo*FNRΔ14/L28H at 1.43 Å resolution structure yielded an R_{work} of 34.06% and an R_{free} of 34.47%. Initial refinement of the 1.64 Å resolution *holo*FNRΔ14/L28H structure in PHENIX yielded an R_{work} of 36.85% and an R_{free} of 39.54%.

Model building was performed in Coot (Emsley and Cowtan, 2004), and PHENIX was used for real-space refinement, with non-crystallographic symmetry (NCS) restraints applied for *apo* and *holo* FNRΔ14/L28H structural refinements. Positional and individual B-factor refinement continued at the full resolution until the model was complete. Waters were added by the "Update Water" function in PHENIX and visual inspection in Coot after all the amino acids and [4Fe-4S] clusters were refined. Parameter files for the [4Fe-4S] cluster during refinement were generated using PHENIX eLBOW (Moriarty et al., 2009). Composite omit maps, generated in PHENIX, were also used to verify the FNR structures during model building and refinement. Model geometry was assessed using MolProbity (Williams et al., 2018). Figures were generated in ChimeraX (Pettersen et al., 2021).

The 1.43 Å resolution *apo*FNRΔ14 structure contains two molecules of FNR per asymmetric unit with clear density for residues 15–247 (chain A) and 15–250 (chain B). The 1.64 Å resolution *holo*FNRΔ14 structure contains two FNR chains with clear density for residues 15–248 (chain A) and 15–248 (chain B). Crystallographic software used was compiled by SBGrid (Morin et al., 2013). Statistics for model refinement and final crystallographic models are provided in Table 5.2. PDB accession numbers are pending assignment.

ACKNOWLEDGMENTS

We thank R. Grant for assistance with in-house screening; P. Kiley, E. Mettert, and A. Goren for helpful advice; and I. Levchenko for peptides.

This work was supported by NIH grants AI-016892 (R.T.S., T.A.B.), R35-GM126982 (C.L.D.), and T32 GM007287 (S.K.); the Howard Hughes Medical Institute (T.A.B., C.L.D.); the National Science Foundation Graduate Research Fellowship grant 1745302 (S.K., L.R.F.B.); the Helen Hay Whitney Foundation Fellowship (C.N.); the Dow Fellowship at MIT (L.R.F.B.); and an HHMI Gilliam Fellowship (L.R.F.B.). Support for research was also provided by NIH EHS grant P30-ES002109 and GMS grant P41 GM103403 (conducted at the Northeastern Collaborative Access Team beamlines). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

Abdelhakim, A.H., Oakes, E.C., Sauer, R.T., and Baker, T.A. (2008). Unique contacts direct high-priority recognition of the tetrameric transposase-DNA complex by the AAA+ unfoldase ClpX. *30*, 1, 39–50.

Adams, P.D., Afonine, P. V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 2, 213–221.

Ahn, B.-E., and Baker, T.A. (2016). Oxidization without substrate unfolding triggers proteolysis of the peroxide-sensor, PerR. Proc. Natl. Acad. Sci. U. S. A. *113*, 1, E23-31.

Ainavarapu, S.R.K., Li, L., Badilla, C.L., and Fernandez, J.M. (2005). Ligand binding modulates the mechanical stability of dihydrofolate reductase. Biophys. J. *89*, 5, 3337–3344.

Baez, A., and Shiloach, J. (2014). Effect of elevated oxygen concentration on bacteria, yeasts, and cells propagated for production of biological compounds. Microb. Cell Fact. *13*, 181.

Bagnall, J., Leedale, J., Taylor, S.E., Spiller, D.G., White, M.R.H., Sharkey, K.J., Bearon, R.N., and Sée, V. (2014). Tight control of hypoxia-inducible factor-α transient dynamics is essential for cell survival in hypoxia. J. Biol. Chem. *289*, 9, 5549–5564.

Baker, T.A., and Sauer, R.T. (2012). ClpXP, an ATP-powered unfolding and protein-degradation machine. Biochim. Biophys. Acta - Mol. Cell Res. *1823*, 1, 15–28.

Barbieri, N.L., Nicholson, B., Hussein, A., Cai, W., Wannemuehler, Y.M., Dell'Anna, G., Logue, C.M., Horn, F., Nolan, L.K., and Li, G. (2014). FNR regulates expression of important virulence factors contributing to pathogenicity of uropathogenic *Escherichia coli*. Infect. Immun. *82*, 12, 5086–5098.

Barbieri, N.L., Vande Vorde, J.A., Baker, A.R., Horn, F., Li, G., Logue, C.M., and Nolan, L.K. (2017). FNR regulates the expression of important virulence factors contributing to the pathogenicity of avian pathogenic *Escherichia coli*. Front. Cell. Infect. Microbiol. *7*, 265.

Boshoff, H.I.M., and Barry, C.E. (2005). Tuberculosis—metabolism and respiration in the absence of growth. Nat. Rev. Microbiol. *3*, 1, 70–80.

Camberg, J.L., Hoskins, J.R., and Wickner, S. (2009). ClpXP protease degrades the cytoskeletal protein, FtsZ, and modulates FtsZ polymer dynamics. Proc. Natl. Acad. Sci. U. S. A. *106*, 26, 10614–10619.

Camberg, J.L., Viola, M.G., Rea, L., Hoskins, J.R., and Wickner, S. (2014). Location of dual sites in E. coli FtsZ important for degradation by ClpXP; one at the C-terminus and one in the disordered linker. PLoS One *9*, 4, e94964.

Chien, P., Perchuk, B.S., Laub, M.T., Sauer, R.T., and Baker, T.A. (2007). Direct and adaptor-mediated substrate recognition by an essential AAA protease.

Chowdhury, T., Chien, P., Ebrahim, S., Sauer, R.T., and Baker, T.A. (2010). Versatile modes of peptide recognition by the ClpX N domain mediate alternative adaptor-binding specificities in different bacterial species. Protein Sci. *19*, 2, 242–254.

Emsley, P., and Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. *60*, 12 I, 2126–2132.

Fei, X., Bell, T.A., Barkow, S.R., Baker, T.A., and Sauer, R.T. (2020). Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. eLife *9*, e61496, 1–39.

Flynn, J.M., Neher, S.B., Kim, Y.I., Sauer, R.T., and Baker, T.A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. Mol. Cell *11*, 3, 671–683.

Green, J., Crack, J.C., Thomson, A.J., and LeBrun, N.E. (2009). Bacterial sensors of oxygen. Curr. Opin. Microbiol. *12*, 2, 145–151.

Guest, J.R., Green, J., Irvine, A.S., and Spiro, S. (1996). The FNR modulon and FNR-regulated gene expression. In: Regulation of Gene Expression in *Escherichia coli*. Springer, Boston, MA. pp. 317–342.

Gur, E., Vishkautzan, M., and Sauer, R.T. (2012). Protein unfolding and degradation by the AAA+ Lon protease. Protein Sci. *21*, 2, 268–278.

Harris, A.L. (2002). Hypoxia — a Key Regulatory Factor in Tumour Growth. Nat. Rev. Cancer *2*, 1, 38–47.

Jacquamet, L., Traoré, D.A.K., Ferrer, J.L., Proux, O., Testemale, D., Hazemann, J.L., Nazarenko, E., El Ghazouani, A., Caux-Thang, C., Duarte, V., et al. (2009). Structural characterization of the active form of PerR: Insights into the metal-induced activation of PerR and Fur proteins for DNA binding. Mol. Microbiol. *73*, 1, 20–31.

Jervis, A.J., and Green, J. (2007). In vivo demonstration of FNR dimers in response to lower O(2) availability. J. Bacteriol. *189*, 7, 2930–2932.

Jervis, A.J., Crack, J.C., White, G., Artymiuk, P.J., Cheesman, M.R., Thomson, A.J., Brun, N.E.L., and Green, J. (2009). The O_2 sensitivity of the transcription factor FNR is controlled by Ser24 modulating the kinetics of [4Fe-4S] to [2Fe-2S] conversion. Proc. Natl. Acad. Sci. U. S. A. *106*, 12, 4659–4664.

Jones, S.A., Chowdhury, F.Z., Fabich, A.J., Anderson, A., Schreiner, D.M., House, A.L., Autieri, S.M., Leatham, M.P., Lins, J.J., Jorgensen, M., et al. (2007). Respiration of *Escherichia coli* in the mouse intestine. Infect. Immun. *75*, 10, 4891–4899.

Kabsch, W. (2010a). XDS. Acta Crystallogr. D. Biol. Crystallogr. *66*, Pt 2, 125–132.

Kabsch, W. (2010b). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. Sect. D Biol. Crystallogr. *66*, 2, 133–144.

Kang, Y., Weber, K.D., Qiu, Y., Kiley, P.J., and Blattner, F.R. (2005). Genome-wide expression analysis indicates that FNR of *Escherichia coli* K-12 regulates a large number of genes of unknown function. J. Bacteriol. *187*, 3, 1135–1160.

Kenniston, J.A., Baker, T.A., and Sauer, R.T. (2005). Partitioning between unfolding and release of native domains during ClpXP degradation determines substrate selectivity and partial processing. Proc. Natl. Acad. Sci. U. S. A. *102*, 5, 1390–1395.

Kim, Y.I., Burton, R.E., Burton, B.M., Sauer, R.T., and Baker, T.A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. Mol. Cell *5*, 4, 639–648.

Kohn, K.W., Riss, J., Aprelikova, O., Weinstein, J.N., Pommier, Y., and Barrett, J.C. (2004). Properties of switch-like bioregulatory networks studied by simulation of the hypoxia response control system. Mol. Biol. Cell *15*, 7, 3042–3052.

Koodithangal, P., Jaffe, N.E., Kraut, D.A., Fishbain, S., Herman, C., and Matouschek, A. (2009). ATP-dependent proteases differ substantially in their ability to unfold globular proteins. J. Biol. Chem. *284*, 28, 18674–18684.

Lazazzera, B.A., Bates, D.M., and Kiley, P.J. (1993). The activity of the *Escherichia coli* transcription factor FNR is regulated by a change in oligomeric state. Genes Dev. *7*, 10, 1993–2005.

Lazazzera, B.A., Beinert, H., Khoroshilova, N., Kennedy, M.C., Kiley, P.J., Lazazzera, B.A., Kiley, P.J., and Beinert, H. (2002). DNA binding and dimerization of the Fe-S-containing FNR protein from *Escherichia coli* are regulated by oxygen. J. Biol. Chem. *271*, 5, 2762–2768.

Lee, J.-W., and Helmann, J.D. (2006). The PerR transcription factor senses H_2O_2 by metal-catalysed histidine oxidation. Nature *440*, 7082, 363–367.

Lee, C., Schwartz, M.P., Prakash, S., Iwakura, M., and Matouschek, A. (2001). ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. Mol. Cell *7*, 3, 627–637.

Ling, L., Montaño, S.P., Sauer, R.T., Rice, P.A., and Baker, T.A. (2015). Deciphering the Roles of Multicomponent Recognition Signals by the AAA+ Unfoldase ClpX. J. Mol. Biol. *427*, 18, 2966–2982.

Mahmoud, S.A., and Chien, P. (2018). Regulated proteolysis in bacteria. Annu. Rev. Biochem. *87*, 677–696.

Marteyn, B., Scorza, F.B., Sansonetti, P.J., and Tang, C. (2011). Breathing life into pathogens: the influence of oxygen on bacterial virulence and host responses in the gastrointestinal tract. Cell. Microbiol. *13*, 2, 171–176.

Martin, A., Baker, T.A., and Sauer, R.T. (2008). Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. Mol. Cell *29*, 4, 441–450.

Marxsen, J.H., Stengel, P., Doege, K., Heikkinen, P., Jokilehto, T., Wagner, T., Jelkmann, W., Jaakkola, P., and Metzen, E. (2004). Hypoxia-inducible factor-1 (HIF-1) promotes its degradation by induction of HIF-α-prolyl-4-hydroxylases. Biochem. J. *381*, 3, 761–767.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 4, 658–674.

Mettert, E.L., and Kiley, P.J. (2005). ClpXP-dependent proteolysis of FNR upon loss of its O_2-sensing [4Fe–4S] cluster. J. Mol. Biol. *354*, 2, 220–232.

Mettert, E.L., and Kiley, P.J. (2018). Reassessing the structure and function relationship of the O2 sensing transcription factor FNR. Antioxidants Redox Signal. *29*, 18, 1830–1840.

Mettert, E.L., Outten, F.W., Wanta, B., and Kiley, P.J. (2008). The impact of O_2 on the Fe-S cluster biogenesis requirements of *Escherichia coli* FNR. J. Mol. Biol. *384*, 4, 798–811.

Moore, L.J., and Kiley, P.J. (2001). Characterization of the dimerization domain in the FNR transcription factor. J. Biol. Chem. *276*, 49, 45744–45750.

Moore, L.J., Mettert, E.L., and Kiley, P.J. (2006). Regulation of FNR dimerization by subunit charge repulsion. J. Biol. Chem. *281*, 44, 33268–33275.

Moriarty, N.W., Grosse-Kunstleve, R.W., and Adams, P.D. (2009). Electronic ligand builder and optimization workbench (eLBOW): A tool for ligand coordinate and restraint generation. Acta Crystallogr. Sect. D Biol. Crystallogr. *65*, 10, 1074–1080.

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P.C., Timony, M.A., Ottaviano, M., and Sliz, P. (2013). Collaboration gets the most out of software. eLife *2*, e01456.

Neher, S.B., Sauer, R.T., and Baker, T.A. (2003). Distinct peptide signals in the UmuD and UmuD' subunits of UmuD/D' mediate tethering and substrate processing by the ClpXP protease. Proc. Natl. Acad. Sci. *100*, 23, 13219–13224.

Ohh, M., Park, C.W., Ivan, M., Hoffman, M.A., Kim, T.Y., Huang, L.E., Pavletich, N., Chau, V., and Kaelin, W.G. (2000). Ubiquitination of hypoxia-inducible factor requires direct binding to the β-domain of the von Hippel - Lindau protein. Nat. Cell Biol. *2*, 7, 423–427.

Olivares, A.O., Baker, T.A., and Sauer, R.T. (2015). Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. Nat. Rev. Microbiol. *14*, 1, 33–44.

Pan, Q., Shan, Y., and Yan, A. (2012). A Region at the C-Terminus of the *Escherichia coli* Global Transcription Factor FNR Negatively Mediates Its Degradation by the ClpXP Protease. Biochemistry *51*, 25, 5061–5071.

Partridge, J.D., Sanguinetti, G., Dibden, D.P., Roberts, R.E., Poole, R.K., and Green, J. (2007). Transition of *Escherichia coli* from aerobic to micro-aerobic conditions involves fast and slow reacting regulatory components. J. Biol. Chem. *282*, 15, 11230–11237.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. *30*, 1, 70–82.

Qutub, A.A., and Popel, A.S. (2006). A computational model of intracellular oxygen sensing by hypoxia-inducible factor HIF1α. J. Cell Sci. *119*, 16, 3467–3480.

Rivera-Rivera, I., Román-Hernández, G., Sauer, R.T., and Baker, T.A. (2014). Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. Proc. Natl. Acad. Sci. U. S. A. *111*, 37, E3853-9.

Salceda, S., and Caro, J. (1997). Hypoxia-inducible factor 1α (HIF-1α) protein is rapidly degraded by the ubiquitin-proteasome system under normoxic conditions. Its stabilization by hypoxia depends on redox-induced changes. J. Biol. Chem. *272*, 36, 22642–22647.

Salmon, K., Hung, S., Mekjian, K., Baldi, P., Hatfield, G.W., and Gunsalus, R.P. (2003). Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and FNR. J. Biol. Chem. *278*, 32, 29837–29855.

Sauer, R.T., and Baker, T.A. (2011). AAA+ Proteases: ATP-fueled machines of protein destruction. Annu. Rev. Biochem. *80*, 1, 587–612.

Sauer, R.T., Fei, X., Bell, T.A., and Baker, T.A. (2021). Structure and function of ClpXP, a AAA+ proteolytic machine powered by probabilistic ATP hydrolysis. Crit. Rev. Biochem. Mol. Biol. 1–17.

Semenza, G.L. (2004). Hydroxylation of HIF-1: Oxygen sensing at the molecular level. Physiology *19*, 4, 176–182.

Spiro, S., and Guest, J.R. (1987). Regulation and over-expression of the *fnr* gene of *Escherichia coli.* J. Gen. Microbiol. *133*, 12, 3279–3288.

Studemann, A., Noirclerc-Savoye, M., Klauck, E., Becker, G., Schneider, D., and Hengge, R. (2003). Sequential recognition of two distinct sites in σS by the proteolytic targeting factor RssB and ClpX. EMBO J. *22*, 16, 4111–4120.

Sutton, V.R., Stubna, A., Patschkowski, T., Münck, E., Beinert, H., and Kiley, P.J. (2004). Superoxide destroys the $[2Fe\text{-}2S]^{2+}$ cluster of FNR from *Escherichia coli*. Biochemistry *43*, 3, 791–798.

Tolla, D.A., and Savageau, M.A. (2010). Regulation of aerobic-to-anaerobic transitions by the FNR cycle in *Escherichia coli*. J. Mol. Biol. *397*, 4, 893–905.

Tolla, D.A., Kiley, P.J., Lomnitz, J.G., and Savageau, M.A. (2015). Design principles of a conditional futile cycle exploited for regulation. Mol. Biosyst. *11*, 7, 1841–1849.

Traoré, D.A.K., Ghazouani, A. El, Jacquamet, L., Borel, F., Ferrer, J.L., Lascoux, D., Ravanat, J.L., Jaquinod, M., Blondin, G., Caux-Thang, C., et al. (2009). Structural and functional characterization of 2-oxo-histidine in oxidized PerR protein. Nat. Chem. Biol. *5*, 1, 53–59.

