Intersubunit Communication and Coordinated Mechanical Activity in the AAA+ Protease ClpXP

by

Tristan A. Bell

B.A., Molecular and Cell Biology
University of California, Berkeley, 2014

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

February 2020

Signature of Author……
December 13, 2019
Department of Biology

Certified by……
Robert T. Sauer
Salvador E. Luria Professor of Biology
Thesis Supervisor

Certified by……
Tania A. Baker
E.C. Whitehead Professor of Biology
Thesis Co-supervisor

Accepted by……...
Stephen P. Bell
Uncas and Helen Whitaker Professor of Biology
Co-chair, Biology Graduate Committee

Intersubunit Communication and Coordinated
Mechanical Activity in the AAA+ Protease ClpXP

by

Tristan A. Bell

Submitted to the Department of Biology on December 13, 2019
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ABSTRACT

Proteases belonging to the AAA+ (ATPases associated with various cellular activities) family perform regulated proteolysis in all domains of life by binding, mechanically unfolding, and degrading target proteins. The bacterial AAA+ protease ClpXP is composed of two distinct proteins: ClpX, a ring-hexamer protein unfoldase; and ClpP, a barrel-shaped tetradecameric peptidase. The assembly of ClpX ring hexamers results in extensive interaction between subunits, and the motor exhibits positive cooperativity in both ATP hydrolysis and mechanical activity against substrates. Despite general understanding of the mechanism of protein unfolding and degradation by ClpXP and other AAA+ proteases, how the six unfoldase subunits coordinate their mechanical activity to produce the force required to quickly and efficiently degrade stably folded substrates is unclear. Here, I present experiments that interrogate intersubunit communication and coordination of mechanical activity by *Escherichia coli* ClpXP.

In Chapter I, I review the current understanding of AAA+ protease structure and function as background to contextualize the findings presented in later chapters. In Chapter II, I present structural and functional characterization of a ClpX structural element, termed the hinge-linker, in facilitating communication between subunits of the ring hexamer. In Chapter III and two related Appendices, I present experiments that systematically identified determinants of grip between ClpX and its substrates. These experiments also identified distinct functions for different unfoldase subunits during application of force to bound substrates. In Chapter IV, I present results from a collaborative project that determined structures of ClpXP bound to a protein substrate and biochemically characterized several previously unvisualized elements of ClpX and ClpP. In chapter V, I report the effects of inhibiting relative rotation of ClpX and ClpP on ATP hydrolysis and mechanical activity of the ClpX unfoldase. Using the constraints on mechanism inferred from these findings, I also propose molecular models for processive mechanical activity. Finally, in Chapter VI, I discuss the results presented in previous chapters in the larger context of communication and coordination between subunits of AAA+ protein unfolding motors.

Thesis Supervisor: Robert T. Sauer
Title: Salvador E. Luria Professor of Biology

Thesis Co-supervisor: Tania A. Baker
Title: E.C. Whitehead Professor of Biology

ACKNOWLEDGMENTS

First and foremost, I am grateful to my doctoral advisor, Bob Sauer. Bob has devoted a substantial amount of time and energy not only to the papers we've authored together, but also to my professional development as a scientist. I couldn't have asked for a better mentor. I'm also grateful to my co-advisor Tania Baker, whose insightful perspectives have guided this work. My thesis committee members Steve Bell and Amy Keating have been incredibly generous with their time and counsel throughout my graduate career. I also owe a great deal to my undergraduate mentors James Berger and Kathleen Collins, who encouraged me to pursue this path.

My work on this project would not have developed as it did without the input and encouragement of many colleagues. In particular, I'm indebted to my labmate and collaborator Xue Fei, who has unreservedly shared her expertise in structural biology. I also thank Karl Schmitz, who gave generously of his time to help me define my first projects. My labmates in the Sauer and Baker labs helped bring joy to my research even in difficult times. I'm especially thankful for the companionship of my fellow graduate students: A.J. Amor, Vlad Baytshtok, Juhee Morehouse, Gina Mawla, Sora Kim, Kristin Zuromski, Irene Shih, and Meghann Kasal.

My doctoral work was also shaped by the mentorship of my first advisor, Wendy Gilbert. I'm grateful to Wendy and all the members of the Gilbert lab I had the privilege of working with, especially Cassandra Schaening, Julia Wang, Gina Mawla, Thomas Carlile, Kristen Bartoli, Paritosh Gangaramani, Maria Fernanda, and M.K. Thompson. I also thank Uttam RajBhandary and David Bartel, who provided excellent advice during my time in the Gilbert lab.