Wah, D.A., Levchenko, I., Rieckhof, G.E., Bolon, D.N., Baker, T.A., and Sauer, R.T. (2003). Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the AAA+ ClpXP protease. Mol. Cell *12*, 2, 355–363.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci. *27*, 1, 293–315.

Yan, A., and Kiley, P.J. (2009). Chapter 42 Techniques to Isolate O_2-Sensitive Proteins. In Methods in Enzymology, pp. 787–805.

Zheng, L., Cash, V.L., Flint, D.H., and Dean, D.R. (1998). Assembly of iron-sulfur clusters. Identification of an *iscSUA-hscBA-fdx* gene cluster from *Azotobacter vinelandii*. J. Biol. Chem. *273*, 21, 13264–13272.

Zhou, Y., Gottesman, S., Hoskins, J.R., Maurizi, M.R., and Wickner, S. (2001). The RssB response regulator directly targets sigma(S) for degradation by ClpXP. Genes Dev. *15*, 5, 627–637.

CHAPTER 6

Conclusions and Perspectives

OVERVIEW

In this chapter, I will provide my outlook on the results described in Chapters 2–5. The conclusions described in the discussion section of each chapter raise additional questions that should be addressed in future studies. I will also discuss the biological implications of my research findings, informing our current understanding of AAA+ proteolytic mechanism and the role of AAA+ proteases in protein quality control and regulated proteolysis.

CONFORMATIONAL PLASTICITY IN THE INTERFACE BETWEEN THE AAA+ UNFOLDASE AND THE PEPTIDASE

In Chapter 2, I present work showing that rotation of the AAA+ unfoldase $ClpA_6$ with respect to the ClpP heptameric ring is not required for proteolytic function. However, these experiments also established that crosslinking ClpA to ClpP subunits slows unfolding and degradation rates of ssrA-tagged substrates, especially for highly stable proteins, suggesting that movement of the ClpA–ClpP interface is important for efficient proteolysis. Multiple structures of ClpXP and ClpAP highlight the conformational plasticity of IGF (in ClpX) and IGL (in ClpA) loops used to bind ClpP clefts (Gatsogiannis et al., 2019; Fei et al., 2020b, 2020a; Lopez et al., 2020; Ripstein et al., 2020). These flexible IGF/L loops are critical for the AAA+ unfoldase to reach the binding clefts in its partner peptidase, and thus, to overcome challenges in protease assembly due to the hexamer–heptamer symmetry mismatch (Martin et al., 2007; Amor et al., 2019). Movement between the unfoldase and peptidase is also observed in genetically-tethered AAA+ proteases, such as Lon, which do not feature a subunit symmetry mismatch. Bacterial and human Lon and other AAA+ proteases, such as Yme1 (in mitochondria) and FtsH, contain a flexible interdomain linker that is proposed to facilitate polypeptide translocation by enabling movement of the AAA+ unfoldase relative to the planar peptidase domains (Bieniossek et al., 2009; Puchades et al., 2017; Shin et al., 2020, 2021). Consequently, conformational flexibility between the two components of a AAA+

protease appears to be a conserved feature shared by enzymes that are expressed by a single gene or by multiple genes.

Real-time kinetic assays of ClpXP assembly and disassembly demonstrate that interactions between ClpX and ClpP are highly dynamic, suggesting that transient unbinding/rebinding of IGF loops occurs frequently (Amor et al., 2016). Furthermore, deletion of even a single ClpX IGF loop results in slower degradation of both unfolded and native substrates, which occurs more slowly as more loops are removed (Joshi et al., 2004; Martin et al., 2007; Amor et al., 2019). Unfolding assays performed in the absence of ClpP elucidated that the impaired degradation of ClpX IGF-loop mutants likely arises due to slower polypeptide translocation but not reduced rates of protein unfolding (Amor et al., 2019). To characterize the crosslinked ClpA–ClpP complexes, which contain a static IGL–ClpP cleft interface due to multiple crosslinked subunits, future studies should probe the reactions steps responsible for the slower rates of substrate degradation *i.e.* protein unfolding or polypeptide translocation using similar kinetic methods described in Amor *et al.* 2016, 2019 or through single-molecule optical trapping.

PORE-LOOP CONFORMATIONS DURING AAA+ ENZYME OPERATION

A subset of ClpAPS delivery complex cryo-EM structures (~20% of the final particles) features tucked pore-loop conformations in the D2 ring (described in Chapter 3). As discussed in this chapter, the rotation and altered orientation of the key aromatic pore-1 loop residue in most of the D2 AAA+ modules in the ClpA hexamer suggests that the D2 pore-1 loop is an ‘inactive’ state (does not bind the NTE polypeptide). The presence of tucked pore loops correlates with a lack of observable NTE density in the D2, but not the D1 portion of the ClpA channel. Thus, these ClpAPS structures likely capture ClpA partially releasing the ClpS NTE or an assembly intermediate (‘half-engaged’) during the process of binding the ClpS NTE in both D1 and D2 rings. Briefly (to summarize the discussion in Chapter 3), transient loss of polypeptide contact via tucked pore

loops could occur during enzyme pausing by double-ring enzymes, such as ClpA, which would need to temporarily disengage from the polypeptide in the stalled AAA+ ring while allowing the other ring to continue unfolding/translocation. Single-molecule optical trapping studies established this scenario likely occurs to promote processivity in ClpAP; frequent pausing is observed in a D1 ATPase-inactive ClpA variant, indicating that the D1 ring rescues stalling in the D2 major motor (Kotamarthi et al., 2020). Furthermore, pore loop tucking could be used more generally as a reaction step in enzyme dissociation from the substrate polypeptide *e.g.*, during failed unfolding attempts and substrate release, which occurs frequently for stable native substrates (Kenniston et al., 2005). As substrate release is a required step for AAA+ disaggregases, including the ClpA homologs ClpB/Hsp104 (Shorter and Southworth, 2019), the pore-loop tucking mechanism revealed in our structures may be especially relevant for this class of AAA+ unfoldases/remodeling enzymes that do not processively translocate and degrade substrate proteins.

Although there are many mechanistic implications of ClpA pore loops in a partially released or partially bound state, these cryo-EM structures (found in a single AAA+ protease) are insufficient evidence that tucked pore loops are universal features of all AAA+ molecular machines. One possibility explaining why pore-loop tucking has not yet been reported in previous structures of AAA+ enzymes could be due to limitations of current cryo-EM methodology. Reconstructions of continuous heterogeneity, such as the molecular motions involved in substrate recognition, unfolding, and translocation, remain challenging and an area of technological development (Nakane et al., 2018; Punjani and Fleet, 2021; Zhong et al., 2021). Because our structures were assembled in ATPγS and represent initial stages of adaptor-assisted N-end-rule substrate delivery, our ClpAPS complexes were likely more homogenous than those in prior studies, allowing us to better elucidate ClpS NTE interactions with the ClpA axial pore loops in molecular detail. Future cryo-EM studies in other AAA+ proteolytic machines will illuminate whether this new

pore-loop conformation is also observed in other AAA+ proteases, limited to double-ring enzymes, or specific to ClpAPS delivery complexes.

MECHANISMS OF SUBSTRATE RECOGNITION BY ClpAP

In Chapter 3, I compare the polypeptide in the axial channel of ClpAP cryo-EM structures assembled in the presence of ATPγS with ClpS and an N-end-rule substrate to those from a previous study with ClpAP, ATPγS, and a directly recognized substrate (Lopez et al., 2020). In both sets of structures, the axial pore loops of both D1 and D2 rings bind the polypeptide (either the ClpS NTE or the substrate), indicating that ATP-fueled conformational changes are not required for pore-loop engagement throughout the ClpA channel. These data also support biochemical studies demonstrating the broad substrate recognition of ClpAP and its biological function in general protein quality control (Hoskins et al., 1998, 2000; Singh et al., 2000). Because ClpA appears to bind a wide range of peptides and even functions with ClpS bearing an NTE substituted with a Gly-Ser-Lys linker, perhaps the ClpA channel discriminates very little between specific peptide sequences.

However, this permissive open channel state of ClpA presents a paradox: how does ClpA efficiently recognize specific degrons? ClpAP recognizes N-end-rule substrates weakly in the absence of ClpS, as K_M values exceed 10 μM for intrinsic ClpAP degradation of substrates tested in a previous study (Wang et al., 2007). By contrast, ClpS-assisted degradation of these substrates tightens the K_M ~10 to 70-fold, supporting the role of ClpS as a specificity enhancement factor (Wang et al., 2007). Thus, perhaps specific recognition by ClpAP is mediated exclusively by ClpS, in which its hydrophobic binding pocket selects for strong N-degrons containing primary destabilizing N-end residues. Because the ClpS NTE binds to ClpA axial pore loops (Chapter 3), substrates that rely on direct binding to the ClpA pore loops, such as those bearing the ssrA tag,

would be excluded in the presence of ClpS. ClpS therefore would function as a rheostat modulating the substrate specificity of ClpAP.

Nevertheless, a critical flaw remains unaddressed in this rationale that ClpA specificity relies solely on ClpS delivery. ClpAP degrades ssrA-tagged substrates with a K_M of 1.5 – 4 μM and V_{max} of 3 – 5 min^{-1} $ClpA_6^{-1}$, rivaling kinetic parameters of degradation by ClpXP (a highly specific AAA+ protease) of the same class of substrates (Flynn et al., 2001; Kim et al., 2020). This efficient recognition of ssrA-tagged substrates by ClpAP cannot be accounted by generic recognition of an unstructured peptide, as demonstrated by the weak intrinsic recognition of N-end-rule substrates. Although ClpA does not appear to recognize a specific region of the ssrA tag, mutations in distinct residues of the ssrA tag, AANDENYALAA-COO^- (underlined residues), increase the K_M for degradation of GFP-ssrA variants (Flynn et al., 2001). To elucidate the mechanism of specific recognition by ClpAP, future studies should identify the chemical properties of key ssrA residues required for ClpA recognition, expanding upon a previous study that mutagenized each Ala residue to Asp and every non-Ala residue to Ala (Flynn et al., 2001). Additional cryo-EM structures using an ssrA-tagged substrate may also illuminate the structural basis of ClpA interactions with the ssrA tag, as is the case for the ClpXP–ssrA recognition complex structure (Fei et al., 2020b). More broadly, a screening-based approach (either by *in vivo* mutagenesis or generation of a candidate degron library) could be used to identify strong intrinsic ClpA degrons and consensus motifs, and thus, to better classify whether a given protein is degraded by ClpAP via ‘specific’ (distinct recognition of degrons) or ‘nonspecific’ (general recognition of an unstructured polypeptide) pathways.

MECHANISM OF ClpS-ASSISTED DEGRADATION OF N-END-RULE SUBSTRATES

The cryo-EM structures and biochemical experiments using D1 ring pore-2 loop variants provide insight into two steps in the ClpS-assisted mechanism of N-end-rule substrate delivery: *(i)* initial

formation of the high-affinity delivery complex containing ClpAPS•N-end-rule substrate and (*ii)* remodeling of the ClpS core during substrate transfer. In our model of ClpS-assisted substrate delivery, we propose that ClpA must remodel ClpS and then also unfold native structure in the N-end-rule substrate (Chapter 3, pp. 108–112). To better understand ClpS remodeling, single-molecule dual optical trapping could be performed, in which a ClpS chimera is attached to one bead and ClpAP to the other (**Figure 6.1A**). Since a ClpS variant bearing a duplicated NTE is engaged, translocated, and clipped by ClpAP (Rivera-Rivera et al., 2014), a single-molecule assay with this NTE_2-ClpS could monitor conformational changes in ClpS and attempts by ClpAP to translocate the ClpS NTE. By tracking the bead-to-bead distance corresponding to the length between ClpS and ClpAP, this single-molecule approach should determine if ClpS is fully unfolded (results in a sharp increase in bead-to-bead distance) or undergoes more subtle conformational changes induced by ClpA.

Furthermore, if ClpS is fully unfolded, it is possible that the Pro-Pro-junction motif (Chapter 4) helps ClpS escape ClpAP proteolysis via 'back-slipping' (**Figure 6.1B**). As described by Kotamarthi et al., 2020, if ClpAP 'slips' on the ClpS polypeptide, the bead-to-bead distance would appear to increase during the process of translocation, indicating that the ClpS polypeptide was threaded back out of the ClpA channel. Alternatively, ClpS resistance to ClpAP degradation may be reflected by slower unfolding (resulting in longer unfolding dwell times) and/or translocation velocities. Slower unfolding rates, compared to those of previously characterized substrates (Aubin-Tam et al., 2011; Olivares et al., 2014, 2017; Kotamarthi et al., 2020), would indicate that ClpAP consumes more energy attempting to unfold and translocate ClpS, and consequently performs more 'non-productive' mechanical work in these futile cycles. As highly stable proteins partition between enzyme-catalyzed denaturation and substrate release (following repeated, failed unfolding attempts), longer unfolding dwell times could correspond to ClpS remodeling and/or greater mechanical stability of the ClpS adaptor that promotes its subsequent release

(discussed below). Slower translocation velocities could also result from the ClpS polypeptide sequence, as observed for other 'slippery' sequences (Barkow et al., 2009).

Following *(i)* assembly of the high-affinity delivery complex, our current model proposes subsequent reaction steps during adaptor-assisted N-end-rule substrate delivery (**Figure 3.6**, p. 114): (*ii)* ClpS remodeling, (*iii)* substrate transfer from the ClpS binding pocket to the ClpA axial channel and ClpS release, and *(iv)* N-end-rule substrate unfolding/translocation/degradation. However, the interactions between ClpA, ClpS, and the N-end-rule substrate involved in substrate-handoff are yet to be directly observed. In our model, I presume a key step during substrate transfer is the dissociation of ClpS from the ClpA axial channel, to free the ClpA pore loops to engage the N-end-rule substrate. ClpS release should be detected using real-time kinetic assays that monitor the half-life of ClpS association with ClpAP during degradation of N-end-rule substrates. As proposed by Rivera-Rivera (2015), an initial experiment could measure changes in fluorescence anisotropy due to dissociation of fluorescently labeled ClpS (ClpS*) over the course of N-end-rule substrate degradation. Briefly, ClpAP•ClpS*•N-end-rule substrate complexes would be assembled in the presence of ATPγS, followed by addition of ATP to initiate ClpS-assisted delivery and subsequent substrate degradation by ClpAP. If the rate of ternary ClpAP•ClpS*•N-end-rule substrate complex dissociation (monitored by decreases in anisotropy) is faster than the steady-state rate of ClpS-assisted N-end-rule substrate degradation, this experiment would suggest ClpS is released prior to substrate transfer and degradation (Rivera-Rivera, 2015). To track individual ClpS*-bound complexes, follow-up experiments using single-molecule fluorescence microscopy should confirm results from the bulk kinetic assay. Alternatively, additional cryo-EM studies, for example by preparing complexes similarly to those in Chapter 3 using ATP instead of ATPγS, may provide structural insights into the substrate transfer and N-end-rule substrate degradation reaction steps.

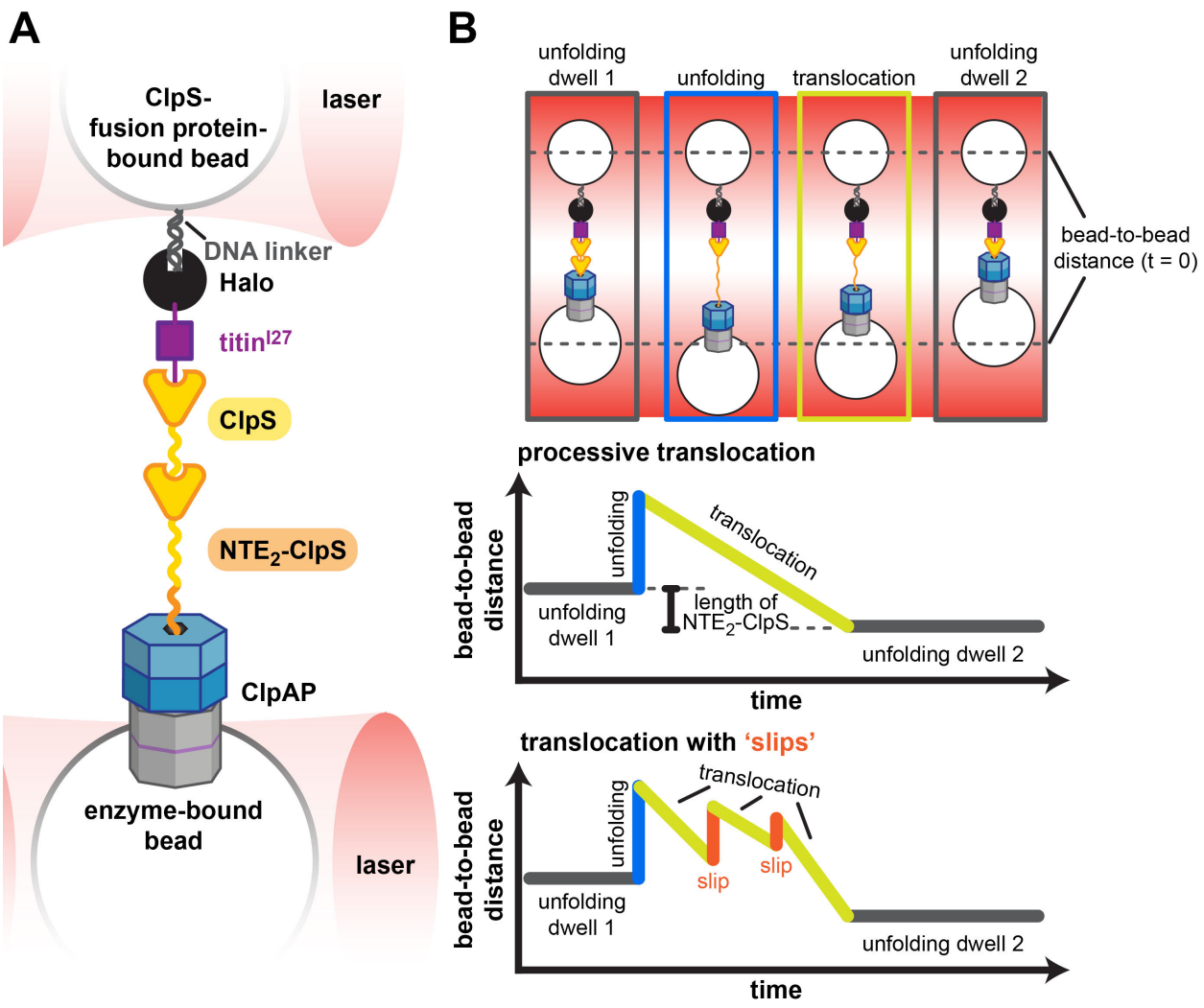


Figure 6.1 ClpS 'slipping' single-molecule optical-trapping technique. (A) Optical-trap assay design with ClpAP attached to one bead via ClpP^{platform} and a multi-domain chimera containing NTE$_2$-ClpS, ClpS, titin^{I27}, and the Halo domain attached to a second bead via a HaloTag-DNA linker, as established by previous methods (Aubin-Tam et al., 2011; Olivares et al., 2014, 2017; Kotamarthi et al., 2020). To capture single-molecule traces of bead-to-bead distance (nm between the ClpS-bound and enzyme-bound beads), the longer NTE$_2$-ClpS variant may be required to form stable tethers between the two beads. The titin^{I27} domain is an internal control that has been characterized previously by optical trapping. **(B)** Cartoon of single-molecule ClpAP traces showing processive translocation (no 'slipping', as predicted for the ClpS Ala^{24}-Ala^{25} junction variant) or ClpS 'slipping' (ClpS bearing its native Pro-Pro junction sequence is not degraded by ClpAP). The increase in bead-to-bead distance is longer for unfolding events (blue) than for 'slipping' (orange). Translocation results in decreases in the bead-to-bead distance. Little (or no) bead movement occurs during unfolding dwell times, which correlate with the mechanical stability of local protein structure, such that ClpAP continues to hydrolyze many ATPs until the next successful protein unfolding event. Adapted from Figure 2 and 6, Kotamarthi *et al.*, 2020 and Figure 4, Olivares *et al.*, 2015.

FUNCTIONS OF PORE-2 LOOPS

In the ClpA D1 ring, pore-2 loops collaborate with pore-1 loops, contacting a large surface area of the ClpS NTE (in structures of ClpAPS delivery complexes) or the substrate polypeptide (Lopez et al., 2020). The D1 pore-2 loops are critical for degradation of N-end-rule substrates but fully dispensable for N-end-rule peptide binding (Chapter 3). By contrast, the D1 pore-1 loops are essential for binding ssrA-tagged substrates and assembly into ClpAP•ClpS•N-end-rule substrate complexes (Hinnerwisch et al., 2005; Zuromski et al., 2021). Thus, from my experiments in Chapter 3, I propose that D1 pore-2 loops function predominantly in substrate unfolding/translocation, or mechanical work. To further dissect whether these pore loops affect protein unfolding or polypeptide translocation, D1 pore-2 loop variants could be characterized by kinetic assays that specifically monitor unfolding and translocation (rather than only the end result of substrate degradation) in bulk studies or by single-molecule methods, as described previously (Olivares et al., 2014, 2017; Baytshtok et al., 2015; Amor et al., 2019; Kotamarthi et al., 2020). However, pore-2 loops in either the D1 or D2 ring may have multiple functions. For example, in the ClpXP recognition complex, a pore-2 loop blocks the ssrA tag from further access to the ClpX channel, suggesting that the pore-2 loops may function as gatekeepers controlling enzyme commitment to processive unfolding/translocation (Fei et al., 2020b). Although current ClpAP structures do not show pore-2 gatekeeping, pore-2 loop variants, such as those tested in Chapter 3, could be characterized in further detail to determine whether mutating these loops results in expanded, more ‘nonspecific’ recognition by ClpAP.

THE PRO-PRO-FOLDED STRUCTURE MOTIF AND OTHER ‘ANTI-DEGRONS’

In Chapter 4, I demonstrate that the Pro-Pro dipeptide located within a few residues of a stably folded domain protects ssrA-tagged fusion proteins from ClpAP degradation. ClpAP may ‘slip’ on the Pro-Pro sequence in the native ClpS NTE–core junction, allowing ClpS to escape proteolysis. As mentioned earlier on pp. 200–203 (“Mechanism of ClpS-assisted degradation of N-end-rule

substrates"), ClpS-associated 'slipping' is likely to be detected by single-molecule optical trapping. In this technique, the sequence dependence of 'slipping' on the Pro-Pro motif could be confirmed using a similar ClpS variant bearing an Ala-Ala substitution in the Pro^{24}-Pro^{25} junction. Based on the results in Chapter 4, I predict that this ClpS Ala-Ala-junction variant would be efficiently degraded, without 'back-slips' and/or other defects in unfolding/translocation (**Figure 6.1B**, upper trace) likely to be observed with the wild-type junction sequence (lower trace). To address a potential caveat of using Ala-Ala ClpS, in which ClpAP may rapidly degrade this ClpS variant before single-molecule traces can be recorded, the ClpS variants used to compare Pro-Pro vs. Ala-Ala ClpS may need to be designed as multi-domain chimeras, in which multiple ClpS proteins are fused in tandem (**Figure 6.1A**). In this case, ClpAP may only engage the first Pro-Pro ClpS, whereas Ala-Ala ClpS multidomain variant would be processively unfolded, translocated, and degraded. To help determine how the Pro-Pro sequence protects against ClpAP proteolysis, an experimental follow-up to the ssrA-tagged proteins in Chapter 4 could involve permutation to all other amino acid dipeptides, which may reveal shared characteristics amongst protective sequence.

The most surprising result highlighted in Chapter 4 is the ability of Pro-Pro followed by a stable folded domain to effectively block ssrA-tagged degradation by ClpAP and ClpXP. The ssrA tag is a well-studied, conserved degron that targets ssrA-tagged proteins for efficient, processive degradation by multiple AAA+ proteases, including ClpAP and ClpXP. Therefore, I refer to the Pro-Pro-folded structure motif as an 'anti-degron' due to antagonization of degradation conferred by the ssrA degron. These results, in context with the natural resistance of ClpS to degradation by ClpAP, thus raise the question: what other 'anti-degrons' exist, in addition to the ClpS junction?

Although a purely bioinformatic approach to search for other proline-rich proteins may not be informative, a targeted search among ClpS orthologs across prokaryotes could illuminate patterns

of residues located at other ClpS junctions. In some bacteria, for instance, rather than two or three tandem prolines, only a single proline is present at the predicted junction sequence (Hou et al., 2008; Román-Hernández et al., 2011). In these single-proline ClpS proteins, additional residues flanking the proline may be protective against ClpAP proteolysis, or conversely, may not prevent ClpAP degradation of ClpS. This latter scenario would imply that ClpS resistance to its cognate protease is not evolutionarily conserved, indicating that these naturally degradable ClpS orthologs may behave like MecA, an adaptor that is degraded by its partner AAA+ protease, ClpCP, in *Bacillus subtilis* (Turgay et al., 1998; Schlothauer et al., 2003; Mei et al., 2009). Furthermore, a library screening approach to systematically mutate a stretch of four to five residues (to mimic the approximate length of the ClpS junction sequence) or even more residues followed by the ssrA tag could be used for comprehensive identification of other 'anti-degron' motifs adjoining a stable folded structure. Characterization of 'anti-degron' sequences for distinct AAA+ proteases will provide insight into the mechanisms controlling partial proteolysis, as some substrates, such as DnaX, are not completely degraded, resulting in selective degradation of specific domains and subsequent release of intact folded domains (Vass and Chien, 2013).

REDOX-SENSITIVE DEGRADATION BY AAA+ PROTEASES

As alluded to earlier in the introduction (Chapter 1), conformational access to degrons (pp. 35–36), e-tags (p. 38), and even adaptor-binding sites (pp. 41–42) can determine whether or not a substrate is targeted for degradation by a AAA+ protease. Consequently, substrate recognition by AAA+ proteases is often conformationally selective and regulated by environmental changes, for example, via exposure to reactive oxygen species (ROS). In the bacterial transcriptional regulator FNR, exposure of its redox-sensitive [4Fe-4S] cluster to oxygen controls both accessibility of ClpXP protease recognition signals and FNR's site-specific DNA binding activity (Chapter 5). Because the e-tag and pore-binding tag are hidden in the cluster-bound *holo*FNR dimer, ClpXP selectively degrades the *apo* FNR monomeric form that is no longer bound to the

[4Fe-4S] cluster. Lon degradation of the hydrogen peroxide sensor PerR is also controlled by oxidation-sensitive conformational changes (Ahn and Baker, 2016). In response to rising peroxide levels, 2-oxo-histidine adducts form at His^{37} and His^{91}, leading PerR to release its Fe^{2+} ligand and to adopt a protease-accessible 'open' conformation (Lee and Helmann, 2006; Jacquamet et al., 2009; Traoré et al., 2009).

The role of AAA+ proteases in oxygen-sensing may be widely conserved throughout all domains of life. Under normoxic conditions (generally 21% pO2, 160 mmHg at sea level), the eukaryotic proteasome degrades HIF-1α (Hypoxia-Inducible Factor-1), the α subunit of a global regulator of oxygen homeostasis found in all metazoans (Semenza, 2004). Unlike the metal-binding transcriptional regulators, FNR and PerR, discussed above, HIF-1α is not directly regulated by oxygen exposure. Instead oxygen levels modulate post-translational modification of HIF-1α by prolyl hydroxylases, which use oxygen as a substrate to generate 4-hydroxyproline at Pro^{402} and/or Pro^{564} in HIF-1α. Hydroxylated HIF-1α is selectively recognized by VHL (von Hippel-Lindau tumor suppressor protein), a subunit of an E3 ubiquitin-protein ligase complex, for destruction by the 26S proteasome (Ohh et al., 2000; Marxsen et al., 2004; Semenza, 2004).

Conversely, ubiquitin-dependent degradation of transcriptional regulators can also be negatively regulated under oxygen-replete conditions. In response to oxidative stress (excess ROS levels), inactivation of proteasomal degradation stabilizes protein levels of NRF2 (Nuclear factor Erythroid 2-related Factor 2), allowing this transcription factor to drive expression of antioxidant signaling pathways (Suzuki and Yamamoto, 2017). In unstressed cells, free NRF2 is constitutively targeted for degradation by the E3 $CUL3^{KEAP1}$ ligase. The KEAP1 subunit contains redox-sensitive Cys residues that prevent its assembly into the functional ubiquitin ligase complex when oxidized. Thus, oxidative stress directly controls NRF2 ubiquitination and proteasomal degradation (McMahon et al., 2003; Stewart et al., 2003).

Furthermore, a recent study demonstrated reductive stress, in addition to oxidative stress, also regulates proteasomal degradation (Manford et al., 2020). The reductive stress response is caused by prolonged antioxidant signaling in response to ROS depletion below normal physiological levels. FNIP1 (Folliculin-interacting Protein 1), associates with Folliculin to help regulate mTORC1 signaling and mitochondrial biogenesis (Baba et al., 2006; Hasumi et al., 2012). Importantly, FNIP1 is a substrate of the E3 ligase $CUL2^{FEM1B}$, and its degradation counteracts reductive stress by activating mitochondria (thereby increasing ROS production) to restore redox homeostasis (Manford et al., 2020). Because reduced Cys residues in FNIP1 (present under reductive stress) are required for $CUL2^{FEM1B}$ binding, only reduced FNIP1 is targeted for proteasomal degradation.

In summary, redox-dependent mechanisms regulating degradation by AAA+ proteases are as complex as the multiple strategies used for substrate recognition. Similar to the eukaryotic substrates and ubiquitin ligases mentioned above, bacteria likely feature redox-dependent recognition beyond direct modulation of the substrate itself, as it is plausible that adaptors and/or substrate binding partners could also be regulated in response to ROS levels. Perhaps prospective genetic screens for substrates, adaptors, and substrate-interacting proteins that are degraded by bacterial AAA+ proteases in a ROS-dependent manner will identify additional examples of targeted degradation of redox stress response regulators. Such proteins may be useful therapeutic targets to help mitigate antimicrobial resistance, as pathogenic microbes often alter redox stress response pathways caused by the host oxidative immune response and treatment with antimicrobial drugs (Grant and Hung, 2013; Dwyer et al., 2014; Whiteley et al., 2017).

REFERENCES

Ahn, B.-E., and Baker, T.A. (2016). Oxidization without substrate unfolding triggers proteolysis of the peroxide-sensor, PerR. Proc. Natl. Acad. Sci. U. S. A. *113*, E23-31.

Amor, A.J., Schmitz, K.R., Sello, J.K., Baker, T.A., and Sauer, R.T. (2016). Highly dynamic interactions maintain kinetic stability of the ClpXP protease during the ATP-fueled mechanical cycle. ACS Chem. Biol. *11*, 1552–1560.

Amor, A.J., Schmitz, K.R., Baker, T.A., and Sauer, R.T. (2019). Roles of the ClpX IGF loops in ClpP association, dissociation, and protein degradation. Protein Sci. *28*, 756–765.

Aubin-Tam, M.E., Olivares, A.O., Sauer, R.T., Baker, T.A., and Lang, M.J. (2011). Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. Cell *145*, 257–267.

Baba, M., Hong, S.B., Sharma, N., Warren, M.B., Nickerson, M.L., Iwamatsu, A., Esposito, D., Gillette, W.K., Hopkins, R.F., Hartley, J.L., et al. (2006). Folliculin encoded by the BHD gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling. Proc. Natl. Acad. Sci. U. S. A. *103*, 15552–15557.

Barkow, S.R., Levchenko, I., Baker, T.A., and Sauer, R.T. (2009). Polypeptide translocation by the AAA+ ClpXP protease machine. Chem. Biol. *16*, 605.

Baytshtok, V., Baker, T.A., and Sauer, R.T. (2015). Assaying the kinetics of protein denaturation catalyzed by AAA+ unfolding machines and proteases. Proc. Natl. Acad. Sci. U. S. A. *112*, 5377–5382.

Bieniossek, C., Niederhauser, B., and Baumann, U.M. (2009). The crystal structure of apo-FtsH reveals domain movements necessary for substrate unfolding and translocation. Proc. Natl. Acad. Sci. U. S. A. *106*, 21579–21584.

Dwyer, D.J., Belenky, P.A., Yang, J.H., Cody MacDonald, I., Martell, J.D., Takahashi, N., Chan, C.T.Y., Lobritz, M.A., Braff, D., Schwarz, E.G., et al. (2014). Antibiotics induce redox-related physiological alterations as part of their lethality. Proc. Natl. Acad. Sci. U. S. A. *111*, E2100–E2109.

Fei, X., Bell, T.A., Jenni, S., Stinson, B.M., Baker, T.A., Harrison, S.C., and Sauer, R.T. (2020a). Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. eLife *9*, e52774.

Fei, X., Bell, T.A., Barkow, S.R., Baker, T.A., and Sauer, R.T. (2020b). Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. eLife *9*, e61496, 1–39.

Flynn, J.M., Levchenko, I., Seidel, M., Wickner, S.H., Sauer, R.T., and Baker, T.A. (2001). Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. Proc. Natl. Acad. Sci. U. S. A. *98*, 10584–10589.

Gatsogiannis, C., Balogh, D., Merino, F., Sieber, S.A., and Raunser, S. (2019). Cryo-EM structure of the ClpXP protein degradation machinery. Nat. Struct. Mol. Biol. *26*, 946–954.

Grant, S.S., and Hung, D.T. (2013). Persistent bacterial infections, antibiotic tolerance, and the oxidative stress response. Virulence *4*, 273–283.

Hasumi, H., Baba, M., Hasumi, Y., Huang, Y., Oh, H., Hughes, R.M., Klein, M.E., Takikita, S., Nagashima, K., Schmidt, L.S., et al. (2012). Regulation of mitochondrial oxidative metabolism by tumor suppressor FLCN. J. Natl. Cancer Inst. *104*, 1750–1764.

Hinnerwisch, J., Fenton, W.A., Furtak, K.J., Farr, G.W., and Horwich, A.L. (2005). Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. Cell *121*, 1029–1041.

Hoskins, J.R., Pak, M., Maurizi, M.R., and Wickner, S. (1998). The role of the ClpA chaperone in proteolysis by ClpAP. Proc. Natl. Acad. Sci. U. S. A. *95*, 12135–12140.

Hoskins, J.R., Singh, S.K., Maurizi, M.R., and Wickner, S. (2000). Protein binding and unfolding by the chaperone ClpA and degradation by the protease ClpAP.

Hou, J.Y., Sauer, R.T., and Baker, T.A. (2008). Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. Nat. Struct. Mol. Biol. *15*, 288–294.

Jacquamet, L., Traoré, D.A.K., Ferrer, J.L., Proux, O., Testemale, D., Hazemann, J.L., Nazarenko, E., El Ghazouani, A., Caux-Thang, C., Duarte, V., et al. (2009). Structural characterization of the active form of PerR: Insights into the metal-induced activation of PerR and Fur proteins for DNA binding. Mol. Microbiol. *73*, 20–31.

Joshi, S.A., Hersch, G.L., Baker, T.A., and Sauer, R.T. (2004). Communication between ClpX and ClpP during substrate processing and degradation. Nat. Struct. Mol. Biol. *11*, 404–411.