This work would have been nowhere near as joyful without the many friends who have come along on the journey. Thanks to everyone who shared in the running, hiking, climbing, and beer brewing, especially Julia Wang, Cassandra Schaening, and Josh Saul. I'm deeply grateful to Rohan Jonnalagadda, who has shared in this journey during its highest and lowest moments. Thank you to my brothers Kyle Mann, Aaron Gipson, Harry Heng, and James Frederick, who have kept me grounded through times of want and times of plenty. Finally, my thanks to Karen and Steve Cunningham. Your hope and optimism have uplifted and inspired me on this path, and I'm glad to be your friend.

I'm especially grateful to my parents and grandparents, who encouraged and enabled me to pursue higher education and doctoral training as a scientist. This thesis is the culmination of much time and toil, both my own and theirs. Thank you.

My sister Grace has been my friend and co-conspirator over the past five years. Grace, more than anyone else, you have inspired me to keep pressing forward during difficult times. I'm proud of all your hard work and achievement, and excited for the promise that lies ahead of you.

Finally, I'm grateful to my teammate Kathleen Davis. Kathleen and I met in a parking lot my first week in Boston, where she told me a story about running into a grizzly bear while hiking in Montana. Little did I know then that soon we'd be traveling across the country and having unexpected encounters with wildlife together. Thank you Kathleen – this journey wouldn't have been half as wonderful without you. Let's keep adventuring.

TABLE OF CONTENTS

Chapter I

Introduction

All cells have proteomes that are in continual flux, and must constantly adapt to environmental changes. In bacteria, eukaryotes, and archaea, proteases belonging to the AAA+ (ATPases associated with various cellular activities) protein family bear primary responsibility for targeting and destroying proteins that are damaged, misfolded, or no longer required. AAA+ proteases are composed of two elements: an unfoldase that consumes cellular energy to unfold substrates, and an energy-independent peptidase that degrades targets after unfolding. In this introductory chapter, I will provide context helpful for understanding the structural and functional studies of *E. coli* ClpXP (a model member of the AAA+ protease family) presented in later chapters. I will first review the broad family of AAA+ proteases and their roles in maintaining cellular proteostasis. I will then detail the current understanding of AAA+ protease structure and function with an emphasis on *E. coli* ClpXP. I will describe what is currently known about how ClpXP binds and grips substrates during the application of unfolding force. I will then discuss current understanding of communication and coordinated action between ClpX subunits. Finally, I will present an overview of structure-based models for substrate unfolding and translocation by AAA+ proteases, laying out the strengths and weaknesses of each model in light of biochemical studies.

AAA+ proteases maintain proteostasis through targeted remodeling and destruction

Members of the AAA+ protein family perform diverse roles in all living cells. These roles include DNA unwinding, homologous recombination, histone remodeling, vesicle fusion, chaperoned protein folding, microtubule severing, and protein unfolding and degradation (Snider et al., 2008). Despite performing a vast array of cellular functions, AAA+ enzymes share many common structural and functional attributes. First, all family members couple the energy of ATP hydrolysis to conformational changes that perform work on macromolecular substrates (Hanson and