Kenniston, J.A., Baker, T.A., and Sauer, R.T. (2005). Partitioning between unfolding and release of native domains during ClpXP degradation determines substrate selectivity and partial processing. Proc. Natl. Acad. Sci. U. S. A. *102*, 1390–1395.

Kim, S., Zuromski, K.L., Bell, T.A., Sauer, R.T., and Baker, T.A. (2020). ClpAP proteolysis does not require rotation of the ClpA unfoldase relative to ClpP. eLife *9*, e61451.

Kotamarthi, H.C., Sauer, R.T., and Baker, T.A. (2020). The non-dominant AAA+ ring in the ClpAP protease functions as an anti-stalling motor to accelerate protein unfolding and translocation. Cell Rep. *30*, 2644-2654.e3.

Lee, J.-W., and Helmann, J.D. (2006). The PerR transcription factor senses H2O2 by metal-catalysed histidine oxidation. Nature *440*, 363–367.

Lopez, K.E., Rizo, A.N., Tse, E., Lin, J.B., Scull, N.W., Thwin, A.C., Lucius, A.L., Shorter, J., and Southworth, D.R. (2020). Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis. Nat. Struct. Mol. Biol. *27*, 406–416.

Manford, A.G., Rodríguez-Pérez, F., Shih, K.Y., Shi, Z., Berdan, C.A., Choe, M., Titov, D. V., Nomura, D.K., and Rape, M. (2020). A cellular mechanism to detect and alleviate reductive stress. Cell *183*, 46-61.e21.

Martin, A., Baker, T.A., and Sauer, R.T. (2007). Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease. Mol. Cell *27*, 41–52.

Marxsen, J.H., Stengel, P., Doege, K., Heikkinen, P., Jokilehto, T., Wagner, T., Jelkmann, W., Jaakkola, P., and Metzen, E. (2004). Hypoxia-inducible factor-1 (HIF-1) promotes its degradation by induction of HIF-α-prolyl-4-hydroxylases. Biochem. J. *381*, 761–767.

McMahon, M., Itoh, K., Yamamoto, M., and Hayes, J.D. (2003). Keap1-dependent proteasomal degradation of transcription factor Nrf2 contributes to the negative regulation of antioxidant response element-driven gene expression. J. Biol. Chem. *278*, 21592–21600.

Mei, Z., Wang, F., Qi, Y., Zhou, Z., Hu, Q., Li, H., Wu, J., and Shi, Y. (2009). Molecular determinants of MecA as a degradation tag for the ClpCP protease. J. Biol. Chem. *284*, 34366–34375.

Nakane, T., Kimanius, D., Lindahl, E., and Scheres, S.H. (2018). Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. eLife *7*, e36861.

Ohh, M., Park, C.W., Ivan, M., Hoffman, M.A., Kim, T.Y., Huang, L.E., Pavletich, N., Chau, V., and Kaelin, W.G. (2000). Ubiquitination of hypoxia-inducible factor requires direct binding to the β-domain of the von Hippel - Lindau protein. Nat. Cell Biol. *2*, 423–427.

Olivares, A.O., Nager, A.R., Iosefson, O., Sauer, R.T., and Baker, T.A. (2014). Mechanochemical basis of protein degradation by a double-ring AAA+ machine. Nat. Struct. Mol. Biol. *21*, 871–875.

Olivares, A.O., Kotamarthi, H.C., Stein, B.J., Sauer, R.T., and Baker, T.A. (2017). Effect of directional pulling on mechanical protein degradation by ATP-dependent proteolytic machines. Proc. Natl. Acad. Sci. *114*, E6306–E6313.

Puchades, C., Rampello, A.J., Shin, M., Giuliano, C.J., Wiseman, R.L., Glynn, S.E., and Lander, G.C. (2017). Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. Science *358*, eaao0464.

Punjani, A., and Fleet, D.J. (2021). 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. J. Struct. Biol. *213*, 107702.

Ripstein, Z.A., Vahidi, S., Houry, W.A., Rubinstein, J.L., and Kay, L.E. (2020). A processive rotary mechanism couples substrate unfolding and proteolysis in the ClpXP degradation machinery. eLife *9*, e52158.

Rivera-Rivera, I. (2015). Mechanism of active substrate delivery by the AAA+ protease adaptor ClpS. Ph.D. thesis, Massachusetts Institute of Technology. Retrieved from https://dspace.mit.edu/handle/1721.1/101352

Rivera-Rivera, I., Román-Hernández, G., Sauer, R.T., and Baker, T.A. (2014). Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. Proc. Natl. Acad. Sci. U. S. A. *111*, E3853-9.

Román-Hernández, G., Hou, J.Y., Grant, R.A., Sauer, R.T., and Baker, T.A. (2011). The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. Mol. Cell *43*, 217–228.

Schlothauer, T., Mogk, A., Dougan, D.A., Bukau, B., and Turgay, K. (2003). MecA, an adaptor protein necessary for ClpC chaperone activity. Proc. Natl. Acad. Sci. U. S. A. *100*, 2306–2311.

Semenza, G.L. (2004). Hydroxylation of HIF-1: Oxygen sensing at the molecular level. Physiology *19*, 176–182.

Shin, M., Shin, M., Puchades, C., Puchades, C., Asmita, A., Puri, N., Adjei, E., Wiseman, R.L., Karzai, A.W., and Lander, G.C. (2020). Structural basis for distinct operational modes and protease activation in AAA+ protease Lon. Sci. Adv. *6*, eaba8404.

Shin, M., Watson, E.R., Song, A.S., Mindrebo, J.T., Novick, S.J., Griffin, P.R., Wiseman, R.L., and Lander, G.C. (2021). Structures of the human LONP1 protease reveal regulatory steps involved in protease activation. Nat. Commun. *12*, 3239.

Shorter, J., and Southworth, D.R. (2019). Spiraling in control: Structures and mechanisms of the Hsp104 disaggregase. Cold Spring Harb. Perspect. Biol. *11*.

Singh, S.K., Grimaud, R., Hoskins, J.R., Wickner, S., and Maurizi, M.R. (2000). Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. Proc. Natl. Acad. Sci. U. S. A. *97*, 8898–8903.

Stewart, D., Killeen, E., Naquin, R., Alam, S., and Alam, J. (2003). Degradation of transcription factor Nrf2 via the ubiquitin-proteasome pathway and stabilization by cadmium. J. Biol. Chem. *278*, 2396–2402.

Suzuki, T., and Yamamoto, M. (2017). Stress-sensing mechanisms and the physiological roles of the Keap1–Nrf2 system during cellular stress. J. Biol. Chem. *292*, 16817–16824.

Traoré, D.A.K., Ghazouani, A. El, Jacquamet, L., Borel, F., Ferrer, J.L., Lascoux, D., Ravanat, J.L., Jaquinod, M., Blondin, G., Caux-Thang, C., et al. (2009). Structural and functional characterization of 2-oxo-histidine in oxidized PerR protein. Nat. Chem. Biol. *5*, 53–59.

Turgay, K., Hahn, J., Burghoorn, J., and Dubnau, D. (1998). Competence in Bacillus subtilis is controlled by regulated proteolysis of a transcription factor. EMBO J. *17*, 6730–6738.

Vass, R.H., and Chien, P. (2013). Critical clamp loader processing by an essential AAA+ protease in Caulobacter crescentus. Proc. Natl. Acad. Sci. U. S. A. *110*, 18138–18143.

Wang, K.H., Sauer, R.T., and Baker, T.A. (2007). ClpS modulates but is not essential for bacterial N-end rule degradation. Genes Dev. *21*, 403–408.

Whiteley, A.T., Ruhland, B.R., Edrozo, M.B., and Reniere, M.L. (2017). A redox-responsive transcription factor is critical for pathogenesis and aerobic growth of *Listeria monocytogenes*. Infect. Immun. *85*.

Zhong, E.D., Bepler, T., Berger, B., and Davis, J.H. (2021). CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. Nat. Methods *18*, 176–185.

Zuromski, K.L., Kim, S., Sauer, R.T., and Baker, T.A. (2021). Division of labor between the pore-1 loops of the D1 and D2 AAA+ rings coordinates substrate selectivity of the ClpAP protease. J. Biol. Chem. *297*, 101407.

APPENDIX

Regulation of antimycin biosynthesis is controlled by the ClpXP protease

This chapter has been previously published as:
Bilyk B., Kim S., Fazal A., Baker T.A., and Seipke R.F. Regulation of Antimycin Biosynthesis Is Controlled by the ClpXP Protease. (2020). *mSphere*, *5*(2), e00144-20.
doi: 10.1128/mSphere.00144-20

This project was a collaboration with the Seipke lab (University of Leeds, UK). I designed, purified, and assayed the AntA fusion proteins.

Author contributions:

B.B., S.K., and A.F. performed experiments, interpreted data, and wrote sections of the manuscript; T.A.B. wrote sections of the manuscript; R.F.S. designed the study, performed experiments, and wrote the manuscript.

ABSTRACT

The survival of any microbe relies on its ability to respond to environmental change. Use of extracytoplasmic function (ECF) RNA polymerase sigma (σ) factors is a major strategy enabling dynamic responses to extracellular signals. *Streptomyces* species harbor a large number of ECF σ factors, nearly all of which are uncharacterized, but those that have been characterized generally regulate genes required for morphological differentiation and/or response to environmental stress, except for σ^{AntA}, which regulates starter-unit biosynthesis in the production of antimycin, an anticancer compound. Unlike a canonical ECF σ factor, whose activity is regulated by a cognate anti-σ factor, σ^{AntA} is an orphan, raising intriguing questions about how its activity may be controlled. Here, we reconstituted *in vitro* ClpXP proteolysis of σ^{AntA} but not of a variant lacking a C-terminal di-alanine motif. Furthermore, we show that the abundance of σ^{AntA} *in vivo* was enhanced by removal of the ClpXP recognition sequence and that levels of the protein rose when cellular ClpXP protease activity was abolished. These data establish direct proteolysis as an alternative and, thus far, unique control strategy for an ECF RNA polymerase σ factor and expands the paradigmatic understanding of microbial signal transduction regulation.

Keywords

ClpXP, ECF sigma factors, *Streptomyces*, antimycin, proteolysis, regulation of secondary metabolism

INTRODUCTION

The survival of any organism relies on its ability to respond to environmental change. This feature is especially true of bacteria, which often live in hostile and fluctuating environments. *Streptomyces* bacteria thrive in soils. The success of this genus of filamentous, sporulating bacteria is linked to their complex life cycle and keen ability to sense and respond to the surroundings. Notably, a multitude of bioactive secondary or specialized metabolites are produced in response to environmental cues (Zhu et al., 2014). More than half of all small-molecule therapeutics critical for human health and well-being are derived from or inspired by *Streptomyces* natural products (Newman and Cragg, 2012).

Streptomyces species typically harbor a large number of biosynthetic pathways, but only a few of them are expressed under common laboratory conditions. The biochemical diversity encoded by these silent pathways is a tremendous untapped resource for discovery of new antibacterial agents and other therapeutics. All available data indicate that the production of natural products is controlled predominantly at the level of transcription. Although there are complex regulatory cascades that tightly control expression of biosynthetic genes, they are ultimately activated, repressed, or derepressed by so-called cluster-situated regulators—regulatory proteins encoded within the biosynthetic gene cluster (BGC) (van Wezel and McDowall, 2011; van der Heul et al., 2018). Major roadblocks preventing the exploitation of silent biosynthetic pathways are a lack of insight into their regulation and limited technology for activating their expression.

Antimycins have been known for 70 years and are the founding members of a large class of natural products widely produced by *Streptomyces* species (Dunshee et al., 1949; Joynt and Seipke, 2018). Recently, antimycins were shown to be potent and selective inhibitors of the mitochondrial Bcl-2/Bcl-X_L-related antiapoptotic proteins that are overproduced by cancer cells and confer resistance to chemotherapeutic agents whose mode of action is activation of apoptosis

(Tzung et al., 2001). The ~25-kb antimycin (*ant*) BGC harbored by *Streptomyces albus* is composed of 15 genes organized into four polycistronic operons: *antBA*, *antCDE*, *antGF*, and *antHIJKLMNO* (**Figure A.1**) (Seipke et al., 2011, 2014). The regulation of the *ant* BGC is unusual compared to other secondary metabolites. Its expression is regulated by FscRI, a cluster-situated LuxR-family regulator of candicidin biosynthesis; FscRI activates expression of *antBA* and *antCDE* (McLean et al., 2016). Importantly, *antA* is a cluster-situated regulator that encodes an extracytoplasmic function (ECF) RNA polymerase σ factor (σ^{AntA}) that activates expression of the remaining operons: *antGF* and *antHIJKLMNO* (**Figure A.1**) (Seipke et al., 2014).

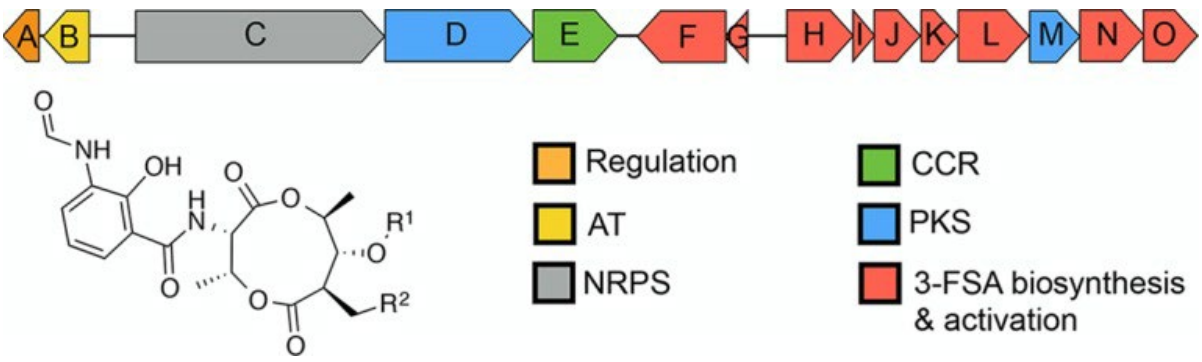


Figure A.1 Schematic representation of the antimycin (*ant*) biosynthetic gene cluster. AT, acyltransferase; NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase; CCR, crotonyl coenzyme A (crotonyl-CoA) carboxylase/reductase; 3-FSA, 3-formamidosalicylate. Antimycins: antimycin A_1, R^1= $COCH(CH_3)CH_2CH_3$, R^2= $(CH_2)_4CH_3$; antimycin A_2, R^1=$COCH(CH_3)_2$, R^2= $(CH_2)_4CH_3$; antimycin A_3, R^1= $COCH_2CH(CH_3)_2$, R^2= $(CH_2)_2CH_3$; antimycin A_4, R^1= $COCH(CH_3)_2$, R^2= $(CH_2)_2CH_3$.

σ^{AntA}, like all ECF σ factors, is similar to members of the housekeeping σ^{70} family but possesses only two of the four highly characteristic sigma domains: domains σ2 and σ4. The σ2 and σ4 regions of sigma factors bind the −10 and −35 promoter elements, respectively, and are sufficient for recruitment of RNA polymerase (Heimann, 2002). Genes encoding ECF σ factors are almost always cotranscribed with their cognate anti-σ factor (Staroń et al., 2009). Anti-σ factors are generally transmembrane proteins, but well-characterized examples of cytoplasmic anti-σ factors are known (Kang et al., 1999; Campbell et al., 2007; Rajasekar et al., 2016). Anti-σ factors

selectively bind to and sequester a partner σ factor until its release is stimulated, usually by an exogenous signal (Staroń et al., 2009; Paget, 2015). After the σ factor is released, it recruits RNA polymerase to express a defined regulon that usually includes the σ factor–anti-σ factor operon itself, which thus establishes a positive auto-feedback loop in the presence of the inducing stimulus. Even when an ECF σ factor does have a cognate anti-σ factor, an additional mechanism of control can also exist, for instance, $\sigma^{R'}$ in *S. coelicolor* is processed by the Clp-protease system (Kim et al., 2009). *Streptomyces* species encode a large number of ECF σ factors (>30 per strain), and the small number of these that have been characterized regulate genes required for morphological differentiation and/or responses to environmental stress and are not dedicated regulators of one biosynthetic pathway. Indeed, cluster-situated ECF σ factors have been observed previously only in the biosynthesis of lantibiotics produced by so-called rare actinomycetes. In *Microbispora corallina*, MibR and σ^{MibX} regulate microbisporicin biosynthesis and σ^{MibX} is controlled by the anti-σ factor MibW (Foulston and Bibb, 2010); in *Planomonospora alba*, PspR and σ^{PspX} regulate planosporicin production and σ^{PspX} is controlled by the anti-σ factor PspW (Sherwood and Bibb, 2013). Interestingly, unlike the canonical ECF σ factors σ^{MibX} and σ^{PspX}, whose activities are controlled by cognate anti-σ factors, σ^{AntA} lacks an identifiable anti-σ factor partner and as a consequence has created curiosity about how its activity is controlled.

The Clp-protease system is essential for normal bacterial proteostasis and is best characterized in *Escherichia coli* (Gur et al., 2011; Baker and Sauer, 2012). The Clp protease is a multienzyme complex composed of a barrel-shaped peptidase, ClpP, and a regulatory enzyme, either ClpA or ClpX (or ClpC in some organisms). ClpA and ClpX (and ClpC) are AAA+-family protein unfoldases that recognize an N-terminal and/or C-terminal recognition signal (degron) and utilize ATP to unfold and translocate proteins to the peptidase chamber, where they are degraded into short peptides (Olivares et al., 2016). In *Streptomyces* species, the peptidase is specified by two genes instead of one and is redundantly encoded (De Crécy-Lagard et al., 1999). The primary peptidase

is encoded by *clpP1P2*, whose corresponding proteins form a complex with ClpX or ClpA to facilitate normal proteostasis; the second peptidase is encoded by *clpP3P4*, but its expression occurs only when the primary system is compromised (Viala et al., 2000; Viala and Mazodier, 2002). The best-understood degron is the SsrA tag from *E. coli* (AANDENYALAA), which is added cotranslationally to polypeptides stalled on ribosomes (Keiler et al., 1996; Gottesman et al., 1998). The *E. coli* SsrA tag has been comprehensively studied, and Ala-Ala-COO-, at the C-terminal region of this motif, is essential for proteolysis by ClpXP (Flynn et al., 2001). Intriguingly, the C terminus of σ^{AntA} harbors the sequence Ala-Ala-COO-, which previously led us to speculate that ClpXP may modulate its level/activity (Seipke et al., 2014).

Here, we reconstituted ClpXP proteolysis of σ^{AntA} *in vitro* and showed that it is dependent upon the C-terminal Ala-Ala. We also found that the abundance of σ^{AntA} *in vivo* was higher when Ala-Ala was changed to Asp-Asp and that the abundance of σ^{AntA} was elevated in the absence of genes encoding the primary peptidase, ClpP, and its unfoldase, ClpX. These data establish direct proteolysis as an alternative, and thus far unique, control strategy of ECF RNA polymerase σ factors, expanding the paradigmatic understanding of microbial signal transduction regulation.

RESULTS AND DISCUSSION

σ^{AntA} orthologues are a new subfamily of ECF σ factors that regulate production of the antimycin biosynthetic starter unit

Since its initial discovery 6 years ago, more than 70 *ant* BGCs have been identified within the members of *Actinobacteria*, including *Actinospica*, *Saccharopolyspora*, *Streptacidiphilus*, and *Streptomyces* (Joynt and Seipke, 2018). Each of these BGCs harbors a single regulator, σ^{AntA} (53% to 100% shared amino acid identity across all orthologues), which lacks a cognate anti-σ factor partner (Seipke et al., 2014; Joynt and Seipke, 2018). Our previous work with *S. albus* S4 established that σ^{AntA} orthologues comprise a new subfamily of ECF σ factors (Seipke and

Hutchings, 2013; Seipke et al., 2014). We demonstrated that σ^{AntA} is required for expression of *antGF* and *antHIJKLMNO*, which encode a standalone ketoreductase (AntM) and proteins required for the production/activation of the starter unit, 3-formamidosalicylate (3-FSA) (**Figure A.1**). We also mapped the transcriptional start sites and identified conserved promoter sequences for these operons in all *ant* BGCs known at the time (Seipke et al., 2014). The conservation of σ^{AntA} and target promoters within *ant* BGCs from taxonomically diverse species suggests that σ^{AntA}-mediated regulation of these genes is direct. To verify this hypothesis, we performed chromatin immunoprecipitation sequencing (ChIP-seq) with a *S. albus* S4 Δ*antA* mutant complemented with a version of σ^{AntA} with a 3xFLAG tag at its N terminus, which we demonstrated restored antimycin production (see **Figure A.S2** in the supplemental material). Analysis of the resulting data revealed only one ChIP-seq peak across the whole chromosome for which the number of mapped reads was enriched for both biological replicates of Δ*antA*/3xFLAG-*antA* compared to that of the wild-type mock-immunoprecipitated control. This region corresponded to the intergenic space (297 bp) between *antG* and *antH*, which upon inspection revealed a prominent peak for the closely spaced and divergent σ^{AntA}-target promoters *antG*p and *antH*p and a second, albeit smaller peak corresponding to the 5′ end of the *antH* coding sequence (**Figure A.2**). Taken together, these data are consistent with the hypothesis that σ^{AntA} is a cluster-situated regulator that directly activates expression of genes for the production of 3-FSA during antimycin biosynthesis.

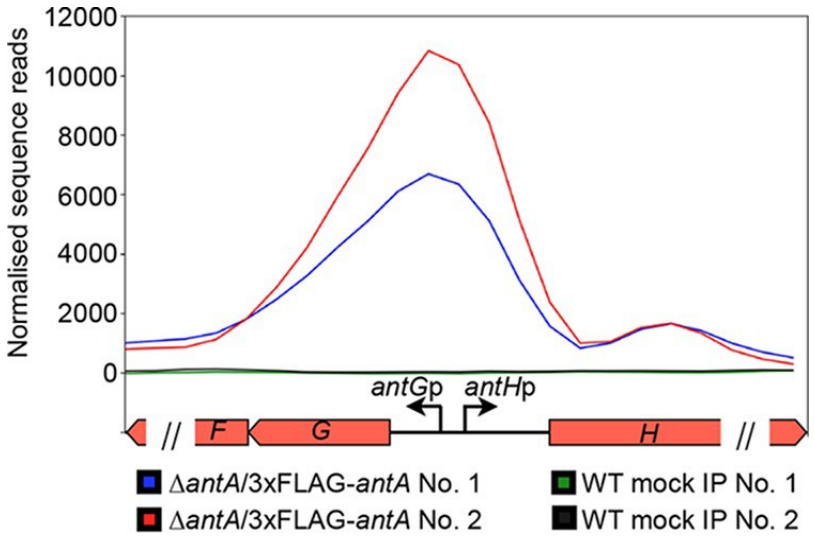


Figure A.2 3xFLAG-σ^{AntA} binds to the *antGF* and *antHIJKLMNO* promoters *in vivo*. Shown is a graphical representation of normalized sequence reads mapped to the intergenic region of *antG-antH* (shown at the bottom). The 1,242-bp genomic window depicts nucleotides 34,430 to 35,671 of contig CADY01000091.1 of the *S. albus* S4 genome. Each double slash denotes that genome window presented does not contain the entire *antF* or *antH* coding sequence. WT, wild type; IP, immunoprecipitation.

σ^{AntA} is degraded by the ClpXP protease *in vitro*

The activities of almost all characterized ECF σ factors are modulated by a cognate anti-σ factor, which is typically a protein coencoded within the same operon. Intriguingly, σ^{AntA} lacks an anti-σ factor and is therefore an orphan, indicating that a unique mechanism is likely at work to control σ^{AntA} activity. An inspection of σ^{AntA} amino acid sequences revealed a C-terminal Ala-Ala in 67 of the 71 orthologues (**Figure A.S1**). A C-terminal Ala-Ala is an important component of a common class of degrons for the ClpXP protease (Flynn et al., 2001). This observation led us to hypothesize that the activity of σ^{AntA} might be modulated by proteolysis instead of by an anti-σ factor. To test this hypothesis, we performed *in vitro* proteolysis. Previous work indicated that *S. albus* S4 σ^{AntA} was insoluble when overproduced by *E. coli*, so we pursued the overproduction and purification of the orthologue from *Streptomyces ambofaciens* ATCC 23877, which has been experimentally demonstrated to be a producer of antimycins (Schoenian et al., 2012). *S. ambofaciens* σ^{AntA} (75% amino acid identity with *S. albus* S4 σ^{AntA}, including 13 of 15 amino acid

residues at the C terminus) was purified as an N-terminal $(His)_6$-SUMO-fusion protein. The $(His)_6$-SUMO tag increases solubility and eases purification of putative substrates, without altering recognition of C-terminal degrons by ClpXP. ClpX orthologues from *E. coli* and *S. ambofaciens* possess 60% shared amino acid identity and therefore likely recognize similar substrates for degradation. Thus, ClpXP from *E. coli* was purified (**Figure A.S3**) and its ability to degrade $(His)_6$-SUMO-σ^{AntA} was assessed. Degradation of $(His)_6$-SUMO-σ^{AntA} was apparent as early as 2.5 min after addition of ATP, and all of the sample was degraded by 15 min (**Figure A.3**). Substrates of ClpXP become resistant to proteolysis by specific alterations of the C-terminal Ala-Ala (Flynn et al., 2001). Therefore, to investigate degradation specificity in the experiment described above, we constructed and tested a variant of *S. ambofaciens* σ^{AntA} in which the C-terminal Ala-Ala was mutated to Asp-Asp [$(His)_6$-SUMO-$\sigma^{AntA\text{-}DD}$]. Strikingly, the Asp-Asp variant was stable against ClpXP degradation over the lifetime of the assay (**Figure A.3**). Thus, the degradation of $(His)_6$-SUMO-σ^{AntA} and the characteristic resistance afforded by the Ala-Ala-to-Asp-Asp mutation demonstrated that σ^{AntA} is a substrate of ClpXP *in vitro*.

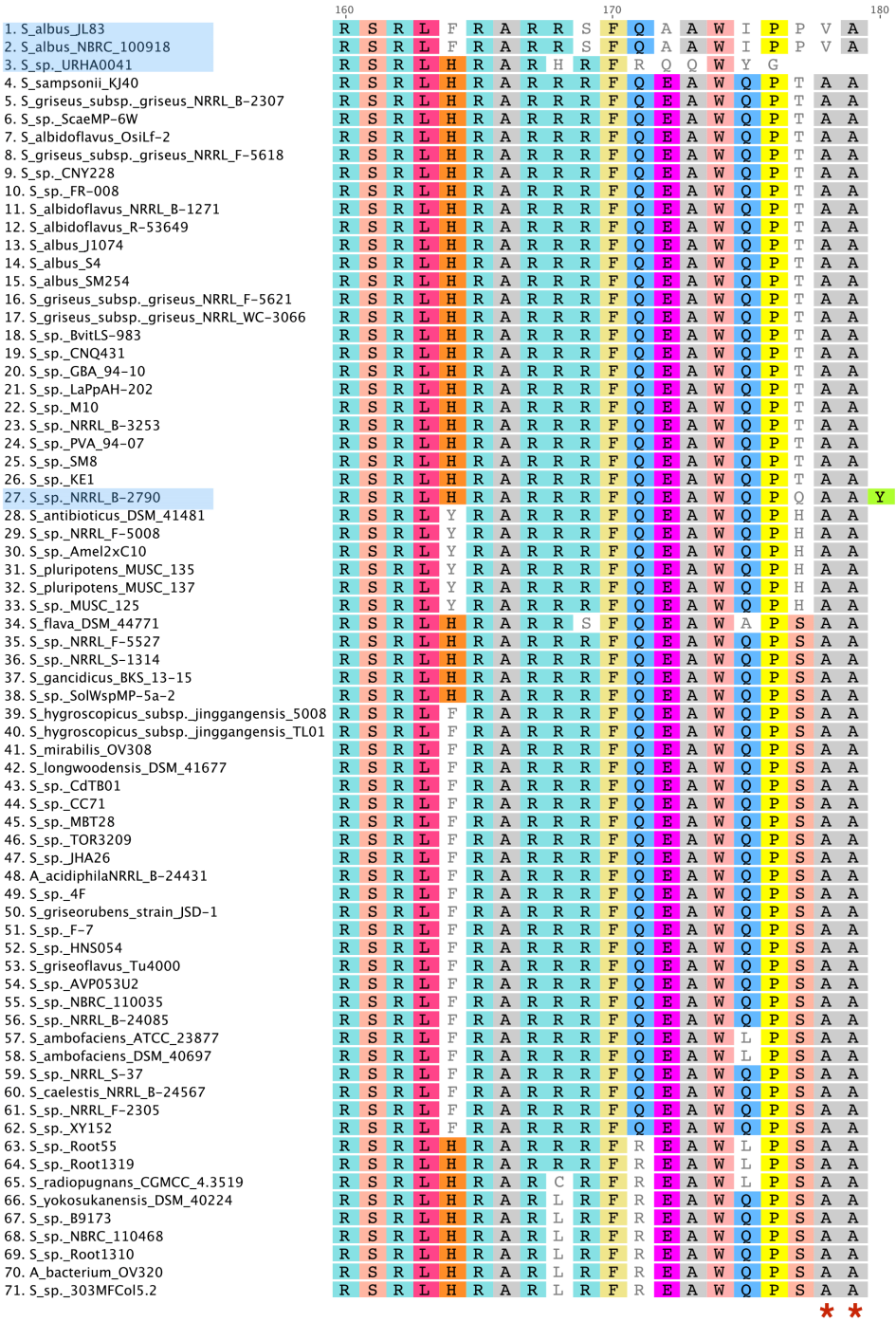


Figure A.S1 MUSCLE alignment of 71 σ^{AntA} orthologues. Only alignments of amino acid residues 160 to 180 are shown. Red asterisks indicate the C-terminal Ala-Ala motif conserved in 67 of 71 orthologues. Blue shading indicates the names of those taxa that do not possess the motif. The orthologue from *Streptomyces* sp. strain NRRL B-2790 terminated in Ala-Ala-Tyr; the inclusion of a terminal Tyr residue likely represents an artefact caused by poor genome sequence quality (>4,000 contigs). The remaining three orthologues lacking the di-alanine motif are *Streptacidiphilus albus* strains (terminating in Val-Ala) and *Streptomyces* sp. strain URHA-0041 (Tyr-Gly).

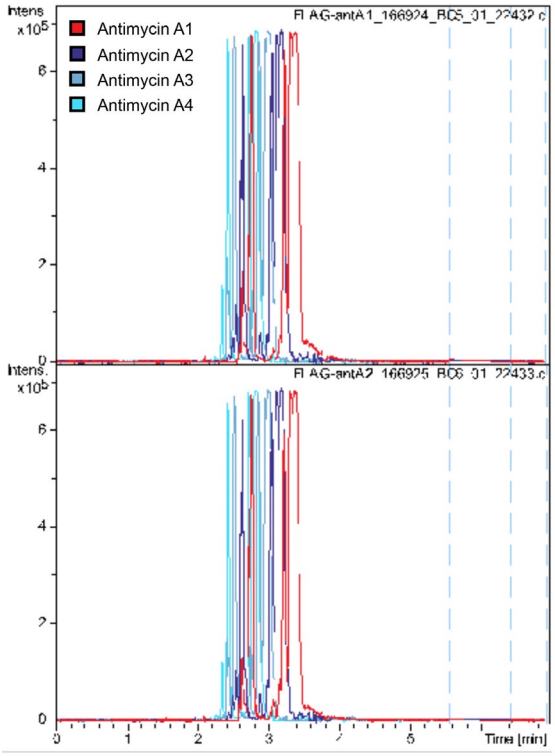


Figure A.S2 LC-HRMS analysis of Δ*antA*/pAU3-45-3xFLAG-*antA* strains. The *m*/*z* values corresponding to the $[M + H]^+$ ions derived from antimycins A1 to A4 are shown.

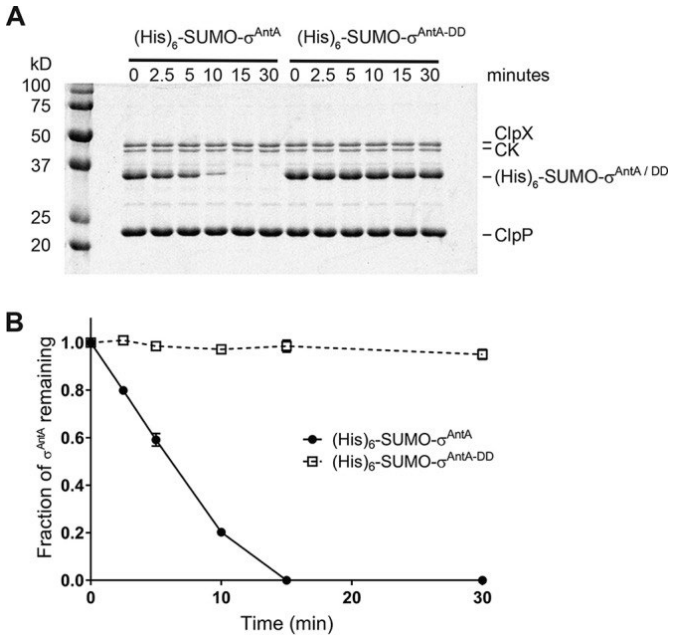


Figure A.3 Proteolysis of *S. ambofaciens* σ^{AntA} by ClpXP *in vitro*.
(A) SDS-PAGE analysis of proteolysis reaction mixtures containing 37 pmol $(His)_6$SUMO-σ^{AntA} or $(His)_6$SUMO-$\sigma^{AntA\text{-}DD}$. (B) Densitometry analysis SDS-PAGE images for three independent proteolysis experiments. The mean is plotted, and error bars illustrate the standard error of the mean (±1 SEM).