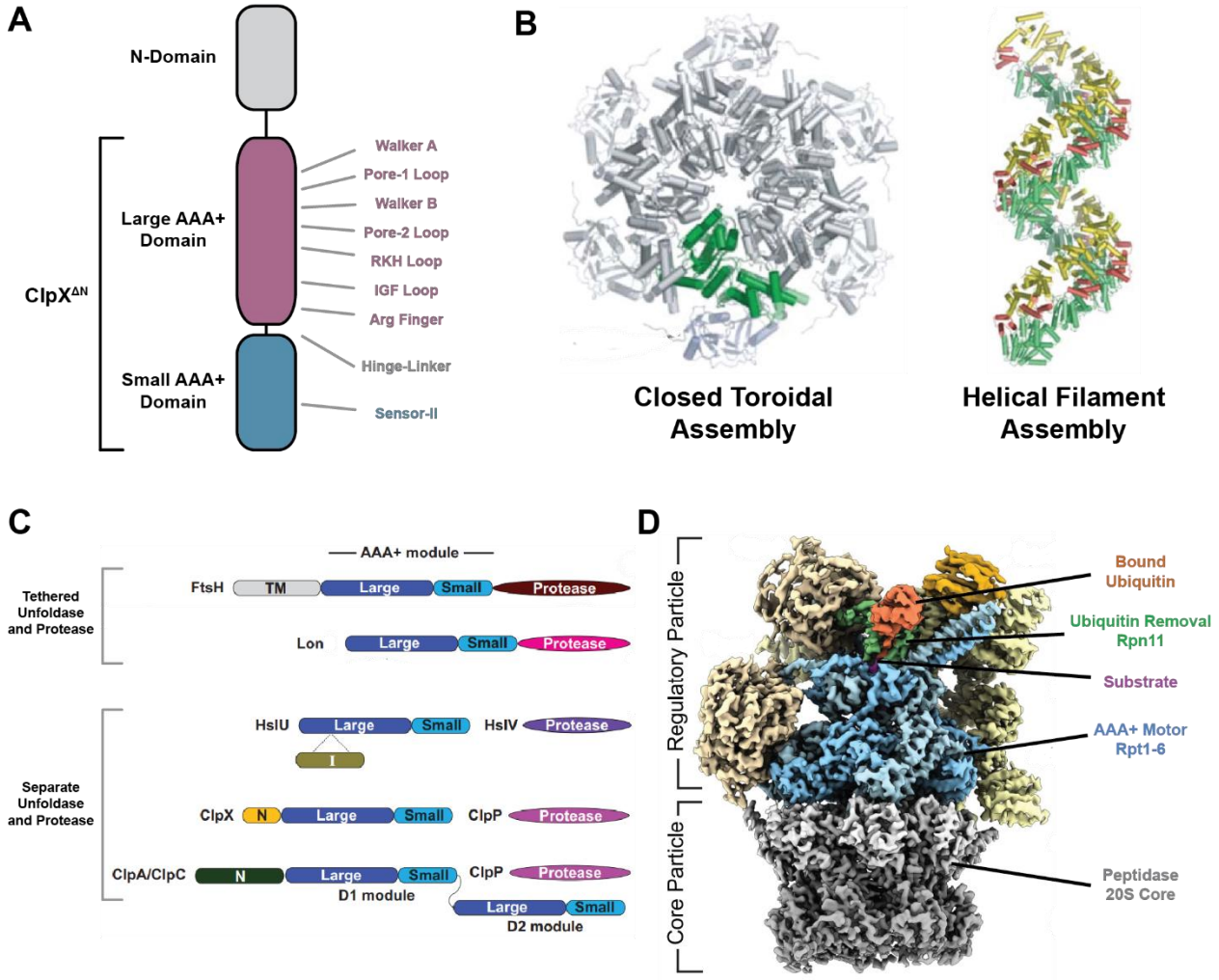


Figure 1.1 – Structure and assembly of AAA+ machines

(A) ClpX domain architecture. **(B)** Assembly of AAAs into hexamer rings or helical filaments. Adapted from Erzberger and Berger, 2006. **(C)** Domain structure of the five *E. coli* AAA+ proteases. "TM", "I", and "N" indicate a transmembrane domain, intermediate domain, and N-terminal accessory domain, respectively. Adapted from Sauer and Baker, 2011. **(D)** Cryo-EM 3D model of the *S. cerevisiae* 26S proteasome. Adapted from de la Peña et al., 2018.

Whiteheart, 2005). Second, all adopt a common overall structure, composed of a Rossman fold (large domain) and an alpha-helical bundle (small domain), connected through a flexible hinge-linker (Figure 1.1A) (Erzberger and Berger, 2006). Finally, all form multimeric assemblies, usually either helical arrays or closed hexameric toroids, stabilized by interactions between the large domain of one subunit and the small domain of its neighbor (Figure 1.1B).

Within the broad family of AAA+ enzymes, AAA+ proteases perform the important work of unfolding and degrading other proteins (Sauer and Baker, 2011). In the cell, proteins are employed in nearly every function, and their production can be highly regulated based on signals from the extracellular environment. AAA+ proteases perform the important regulatory counterpart to protein expression by targeting and degrading proteins that are no longer required, reducing their abundance in the cellular pool (Sauer and Baker, 2011). Most intracellular proteins exist in steady state rather than a static state – the average protein in a growing yeast cell has a cellular half-life of approximately 45 minutes, and approximately 30 hours in a nondividing HeLa (human) cell (Belle et al., 2006; Cambridge et al., 2011).