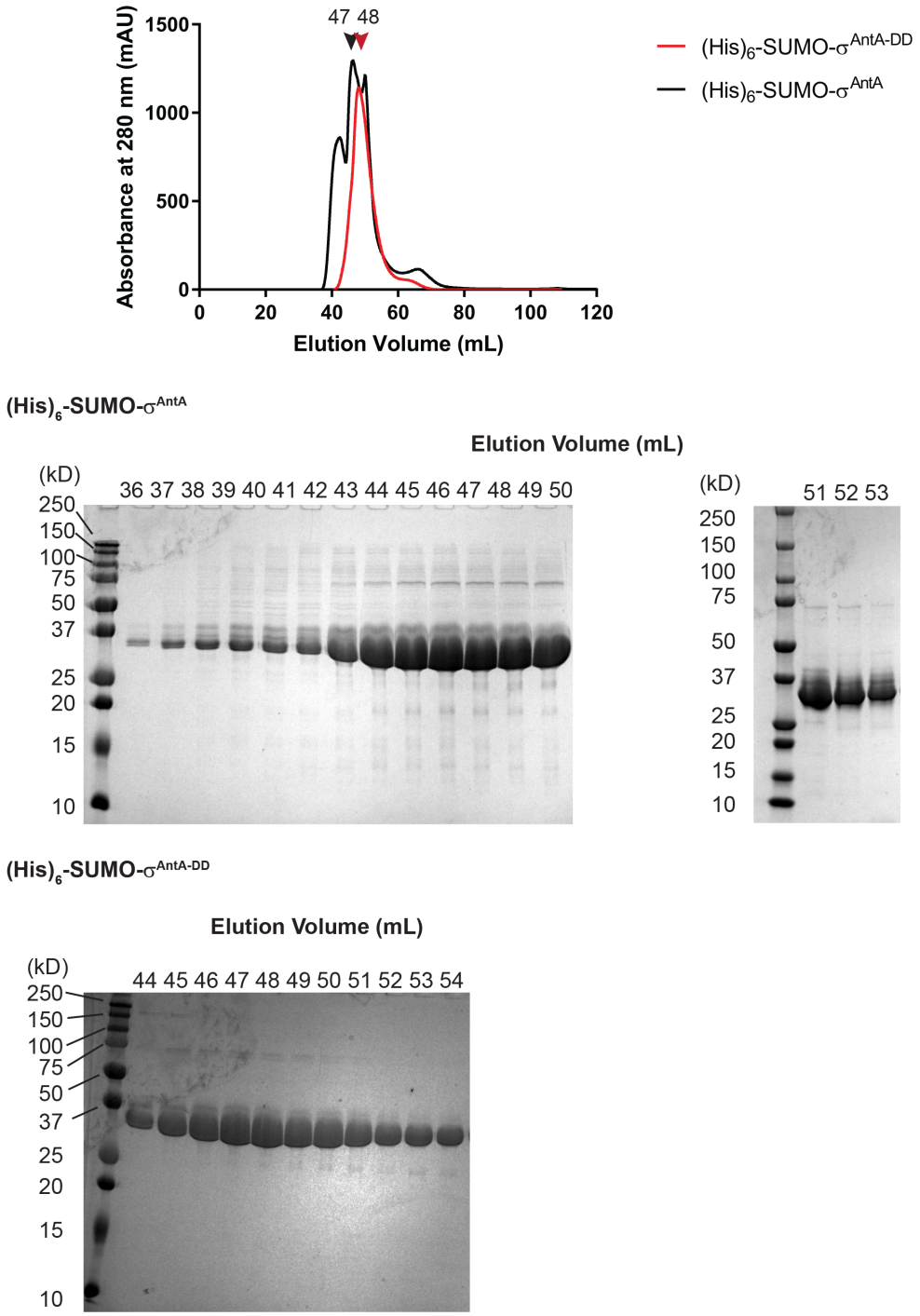


Figure A.S3 Gel filtration of purified (His)$_6$-SUMO-σ^{AntA} proteins. Gel filtration chromatograms show the elution profiles of purified (His)$_6$-SUMO-σ^{AntA} and (His)$_6$-SUMO-σ$^{AntA-DD}$ proteins (upper panel) and results of associated SDS-PAGE analyses (lower panels). The chromatogram of (His)$_6$-SUMO-σ^{AntA} differs slightly from that of (His)$_6$-SUMO-σ$^{AntA-DD}$ because gel filtration of (His)$_6$-SUMO-σ^{AntA} protein was performed in a reaction mixture containing 50 mM HEPES-KOH (pH 7.5), 150 mM KCl, 10% glycerol, and 1 mM DTT, which was later dialyzed into the storage buffer, whereas gel filtration of (His)$_6$-SUMO-σ$^{AntA-DD}$ was performed in the storage buffer.

σ^{AntA} is degraded by the ClpXP protease *in vivo*

To investigate if the *in vitro* degradation of σ^{AntA} demonstrated above is relevant to its regulation *in vivo*, we deleted the operon consisting of the *clpX*, *clpP1*, *and clpP2* genes from *S. albus* S4. The resulting Δ*clpXclpP1clpP2* mutant still harbored the second Clp peptidase encoded by *clpP3clpP4* and underwent a normal developmental cycle, albeit sporulation was less robust, which is consistent with growth characteristics reported for mutation of equivalent genes in *S. coelicolor* (**Figure A.S4**) (Viala and Mazodier, 2003). Next, genes encoding the 3xFLAG-σ^{AntA} or 3xFLAG-$\sigma^{AntA\text{-}DD}$ fusion proteins were generated and introduced into the parental strain and the *ΔclpXclpP1clpP2* mutant so the abundance of these proteins could be assessed over a developmental time course by Western blotting with anti-FLAG antisera. This experiment was initially performed with the σ^{AntA} fusions integrated on the chromosome under the control of the native promoter. However, a reliable signal could not be detected for 3xFLAG-σ^{AntA} and only a trace amount of the Asp-Asp variant was observed, presumably indicating that the cellular level of σ^{AntA} is normally low because the native promoter is relatively weak. The experiment was therefore repeated with 3xFLAG-σ^{AntA} and 3xFLAG-$\sigma^{AntA\text{-}DD}$ expression driven by a stronger, constitutive promoter, *ermE** (Luo et al., 2015). Analysis of the resulting immunoblot revealed that 3xFLAG-$\sigma^{AntA\text{-}DD}$ was more abundant than 3xFLAG-σ^{AntA} in vegetative mycelium of the parent and Δ*clpXclpP1clpP2* strains (**Figure A.4**; see also **Figure A.S5**). Strikingly, in the later stages of development after aerial mycelium had formed (24 h and 30 h), 3xFLAG-σ^{AntA} and 3xFLAG-$\sigma^{AntA\text{-}DD}$ were detected only in the Δ*clpXclpP1clpP2* strain and not the parent; the Asp-Asp variant was also present in greater relative abundance (**Figure A.4**), which was consistent with our previous experiments that showed that the *ant* BGC is downregulated at the level of transcription upon the onset of aerial growth (Seipke et al., 2014). Interestingly, the conspicuous absence of 3xFLAG-σ^{AntA} and the presence of 3xFLAG-$\sigma^{AntA\text{-}DD}$ in protein samples prepared from the latest time point sampled suggest the potential involvement of an additional degradative factor(s). Taken together,

these data support the hypothesis that the levels of σ^{AntA}, and thus its ability to activate gene expression of *antFGHIJKLMNO*, are modulated by the ClpXP protease.

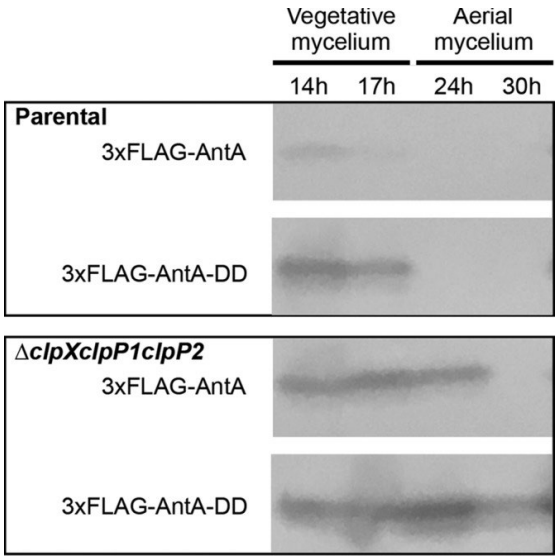


Figure A.4 The abundance of σ^{AntA} is enhanced in the absence of the ClpXP *in vivo*. Cells from the indicated strains were cultivated over a developmental time course atop cellophane discs on agar media. Protein was isolated from 100 mg of either vegetative mycelium (14 and 17 h) or aerial mycelium (24 and 30 h). Thirty-microgram volumes of total protein were analyzed by Western blotting with anti-FLAG antisera. The images shown are derived from uncropped original images shown in **Figure A.S5** in the supplemental material, and the corresponding densitometry analysis is shown in **Figure A.S6**.

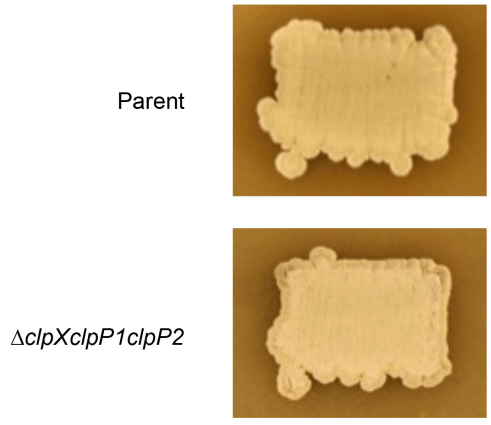


Figure A.S4 Sporulation of *S. albus* S4 *clp* mutants. Photographs were taken after 6 days of growth on the indicated medium. The Δ*clpXclpP1*Δ*clpP2* strain underwent a normal developmental cycle; however, sporulation was less robust on MS agar media than was seen with the parental strain.

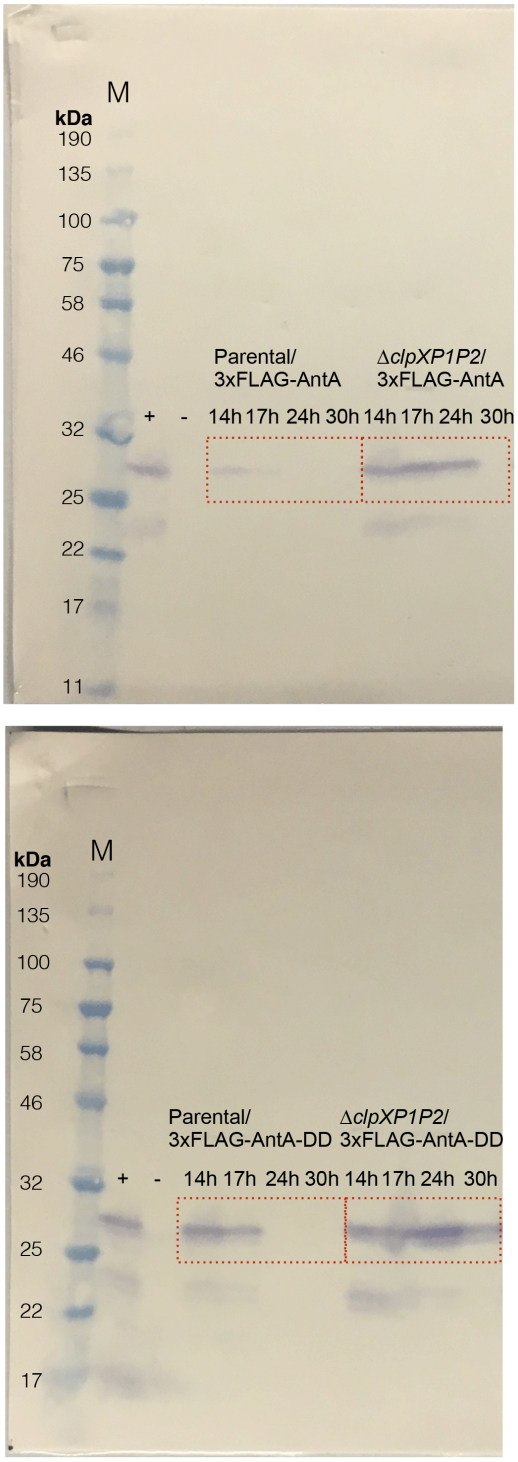


Figure A.S5 Uncropped Western blotting images. Shown are the original Western blots depicted in **Figure A.4**. Red boxes indicate cropped sections for the strains indicated. The molecular weight (Mw) of the protein marker (M) is shown. Plus and minus symbols on each blot indicate positive-testing strains (mutant strain Δ*clpXclpP1clpP2*/pPDD) and negative controls (parental strain Δantall), respectively.

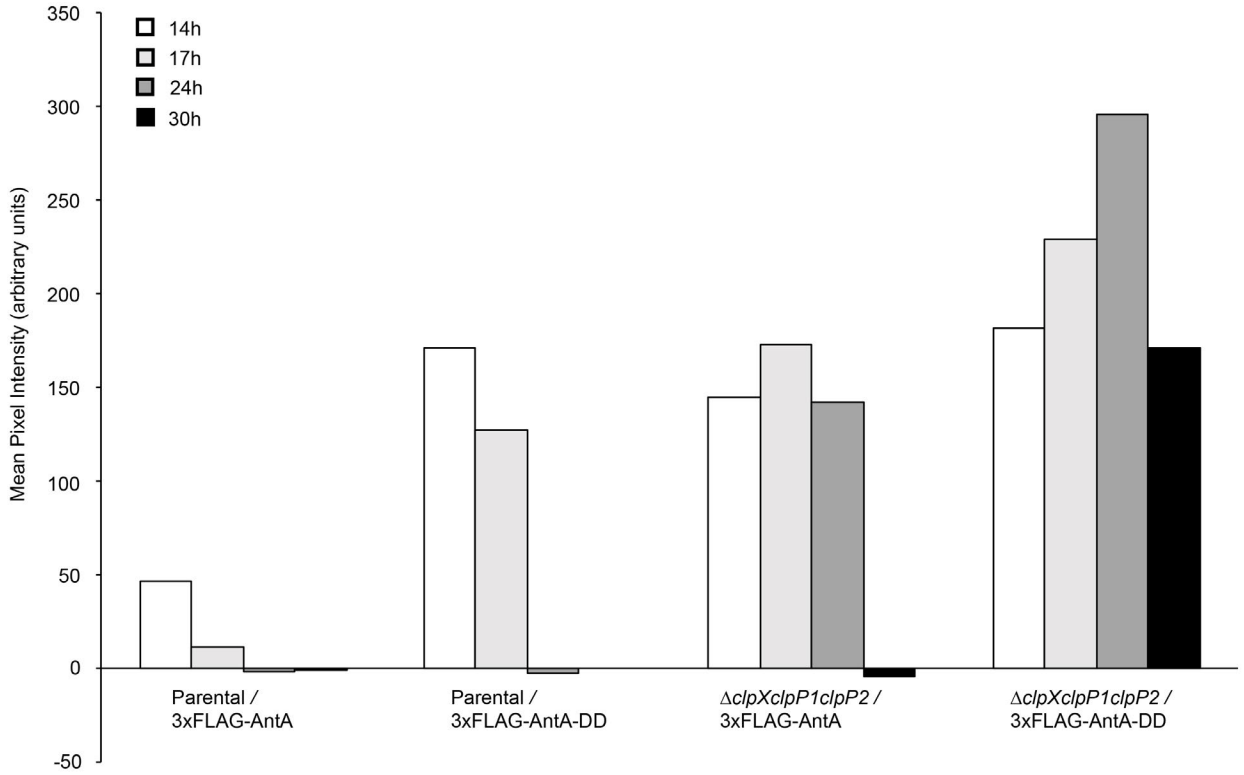


Figure A.S6 Densitometry analysis of Western blotting results. The mean intensities (in pixels) of relevant bands from the Western blots shown in **Figure A.4** are displayed for the indicated strains and time points. The mean pixel intensity was determined using Photoshop v20.0.6, and values were normalized to the intensity of the Δ*clpXclpP1clpP2*/pPDD sample contained on each membrane.

Antimycins are not overproduced in the absence of ClpXP

The results of the experiments described above indicate that the cellular level of σ^{AntA} was more abundant in the absence of the ClpXP protease. In order to determine if an increased level of this transcription factor ultimately influenced the final production titer of antimycins, we used liquid chromatography–high-resolution mass spectrometry (LC-HRMS) to assess the abundance of antimycins in chemical extracts generated from the Δ*clpXclpP1clpP2* and parental strains grown atop a cellophane disk on mannitol-soya flour (MS) agar in triplicate. The extracted ion chromatograms representing antimycin A_1, A_2, A_3, and A_4 were used to determine the peak area for each compound, which was subsequently normalized based on the wet mycelium weight of the sample. Interestingly, the results indicated that the total levels of antimycin production by the

Δ*clpXclpP1clpP2* mutant (15.57 arbitrary units [AU] ± 2.86) and the parental strain (16.59 AU ± 1.12) were not statistically significantly different (*P* value, 0.59) (see **Table A.S1** in the supplemental material). This result is consistent with a previous experiment where overexpression of *antA* did not increase the titer of antimycins, because that experiment showed that it resulted in overexpression of only *antGF* and *antHIJKLMNO* (genes encoding the production of the AntG-*S*-3-formamidosalicylate starter unit) and not the remaining genes (*antABCDE*) in the BGC (Seipke et al., 2014). This also presumably indicates that starter unit biosynthesis is not rate limiting for antimycin production.

Table A.S1 LCMS quantification of antimycin production.

	Total antimycin peak area (arbitrary units)				
***S. albus* S4 strain**	**Replicate 1**	**Replicate 2**	**Replicate 3**	**Average**	**SD***
Parental (Δantall) *attB* ΦC31 cos213	17.24	15.3	17.24	16.59	1.12
Δantall*ΔclpXclpP1clpP2 attB* ΦC31 cos213	18.71	13.11	14.88	15.57	2.86

* SD = Standard Deviation

Model for the regulation of antimycin biosynthesis

Our model for the regulation of antimycin biosynthesis is depicted in **Figure A.5**. Expression of the *ant* BGC is cross-activated by FscRI, a LuxR-family regulator, from the candicidin BGC, which activates expression of *antBA* and *antCDE* (McLean et al., 2016). This regulation in turn enables direct activation of the 3-FSA biosynthetic operons (*antGF* and *antHIJKLMNO*) by σ^{AntA}. The expression of *antBA* and *antCDE* is downregulated following the onset of morphological differentiation, presumably because the ligand sensed by the FscRI PAS domain is no longer available (Seipke et al., 2014; McLean et al., 2016). The cellular level of σ^{AntA} is antagonized by the ClpXP protease, for which it is a direct target, and is ultimately responsible for clearing residual σ^{AntA} when FscRI is inactivated following the onset of morphological differentiation (McLean et al.,

2016). While ClpXP proteolytic control of transcription factor activity, and in particular that of ECF σ factor/anti-σ factors, has been shown previously (Ades et al., 1999; Alba et al., 2002; Bellier and Mazodier, 2004; Flynn et al., 2004; Mettert and Kiley, 2005; Mika and Hengge, 2005; Bellier et al., 2006; Zellmeier et al., 2006; Mao et al., 2013), it has thus far not been directly linked to the control of cluster-situated regulators of natural product biosynthesis. This finding provides a new lens through which to examine microbial signal transduction and the regulation of natural product biosynthesis in *Streptomyces* species. Understanding the diversity of regulatory strategies controlling the expression of these pathways is critical for the development of new tools for exploiting the "silent majority" of biosynthetic pathways harbored by these organisms.