AAA+ proteases in bacteria

Several AAA+ proteases work together to maintain proteostasis in the bacterial cytoplasm. The gram-negative bacterium *E. coli* harbors five such enzymes: ClpAP, ClpXP, FtsH, HslUV, and Lon (Figure 1.1C) (Schmidt et al., 2009). Some bacteria (including *Mycobacteria* and *Actinobacteria*) and archaea harbor additional proteases, including ClpCP, Mpa–20S, and PAN–20S (Bar-Nun and Glickman, 2012; Imkamp et al., 2015). Each of these degradation machines is composed either of genetically tethered AAA+ unfoldase and protease domains (as in FtsH and

Lon) or as a collaboration between separate unfoldase and protease proteins (e.g. ClpA/ClpC/ClpX with ClpP, HslU with HslV, and Mpa/PAN with 20S) (Figure 1.1C).

ClpXP recognizes several endogenous and phage-encoded substrates and regulates their intracellular concentrations through degradation (Baker and Sauer, 2012). These proteins are recognized via specific amino acid sequences, termed degrons, that are sufficient to initiate protein degradation (Flynn et al., 2003). For example, the *C. crescentus* ClpX N-terminal domain binds directly to the SocAB toxin-antitoxin complex, mediating degradation that prevents toxin accumulation and cell growth inhibition (Vass et al., 2017). Adaptor proteins can also program ClpXP degradation activity. CpdR interacts with the N-terminal domains of *C. crescentus* ClpX and recruits several protein substrates for degradation (Smith et al., 2014). Because CpdR is variably phosphorylated throughout the cell cycle and specifically binds ClpXP when unphosphorylated, the adaptor-protease complex mediates cell cycle-dependent degradation of its targets (Lau et al., 2015).

ClpXP also plays a larger role in global protein quality control by recognizing and degrading proteins marked with the ssrA degron, which are generated as a result of nonstop protein translation (Figure 1.2) (Dulebohn et al., 2007). Nonstop protein translation occurs when a translating ribosome reaches the end of an mRNA without encountering a stop codon, due to abortive mRNA transcription, mRNA cleavage, or errors in mRNA decoding (Keiler and Feaga, 2014). Because a stop codon is required for ribosomal subunits to release from mRNA, nonstop translation leads to a stalled ribosome complex, impairing the cell's translational capacity (Karzai et al., 2000). In all bacteria, the majority of such stalled ribosomes are rescued through the tmRNA pathway (Giudice