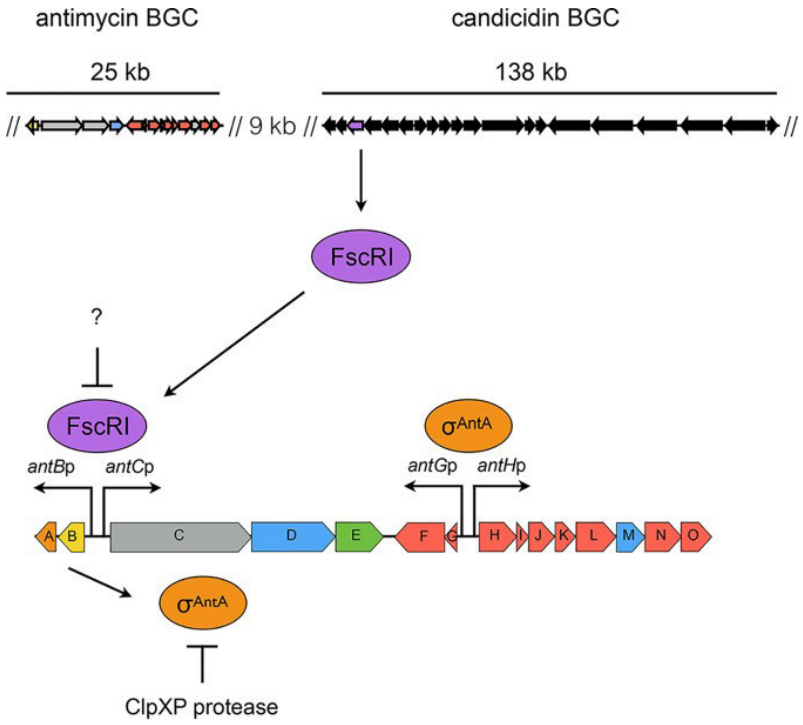


Figure A.5 Model for the regulation of antimycin biosynthesis. The upper panel displays the relative locations of the antimycin and candicidin BGCs in the *S. albus* S4 chromosome. As shown in the lower panel, FscRI, a LuxR-family regulator, from the candicidin BGC, activates expression of *antBA* and *antCDE*. This in turn enables direct activation of the 3-FSA biosynthetic operons (*antGF* and *antHIJKLMNO*) by σ^{AntA}. The cellular level of σ^{AntA} is antagonized by the ClpXP-protease system, for which it is a direct target and is ultimately responsible for clearing residual σ^{AntA} when FscRI is inactivated following the onset of differentiation.

MATERIALS AND METHODS

Table A.S2 Bacterial strains, cosmids and plasmids used in this study.

Name	Description	Reference or source
***Streptomyces* strains**		
S. albus S4	Wild type *Streptomyces albus* S4 strain	(Barke et al., 2010)
S. ambofaciens ATCC 23877	*Streptomyces ambofaciens* wild-type strain	ATCC
Δ*antA*	*S. albus* S4 harbouring an *antA* deletion; Apr^R	(Seipke et al., 2014)
Δ*antA*/pAU3-45-3xFLAG-*antA*	Δ*antA* strain harbouring pAU3-45-3xFLAG-*antA* at the ΦC31 *attB* site; Apr^R, Tsp^R	This study
Δ*antAII*	*S. albus* S4 harbouring an unmarked deletion removing the entire antimycin BGC	(Fazal et al., 2020)
Δ*clpXclpP1clpP2*	Δ*antAII* strain with an unmarked mutation deletion removing the *clpX, clpP1, clpP2* operon	This study
Escherichia coli		
XL10-Gold	General cloning host	Stratagene
ET12567	Non-methylating host for transfer of DNA into *Streptomyces* spp. (*dam, dcm, hsdM*); Cam^R	(MacNeil et al., 1992)
GB05-red	Host for RecET recombination	(Fu et al., 2012)
Rosetta(DE3)	Host for heterologous protein production	Novagene
Cosmids		
cos117	Supercos1 derivative containing the *clpP1P2* and *clpX* genes; $Carb^R$, Kan^R	This study
c117Δ*clpXP*::*aac(3)IV*	The c117-derivative with the *clpP1P2* and *clpX* replaced by the disruption cassette from paac+oriT; $Carb^R$, Kan^R, Apr^R	This study
Plasmids		
patt-saac-oriT	PCR template for *aac3(IV)* oriT cassette used in REDIRECT PCR targeting system; Apr^R, Amp^R	(Myronovskyi et al., 2014)
pAU3-45	pSET152 derivative, integrates into *φ*C31 attachment site; Apr^R Tsp^R	(Bignell et al., 2005)
pAU3-45-3xFLAG-*antA*	pAU3-45 derivative containing the *ermE**p-3xFLAG-*antA* cloned into the NotI and EcoRI sites; Apr^R, Tsp^R	This study
pET23b-His-SUMO	pET23b with a SUMO tag cloned into the NheI-AgeI sites; $Carb^R$	(Wang et al., 2007)
pET23b-His-SUMO-*antA*	pET23b-His-SUMO derivative with the wild-type *antA* gene from *Streptomyces ambofaciens* ATCC 23877 cloned into the AgeI-HindIII sites; $Carb^R$	This study
pET23b-His-SUMO-*antA-DD*	pET23b-His-SUMO-*antA* derivative harbouring with point mutations changing the AntA C-terminal AlaAla to AspAsp; $Carb^R$	This study
pPDA	pSETNFLAG derivative harboring *antA* cloned into KpnI-EcoRI sites; Apr^R	This study

pPDD	pSETNFLAG derivative harbouring *antA* encoding A172D and A173D mutations cloned into KpnI-EcoRI sites; Apr^R	This study
pSET152	*E. coli* – *Streptomyces* integrative shuttle vector, integrates into the ΦC31 attachment site; Apr^R	(Kieser et al., 2000)
pSETNFLAG	pSET152 derivative with an *ermE**p cloned into the EcoRV-EcoRI sites; and an N-terminal 3xFLAG tag and multi-cloning site cloned into the NdeI-KpnI sites; Apr^R	(McLean et al., 2016)
pUWLint31	*Streptomyces* vector for the expression of the ΦC31 integrase; pSG5 temperature sensitive *ori*; Tsp^R	(Myronovskyi et al., 2014)
pUZ8002	Encodes conjugation machinery for mobilization of plasmids from *E. coli* to *Streptomyces*; Kan^R	(MacNeil et al., 1992)

Cam – chloramphenicol; Carb – carbenicillin; Kan – kanamycin; Apr – apramycin; Spr – spectinomycin; Hyg – hygromycin, Tsp - thiostrepton

Table A.S3 Oligonucleotides used in this study.

Name	Sequence (5′–3′)*	Description
PBB001	gtggcacgcatcggtgacggc	PCR: to identify the *clpX*-containing cosmid
PBB002	tcaggccgacttctgctcgcc	
PBB003	gtgacgaatctgatgccctcc	PCR: to confirm presence of the *clpP1P2* operon on the cosmid
PBB004	tcagactccggcgttgagctt	
PBB034	agacggcccggcgccgtcgtaagacgagcaggtggatacttcc ggggatccgtcgaccc	PCR: *clpXP*/*aac3(IV)*+*oriT* recombineering cassette
PBB035	ggtggggcccttccgcgtgcgtctgccgggtgccggccctgtag gctggagctgcttcg	
PBB015	ccaaacggcgggcgcggaccc	PCR: to confirm the *clpXP* deletion
PBB018	gacggaagggccccaccgcgc	
RFS629	tatata*ggtacc*aacaccgcgcacgaactgccc	PCR: to amplify σ^{AntA}. Contains a KpnI site
RFS630	tatata*gaattc*tcaggcggcggtgggctgcc	PCR: to amplify σ^{AntA}. Contains an EcoRI site
RFS663	tatata*gaattc*tcagtcgtcggtgggctgc	PCR: to amplify σ^{AntA}, encodes A172D and A173D mutations. Contains an EcoRI site
SK221	cgc*aagctt*tcacgccgcc	PCR: to amplify σ^{AntA}. Contains a HindIII site
SK222	tat*accggt*ggttccaccgtcagcgaactcc	PCR: to amplify σ^{AntA}. Contains an AgeI site
SK232	ggcatggctgccctcggatgattga*aagcttgcggccgc*	PCR: to mutagenize σ^{AntA}. Contains HindIII and NotI sites
SK233	*gcggccgcaagctt*tcaatcatccgagggcagccatgcc	PCR: to mutagenize σ^{AntA}. Contains HindIII and NotI sites
*Restriction sites are indicated by italics and non-homologous sequences are underlined		

Growth media, strains, cosmids, plasmids, and other reagents.

Escherichia coli strains were propagated on Lennox agar (LA) or broth (LB) (Seipke et al., 2011; Skyrud et al., 2018), and *Streptomyces albus* S4 strains were cultivated using LA, LB, and mannitol-soya flour (MS) agar or broth (Skyrud et al., 2018). Development of *clp* mutants was assessed on MS and ISP2 medium (Skyrud et al., 2018). Culture medium was supplemented with antibiotics as required at the following concentrations: apramycin, 50 μg ml^{-1}; carbenicillin, 100 μg ml^{-1}; chloramphenicol, 25 μg ml^{-1}; hygromycin, 50 μg ml^{-1}; kanamycin, 50 μg ml^{-1}; nalidixic acid, 25 μg ml^{-1}. *Streptomyces* strains were constructed by conjugal mating with *E. coli* ET12567 as previously described (Skyrud et al., 2018). Enzymes were purchased from New England Biolabs unless otherwise stated, and oligonucleotides were purchased from Integrated DNA Technologies, Inc. All of the strains, cosmids, and plasmids used in this study are described in **Table A.S2** in the supplemental material, and all of the oligonucleotides used are provided in **Table A.S3**.

Construction of plasmids

The insertion for each plasmid generated in this study was prepared by PCR amplification with Q5 High-Fidelity DNA polymerase and oligonucleotides containing restriction sites. PCR-amplified insertions were restricted and cloned into the relevant plasmids cut with the same enzymes by standard molecular biology procedures. All clones were sequenced to verify the integrity of insertion DNA. The names of the restriction sites used for cloning are provided with the plasmid descriptions in Table S2.

ChIP-seq and bioinformatics analyses

The *antA* coding sequence was amplified with RFS629 and RFS630, which contain KpnI and EcoRI restriction sites, respectively. The restricted PCR product was cloned into pSETNFLAG (McLean et al., 2016) digested with the same enzymes. The resulting plasmid was then restricted

with NotI and EcoRI to release *ermE**p-3xFLAG-*antA*, which was subsequently cloned into pAU3-45 (Bignell et al., 2005) digested with the same enzymes. pAU3-45-3xFLAG-*antA* was mobilized to an apramycin-marked Δ*antA* strain (Seipke et al., 2014). Cultivations of the wild-type and Δ*antA*/pAUNFLAG-*antA* strains for ChIP-seq were performed exactly as described previously (McLean et al., 2016). Pure DNA resulting from immunoprecipitates from two biological replicates of the wild-type and Δ*antA*/pAUNFLAG-*antA* strains and nonimmunoprecipitated chromosomal DNA were sequenced with a Illumina HiSeq 3000 platform with 150-nucleotide (nt) paired-end reads by the University of Leeds Next Generation Sequencing Facility at the St. James Teaching Hospital NHS Trust. The resulting reads were analyzed exactly as described previously (McLean et al., 2016). The graphic in **Figure A.2** was generated using DeepTools computeMatrix in scale-regions mode with a bin size of 54 and plotProfile functions (Ramírez et al., 2014).

Construction of the *S. albus* S4 Δ*clpXclpP1clpP2* mutant strain

Deletion of *clpXclpP1clpP2* was carried out using RecET recombineering in E. coli as follows. The *clpXclpP1clpP2*-containing cosmid, cos117, was obtained by screening a previously constructed *S. albus* S4 Supercos1 cosmid library (Seipke et al., 2011) by PCR using oligonucleotides PBB001 and PBB002. Cos117 was mutagenized as required using E. coli recombineering with strain GB05-red (Fu et al., 2012) and a deletion cassette. The deletion cassette was generated by PCR from paac-apr-oriT (Myronovskyi et al., 2014) and consisted of the *aac(*3*)IV* apramycin resistance gene and a conjugal origin of transfer (*oriT*), which was flanked by ΦC31-*attL* and ΦC31-*attR* sites for excision of the cassette. The oligonucleotides used to generate deletion cassettes included 39 nt of homology upstream or downstream of the target open reading frame(s) and are listed in **Table A.S3**. The resulting PCR product was digested with DpnI, gel purified, and electroporated into arabinose-induced E. coli GB05-red harboring cos117. Transformants were screened for the presence of mutagenized cosmid by PCR using oligonucleotides listed in **Table A.S3**, and the integrity of the locus was verified by DNA sequencing. The mutagenized cosmid

was electroporated into E. coli ET12567/pUZ8002 and mobilized to a strain of *S. albus* S4 harboring an entire antimycin BGC deletion (Δantall) (Fazal et al., 2020) by conjugation as described previously (Skyrud et al., 2018). Transconjugants were screened for apramycin resistance and kanamycin sensitivity. The integrity of an apramycin-marked mutant was verified by PCR using the oligonucleotides listed in **Table A.S3**. The apramycin deletion cassette was subsequently excised from the chromosome by conjugal introduction of pUWLint31, which is a replicative plasmid with a temperature-sensitive origin of replication that expresses the ΦC31 integrase required for removal of the cassette (Myronovskyi et al., 2014). Transconjugants were screened for loss of apramycin resistance, and excision of the cassette was verified by polymorphic shift PCR and DNA sequencing of the product.

Immunoblot analysis

Spores of the parental strain and of *S. albus* Δantall and the Δ*clpXclpP1clpP2* mutant harboring pPDA or pPDD were grown on MS agar (buffered with 50 mM TES, pH 7.2) covered with cellophane discs. Protein was isolated from mycelium collected during growth at the following regular intervals: 14 h, 17 h, 24 h, and 30 h for the Δantall and Δ*clpXclpP1clpP2* mutants harboring 3xFLAG-AntA constructs and 17 h, 20 h, 23 h, and 30 h for the Δantall and Δ*clpXclpP1clpP2* mutants harboring the 3xFLAG-FscRI construct. Protein samples were generated as follows: 100-mg volumes of cells were resuspended in 200 µl lysis buffer (50 mM sodium phosphate buffer [pH 7.0]; 150 mM sodium chloride; 10 mg ml^{-1} lysozyme; cOmplete, Mini, EDTA-free protease inhibitors [Roche]; 100 mg of 0.1-mm-diameter glass beads [PowerLyzer]) and lysed by vortex mixing for 30 min at 2,000 rpm and 37°C, with a subsequent incubation for another 30 min at 37°C. The obtained suspension was centrifuged for 20 min at 20,000 × *g* at 18°C. Each clarified protein sample (30 µg) was subjected to SDS-PAGE and then transferred to a nitrocellulose membrane (pore size, 0.2 µm) for Western blot analysis. The membrane was probed with mouse monoclonal

anti-FLAG M2-horseradish peroxidase (HRP) antibody (Sigma) (1:10,000, and the signals were detected using Pierce 1-Step Ultra TMB blotting solution (Thermo Scientific).

Protein purification and *in vitro* ClpXP proteolysis assays.

The wild-type *antA* gene was PCR amplified and cloned into the AgeI and HindIII sites of the pET23b-SUMO vector, which harbors an N-terminal $(His)_6$-SUMO tag (Wang et al., 2007). The plasmid for production of $(His)_6$-SUMO-$\sigma^{AntA\text{-}DD}$ was generated by site-directed mutagenesis (Agilent QuikChange) using primers listed in **Table A.S3**. $(His)_6$-SUMO-σ^{AntA} and $(His)_6$-SUMO-$\sigma^{AntA\text{-}DD}$ were produced by E. coli Rosetta(DE3) (Novagen) grown in LB at 37°C until an optical density at 600 nm (OD_{600}) of 0.5 was reached, followed by induction with 0.4 mM IPTG (isopropyl-β-d-thiogalactopyranoside) and growth at 18°C for 16 h. Cells were resuspended in 50 mM sodium phosphate (pH 8)–1 M NaCl–20 mM imidazole–10% glycerol–1 mM dithiothreitol (DTT) and lysed by the use of a French press at 28,000 lb/in^2, followed by treatment with protease inhibitor cocktail set III (EDTA-free) (Calbiochem), and benzonase (Millipore Sigma). $(His)_6$-SUMO-σ^{AntA} and $(His)_6$-SUMO-$\sigma^{AntA\text{-}DD}$ proteins were purified by nickel-nitrilotriacetic acid (Ni-NTA) affinity chromatography and Superdex-75 gel filtration and stored in 50 mM potassium phosphate (pH 6.8)–850 mM KCl–10% glycerol–1 mM DTT. E. coli ClpX and ClpP proteins were purified as described previously (Kim et al., 2000; Neher et al., 2003).