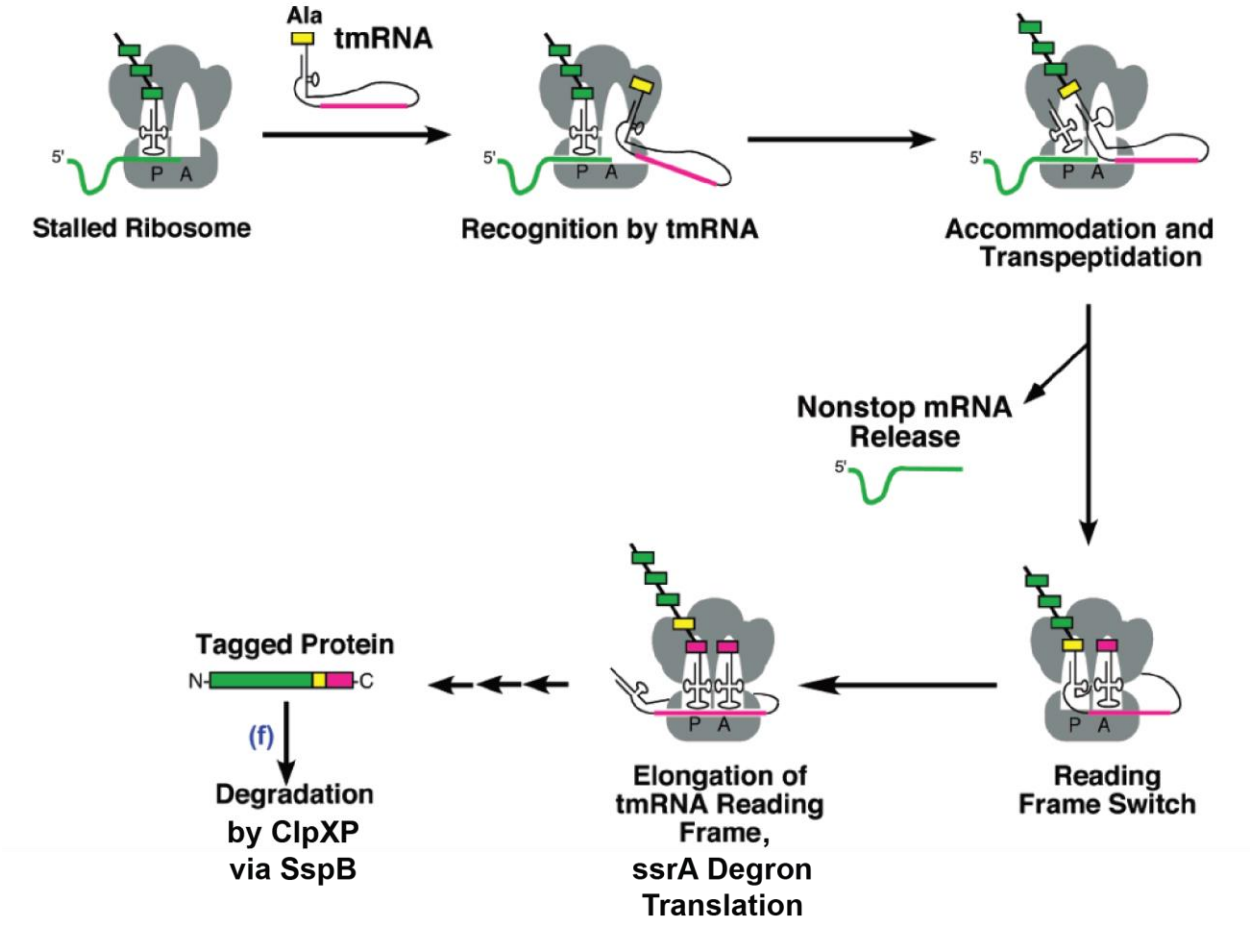


Figure 1.2 – Overview of the tmRNA ribosome rescue pathway
Translation of tmRNA produces protein products with a C-terminal ssrA degron, which is targeted to ClpXP for degradation via the adapter protein SspB. Adapted from Dulebohn et al., 2007.

and Gillet, 2013). Stalled ribosomes recognize and bind tmRNA, a small RNA that exhibits features of both a tRNA and an mRNA (Komine et al., 1994; Ushida et al., 1994). The tmRNA encodes an 11-residue sequence, termed the ssrA tag, which is translated at the C-terminus of the nonstop translation product (Keiler et al., 1996). The ssrA tag is then recruited to ClpXP directly or with the assistance of the adapter protein SspB (Levchenko et al., 2000). ClpX then recognizes the ssrA tag and binds, unfolds, and translocates the protein into ClpP for degradation (Bolon et al., 2004). As a result, ClpXP plays an important role in preventing the accumulation of potentially toxic mistranslated proteins in the cell.

Other AAA+ enzymes in *E. coli* also contribute to proteolysis. Lon performs the majority of protein degradation within the cell, and recognizes damaged or misfolded proteins through exposure of hydrophobic amino acids in unstructured polypeptide segments (Van Melderen and Aertsen, 2009). ClpAP recognizes the ssrA degron like ClpXP, though it also tunes the degradation of folded proteins. This regulation is mediated through the N-end rule, in which proteins are delivered to ClpAP by the adapter protein ClpS based on its affinity for the protein's N-terminal ultimate and penultimate residues (Sauer and Baker, 2011). HslU and HslV are specifically overexpressed in response to heat shock stress and operate most efficiently at high temperatures (Burton et al., 2005; Lien et al., 2009). HslU is thought to recognize and engage free protein termini, possibly as an indicator of partial protein unfolding (Sundar et al., 2010). FtsH is the only membrane-tethered AAA+ protease in *E. coli*, and targets damaged, misfolded, and unneeded membrane proteins for degradation via exposure of their termini (Langklotz et al., 2012). Like ClpXP, ClpAP, and Lon, FtsH also degrades proteins with a C-terminal ssrA tag, suggesting that

it may be involved in quality control of mistranslated membrane or membrane-associated proteins (Hari and Sauer, 2016).

In some bacterial species, post-translational modifications can also target proteins for destruction by AAA+ proteases. For example, the N-terminal auxiliary domain of *B. subtilis* ClpCP recognizes and binds proteins that present phosphoarginine residues, facilitating their degradation (Trentini et al., 2016). In *Mycobacteria*, Mpa–20S recognizes and degrades proteins modified with prokaryotic ubiquitin-like protein (Pearce et al., 2008). Despite being evolutionarily distinct, this mechanism is similar to ubiquitin-mediated substrate recognition in eukaryotes, as discussed below (Imkamp et al., 2015).