In vitro ClpXP proteolysis assays were performed at 30°C by preincubating 0.3 μM $ClpX_6$ and 0.8 μM $ClpP_{14}$ with an ATP regeneration system (4 mM ATP, 50 μg ml^{-1} creatine kinase, 5 mM creatine phosphate) in 25 mM HEPES-KOH (pH 7.5)–20 mM KCl–5 mM $MgCl_2$–10% glycerol–0.032% NP-40–0.2 mM DTT and adding substrate to initiate the reactions. Samples of each reaction mixture were taken at specific time points, and the reactions were stopped by addition of SDS-PAGE loading dye and boiling at 100°C before loading on Tris-Glycine-SDS gels. Bands were visualized by staining with colloidal Coomassie G-250 and quantified by ImageQuant (GE

Healthcare) after scanning by Typhoon FLA 9500 (GE Healthcare). The $(His)_6$-SUMO-σ^{AntA} fraction remaining was calculated by dividing the $(His)_6$-SUMO-σ^{AntA} density at a given time point by the density at time zero and normalizing by ClpX density.

Chemical analysis

S. albus S4 strains were cultivated atop a cellophane disc on MS agar at 30°C for 7 days in triplicate. At the time of harvest, the cellophane disc containing mycelium was removed and the quantity of biomass was determined. Bacterial metabolites were extracted from both the mycelium and the "spent" agar for 1 h using 50 ml of ethyl acetate. Thirty-milliliter volumes of ethyl acetate were evaporated to dryness under conditions of reduced pressure, and the resulting residue was resuspended in 100% methanol (300 µl). Immediately prior to LC-HRMS analysis, methanolic extracts were centrifuged at 16,000 × *g* in a microcentrifuge tube for 5 min to remove insoluble material. Only the supernatant (3 µl) was injected into a Bruker Maxis Impact time of flight (TOF) mass spectrometer equipped with a Dionext Ultimate 3000 HPLC system as previously described (Liu et al., 2015). The peak area associated with the extracted ion chromatograms for antimycin A_1, A_2, A_3, and A_4 present in agar and mycelium extracts was determined and used to calculate the total level of antimycins produced for each replicate. These values were subsequently used to determine the arithmetic mean for total antimycin production for each strain. Statistical significance was assessed in Microsoft Excel by a homoscedastic Student's *t* test with a two-tailed distribution.

Data availability

The next-generation-sequencing data obtained in this study are available under ArrayExpress accession numbers E-MTAB-7700 and E-MTAB-5122.

ACKNOWLEDGMENTS

We thank Matt Hutchings, Paul Hoskisson, Kenneth McDowall, and Alex O'Neill for helpful discussion concerning the manuscript.

B.B., S.K., and A.F. performed experiments, interpreted data, and wrote sections of the manuscript; T.A.B. wrote sections of the manuscript; R.F.S. designed the study, performed experiments, and wrote the manuscript.

B.B. was supported by a grant from the Biotechnology and Biological Sciences Research Council (BB/N007980/1) awarded to R.F.S. A.F. was funded by a Ph.D. studentship funded by the University of Leeds. S.K. was supported by a National Science Foundation Graduate Research Fellowship and the Howard Hughes Medical Foundation; T.A.B. is an employee of the Howard Hughes Medical Foundation.

REFERENCES

Ades, S.E., Connolly, L.E., Alba, B.M., and Gross, C.A. (1999). The *Escherichia coli* σ(E)-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-σ factor. Genes Dev. *13*, 18, 2449–2461.

Alba, B.M., Leeds, J.A., Onufryk, C., Lu, C.Z., and Gross, C.A. (2002). DegS and YaeL participate sequentially in the cleavage of RseA to activate the σE-dependent extracytoplasmic stress response. Genes Dev. *16*, 16, 2156–2168.

Baker, T.A., and Sauer, R.T. (2012). ClpXP, an ATP-powered unfolding and protein-degradation machine. Biochim. Biophys. Acta - Mol. Cell Res. *1823*, 1, 15–28.

Barke, J., Seipke, R.F., Grüschow, S., Heavens, D., Drou, N., Bibb, M.J., Goss, R.J.M., Yu, D.W., and Hutchings, M.I. (2010). A mixed community of actinomycetes produce multiple antibiotics for the fungus farming ant *Acromyrmex octospinosus*. BMC Biol. *8*, 109.

Bellier, A., and Mazodier, P. (2004). ClgR, a novel regulator of *clp* and *lon* expression in *Streptomyces*. J. Bacteriol. *186*, 10, 3238–3248.

Bellier, A., Gominet, M., and Mazodier, P. (2006). Post-translational control of the *Streptomyces lividans* ClgR regulon by ClpP. Microbiology *152*, 4, 1021–1027.

Bignell, D.R.D., Tahlan, K., Colvin, K.R., Jensen, S.E., and Leskiw, B.K. (2005). Expression of *ccaR*, encoding the positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus*, is dependent on *bldG*. Antimicrob. Agents Chemother. *49*, 4, 1529–1541.

Campbell, E.A., Greenwell, R., Anthony, J.R., Wang, S., Lim, L., Das, K., Sofia, H.J., Donohue, T.J., and Darst, S.A. (2007). A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. Mol. Cell *27*, 5, 793–805.

De Crécy-Lagard, V., Servant-Moisson, P., Viala, J., Grandvalet, C., and Mazodier, P. (1999). Alteration of the synthesis of the Clp ATP-dependent protease affects morphological and physiological differentiation in Streptomyces. Mol. Microbiol. *32*, 3, 505–517.

Dunshee, B.R., Leben, C., Keitt, G.W., and Strong, F.M. (1949). The isolation and properties of antimycin A. J. Am. Chem. Soc. *71*, 7, 2436–2437.

Fazal, A., Thankachan, D., Harris, E., and Seipke, R.F. (2020). A chromatogram-simplified *Streptomyces albus* host for heterologous production of natural products. Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol. *113*, 4, 511–520.

Flynn, J.M., Levchenko, I., Seidel, M., Wickner, S.H., Sauer, R.T., and Baker, T.A. (2001). Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. Proc. Natl. Acad. Sci. *98*, 19, 10584–10589.

Flynn, J.M., Levchenko, I., Sauer, R.T., and Baker, T.A. (2004). Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. Genes Dev. *18*, 18, 2292–2301.

Foulston, L.C., and Bibb, M.J. (2010). Microbisporicin gene cluster reveals unusual features of lantibiotic biosynthesis in actinomycetes. Proc. Natl. Acad. Sci. *107*, 30, 13461–13466.

Fu, J., Bian, X., Hu, S., Wang, H., Huang, F., Seibert, P.M., Plaza, A., Xia, L., Müller, R., Stewart, A.F., et al. (2012). Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. Nat. Biotechnol. *30*, 5, 440–446.

Gottesman, S., Roche, E., Zhou, Y., and Sauer, R.T. (1998). The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. Genes Dev. *12*, 9, 1338–1347.

Gur, E., Biran, D., and Ron, E.Z. (2011). Regulated proteolysis in Gram-negative bacteria — how and when? Nat. Rev. Microbiol. *9*, 12, 839–848.

Heimann, J.D. (2002). The extracytoplasmic function (ECF) sigma factors. Adv. Microb. Physiol. *46*, 47–110.

van der Heul, H.U., Bilyk, B.L., McDowall, K.J., Seipke, R.F., and van Wezel, G.P. (2018). Regulation of antibiotic production in *Actinobacteria*: new perspectives from the post-genomic era. Nat. Prod. Rep. *35*, 6, 575–604.

Joynt, R., and Seipke, R.F. (2018). A phylogenetic and evolutionary analysis of antimycin biosynthesis. Microbiology *164*, 1, 28–39.

Kang, J.G., Paget, M.S.B., Seok, Y.J., Hahn, M.Y., Bae, J.B., Hahn, J.S., Kleanthous, C., Buttner, M.J., and Roe, J.H. (1999). RsrA, an anti-sigma factor regulated by redox change. EMBO J. *18*, 15, 4292–4298.

Keiler, K.C., Waller, P.R.H., and Sauer, R.T. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. Science. *271*, 5251, 990–993.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., Hopwood, D.A., and others (2000). Practical streptomyces genetics (John Innes Foundation Norwich).

Kim, M.-S., Hahn, M.-Y., Cho, Y., Cho, S.-N., and Roe, J.-H. (2009). Positive and negative feedback regulatory loops of thiol-oxidative stress response mediated by an unstable isoform of σ^R in actinomycetes. Mol. Microbiol. *73*, 5, 815–825.

Kim, Y.I., Burton, R.E., Burton, B.M., Sauer, R.T., and Baker, T.A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. Mol. Cell *5*, 4, 639–648.

Liu, J., Zhu, X., Seipke, R.F., and Zhang, W. (2015). Biosynthesis of antimycins with a reconstituted 3-formamidosalicylate pharmacophore in *Escherichia coli.* ACS Synth. Biol. *4*, 5, 559–565.

Luo, Y., Zhang, L., Barton, K.W., and Zhao, H. (2015). Systematic identification of a panel of strong constitutive promoters from *Streptomyces albus*. ACS Synth. Biol. *4*, 9, 1001–1010.

MacNeil, D.J., Gewain, K.M., Ruby, C.L., Dezeny, G., Gibbons, P.H., and MacNeil, T. (1992). Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. Gene *111*, 1, 61–68.

Mao, X.M., Sun, N., Wang, F., Luo, S., Zhou, Z., Feng, W.H., Huang, F.L., and Li, Y.Q. (2013). Dual positive feedback regulation of protein degradation of an extra-cytoplasmic function σ factor for cell differentiation in *Streptomyces coelicolor*. J. Biol. Chem. *288*, 43, 31217–31228.

McLean, T.C., Hoskisson, P.A., and Seipke, R.F. (2016). Coordinate regulation of antimycin and candicidin biosynthesis. MSphere *1*, 6.

Mettert, E.L., and Kiley, P.J. (2005). ClpXP-dependent Proteolysis of FNR upon Loss of its O_2-sensing [4Fe–4S] Cluster. J. Mol. Biol. *354*, 2, 220–232.

Mika, F., and Hengge, R. (2005). A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of σS (RpoS) in *E. coli*. Genes Dev. *19*, 22, 2770–2781.

Myronovskyi, M., Rosenkränzer, B., and Luzhetskyy, A. (2014). Iterative marker excision system. Appl. Microbiol. Biotechnol. *98*, 10, 4557–4570.

Neher, S.B., Sauer, R.T., and Baker, T.A. (2003). Distinct peptide signals in the UmuD and UmuD′ subunits of UmuD/D′ mediate tethering and substrate processing by the ClpXP protease. Proc. Natl. Acad. Sci. U. S. A. *100*, 23, 13219–13224.

Newman, D.J., and Cragg, G.M. (2012). Natural products as sources of new drugs over the 30 years from 1981 to 2010. J. Nat. Prod. *75*, 3, 311–335.

Olivares, A.O., Baker, T.A., and Sauer, R.T. (2016). Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. Nat. Rev. Microbiol. *14*, 1, 33–44.

Paget, M.S. (2015). Bacterial sigma factors and anti-sigma factors: Structure, function and distribution. Biomolecules *5*, 3, 1245–1265.

Rajasekar, K. V., Zdanowski, K., Yan, J., Hopper, J.T.S., Francis, M.L.R., Seepersad, C., Sharp, C., Pecqueur, L., Werner, J.M., Robinson, C. V., et al. (2016). The anti-sigma factor RsrA responds to oxidative stress by reburying its hydrophobic core. Nat. Commun. *7*, 1, 12194.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). DeepTools: A flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W1, W187-91.

Schoenian, I., Paetz, C., Dickschat, J.S., Aigle, B., Leblond, P., and Spiteller, D. (2012). An unprecedented 1,2-shift in the biosynthesis of the 3-aminosalicylate moiety of antimycins. ChemBioChem *13*, 6, 769–773.

Seipke, R.F., and Hutchings, M.I. (2013). The regulation and biosynthesis of antimycins. Beilstein J. Org. Chem. *9*, 2556–2563.

Seipke, R.F., Barke, J., Brearley, C., Hill, L., Yu, D.W., Goss, R.J.M., and Hutchings, M.I. (2011). A single *Streptomyces* symbiont makes multiple antifungals to support the fungus farming ant *Acromyrmex octospinosus*. PLoS One *6*, 8, e22028.

Seipke, R.F., Patrick, E., and Hutchings, M.I. (2014). Regulation of antimycin biosynthesis by the orphan ECF RNA polymerase sigma factor σAntA. PeerJ *2014*, 1, e253.

Sherwood, E.J., and Bibb, M.J. (2013). The antibiotic planosporicin coordinates its own production in the actinomycete *Planomonospora alba*. Proc. Natl. Acad. Sci. U. S. A. *110*, 27.

Skyrud, W., Liu, J., Thankachan, D., Cabrera, M., Seipke, R.F., and Zhang, W. (2018). Biosynthesis of the 15-membered ring depsipeptide neoantimycin. ACS Chem. Biol. *13*, 5, 1398–1406.

Staroń, A., Sofia, H.J., Dietrich, S., Ulrich, L.E., Liesegang, H., and Mascher, T. (2009). The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) σ factor protein family. Mol. Microbiol. *74*, 3, 557–581.

Tzung, S.-P., Kim, K.M., Basañez, G., Giedt, C.D., Simon, J., Zimmerberg, J., Zhang, K.Y.J., and Hockenbery, D.M. (2001). Antimycin A mimics a cell-death-inducing Bcl-2 homology domain 3. Nat. Cell Biol. *3*, 2, 183–191.

Viala, J., and Mazodier, P. (2002). ClpP-dependent degradation of PopR allows tightly regulated expression of the *clpP3 clpP4* operon in *Streptomyces lividans*. Mol. Microbiol. *44*, 3, 633–643.

Viala, J., and Mazodier, P. (2003). The ATPase ClpX is conditionally involved in the morphological differentiation of *Streptomyces lividans*. Mol. Genet. Genomics *268*, 5, 563–569.

Viala, J., Rapoport, G., and Mazodier, P. (2000). The clpP multigenic family in *Streptomyces lividans*: conditional expression of the *clpP3 clpP4* operon is controlled by PopR, a novel transcriptional activator. Mol. Microbiol. *38*, 3, 602–612.

Wang, K.H., Sauer, R.T., and Baker, T.A. (2007). ClpS modulates but is not essential for bacterial N-end rule degradation. Genes Dev. *21*, 4, 403–408.

van Wezel, G.P., and McDowall, K.J. (2011). The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances. Nat. Prod. Rep. *28*, 7, 1311.

Zellmeier, S., Schumann, W., and Wiegert, T. (2006). Involvement of Clp protease activity in modulating the *Bacillus subtilis* σW stress response. Mol. Microbiol. *61*, 6, 1569–1582.

Zhu, H., Sandiford, S.K., and van Wezel, G.P. (2014). Triggers and cues that activate antibiotic production by actinomycetes. J. Ind. Microbiol. Biotechnol. *41*, 2, 371–386.

ACKNOWLEDGMENTS

I feel very fortunate to have two thesis advisors to guide and support me over the course of my doctoral training. Thanks to my PhD supervisor, Tania Baker, for pushing me to grow as a scientist and a communicator—for teaching me the many values of clear writing and figures that 'show your work'. Thanks to my co-advisor, Bob Sauer, for illuminating the shortest path in designing experiments and sentences. I am deeply grateful for their sponsorship and patience as I navigated the bench and my entry into structural biology.

Thanks to my long-time committee members, Alan Grossman and Jing-Ke Weng—your philosophies towards basic science, teaching, and your own research fields have been wonderful examples. Thanks to Alexandra Deaconescu for serving on my defense committee. Thanks to Cathy Drennan whose commitment to diversity and inclusion, as well as to completing the FNR structure project, has been a source of optimism and opportunities to collaborate with many talented people, including Chi, Lindsey, Allena, and Becky. Thanks to Frank Solomon, whom I had the pleasure and honor of sharing a Zoom classroom with MSRP students, for uplifting and encouraging me from my beginnings in the Biology program, in Methods and Logic.

I am deeply indebted to my previous research mentors and college professors who gave me multiple chances in my professional development. As a first-gen student, these experiences were precious opportunities. Thanks to the many faculty and staff in the Williams College Biology and Chemistry Departments, including Sarah Goh, who welcomed me as a double major. Thanks to Amy Gehring for laying the foundation and providing a nurturing environment in my second home, the lab bench. Thanks to Lee Park, an excellent role model of a teacher-scholar, for introducing me to independent research and how to pay attention to my surroundings as our experiments are always teaching us something. Thanks to Lois Banta, who developed my scientific curiosity in her transformative classroom that connects the interfaces of biology, chemistry, and computation at the cellular and molecular level. Thanks to Gisela and Maria who cheered me on throughout college, making the chemistry stockroom a family-like work-study. Thanks to Jens Lohr and Todd Golub who listened and trusted me as an early-career research technician, for recognizing my potential and supporting me to pursue graduate school.

Many thanks to the past members of the Baker and Sauer labs: Adrian, AJ, Amaris, Ben, Branwen, Breann, Christina, Elise, Erin, Gina L., Gina M., Hema, Igor, Izarys, Julia, Karl, Kristin, Liz, Reuben, Sanjay, Steph, Ted, Tejas, Tristan, Vlad, Xue, and Zion. I am especially grateful to my co-authors, Izarys for establishing the ClpS junction mutations, Kristin for an excellent tag-team effort in the ClpA–ClpP crosslinking project, Tristan for extremely helpful discussions and feedback, Steph for pioneering the FNR–ClpXP recognition project, and Xue for teaching me the ropes to cryo-EM. Many, many thanks to current members of both labs (Alireza, Jia Jia, Juhee, Irene, Meghann, Reed, SaRa, and Tathagata). Thank you all (past and present) for your advice and friendship. I'm looking forward to hearing about all the great things you're doing once I leave MIT.

Most of all, thanks to my family, both biological and chosen. To my parents who sacrificed everything for their children, thank you, and to my prudent younger sibling who chose a logical career path and keeps an eye on us. To my friends who sustain me, thanks for supporting my entirety, such that some of your parents have even become adoptive family members. You are all inspiring people, and I am so lucky to be in your lives. To E–, who keeps me on the tracks and enriches my thinking and living, thanks.