AAA+ proteases in eukaryotes

In eukaryotic cells, the majority of cytoplasmic and nucleoplasmic proteolysis is carried out by the 26S proteasome, a large protein complex composed of the 19S regulatory particle and the 20S core particle (Figure 1.1D) (Bard et al., 2018). Within 19S, the Rpt1-6 AAA+ motor catalyzes unfolding and translocation of substrates, and feeds them into the proteolytic core of the 20S particle for degradation. Proteolytic degradation is primarily mediated through modification of substrates with ubiquitin, a small protein domain (Oh et al., 2018). A large family of E3 ubiquitin ligase enzymes regulate modification of other proteins with ubiquitin, and each modifies a specific set of targets (Zheng and Shabek, 2017). Substrates can be targeted via their terminal protein sequences, recognition of specific protein structures, recognition of other post-translational modifications, or by direct handoff from chaperones (Koren et al., 2018; Oh et al., 2018; Zheng and Shabek, 2017). Once a protein has been specifically marked as a degradation substrate with a chain of ubiquitin

moieties, the protein is recognized by the ubiquitin-binding protein Rpn13 at the surface of the 19S regulatory particle. Rpn11 and Ubp6 then de-ubiquitinate the substrate concurrent with engagement by the Rpt1-6 motor, liberating the ubiquitin domains for re-use (Bard et al., 2018). The protein is concurrently unfolded by Rpt1-6 and translocated into the 20S core particle, where it is degraded into short peptides.

Although the Ubiquitin-Proteasome Pathway (UPP) is responsible for the majority of intracellular degradation, an additional pathway assists in degradation of membrane proteins and soluble proteins within the endomembrane system. A double-ring AAA+ motor known as Cdc48 in yeast and p97/VCP in humans assists in ER-, mitochondria-, ribosome-, and chromatin-associated degradation (Ye et al., 2017). In most cases, Cdc48/ p97/VCP coordinates with the UPP by handing substrates off to ubiquitin ligases, which mark them for degradation by the 26S proteasome. However, archaeal Cdc48 can directly interact with the 20S core proteasomal particle, and disease-linked mutations in human p97/Cdc48 imply relevance for this mechanism in eukaryotes as well (Barthelme and Sauer, 2012; Barthelme et al., 2015).

Eukaryotes also express relatives of the bacterial AAA+ proteases within plastid organelles. Both mitochondria and chloroplasts express variants of bacterial motors including ClpXP, Lon, and FtsH (Adam et al., 2001; Glynn, 2017; Ostersetzer et al., 2007). Although relatively little is known about these enzymes relative to their cytosolic counterparts, it has been proposed that specific protein quality control functions may be more consequential than their processive degradation activities (van Dyck et al., 1998; Kardon et al., 2015; Röttgers et al., 2002), as discussed below.

Degradation-independent protein remodeling activity

In addition to their primary role in protein degradation, AAA+ proteases also play roles in protein structural remodeling. Against some substrates, ClpX and ClpA can act as chaperones independent of ClpP by unfolding proteins and allowing them a second chance to fold correctly (Burton and Baker, 2005; Wickner et al., 1994). For example, in yeast and human mitochondria, ClpX partially unfolds alpha-linolenic acid synthase to facilitate cofactor loading, activating the enzyme to catalyze formation of an essential heme biosynthesis precursor (Kardon et al., 2015; Yien et al., 2017). As another example, the mitochondrial intermembrane-space protease Yme1 induces cleavage in Opa1 to create distinct membrane-tethered and untethered isoforms, a process necessary for mitochondrial fusion and efficient respiratory energy production (Del Dotto et al., 2018).

Structure and function of ClpXP, a model AAA+ protease

E. coli ClpXP is a well-studied enzyme and serves as a model system for exploring AAA+ protease structure and function. ClpXP is composed of distinct unfoldase (ClpX) and protease components (ClpP). The AAA+ unfoldase ClpX assembles into closed-ring homohexamers, and each subunit contains an N-terminal domain and large and small AAA+ domains (Figure 1.1A-B) (Baker and Sauer, 2012). The two AAA+ domains are connected through a short, flexible hinge-linker (Glynn et al., 2012). Ring hexamers assemble through interactions between the large domain of one subunit and the small domain of its neighbor, with nucleotide binding sites at the interface between protomers (Figure 1.3A) (Hanson and Whiteheart, 2005). The N-terminal domains form three dimers at the surface of the hexamer, and mediate interactions with SspB (Wah et al., 2002). ClpX undergoes conformational changes driven by ATP hydrolysis that initially unfold substrates and

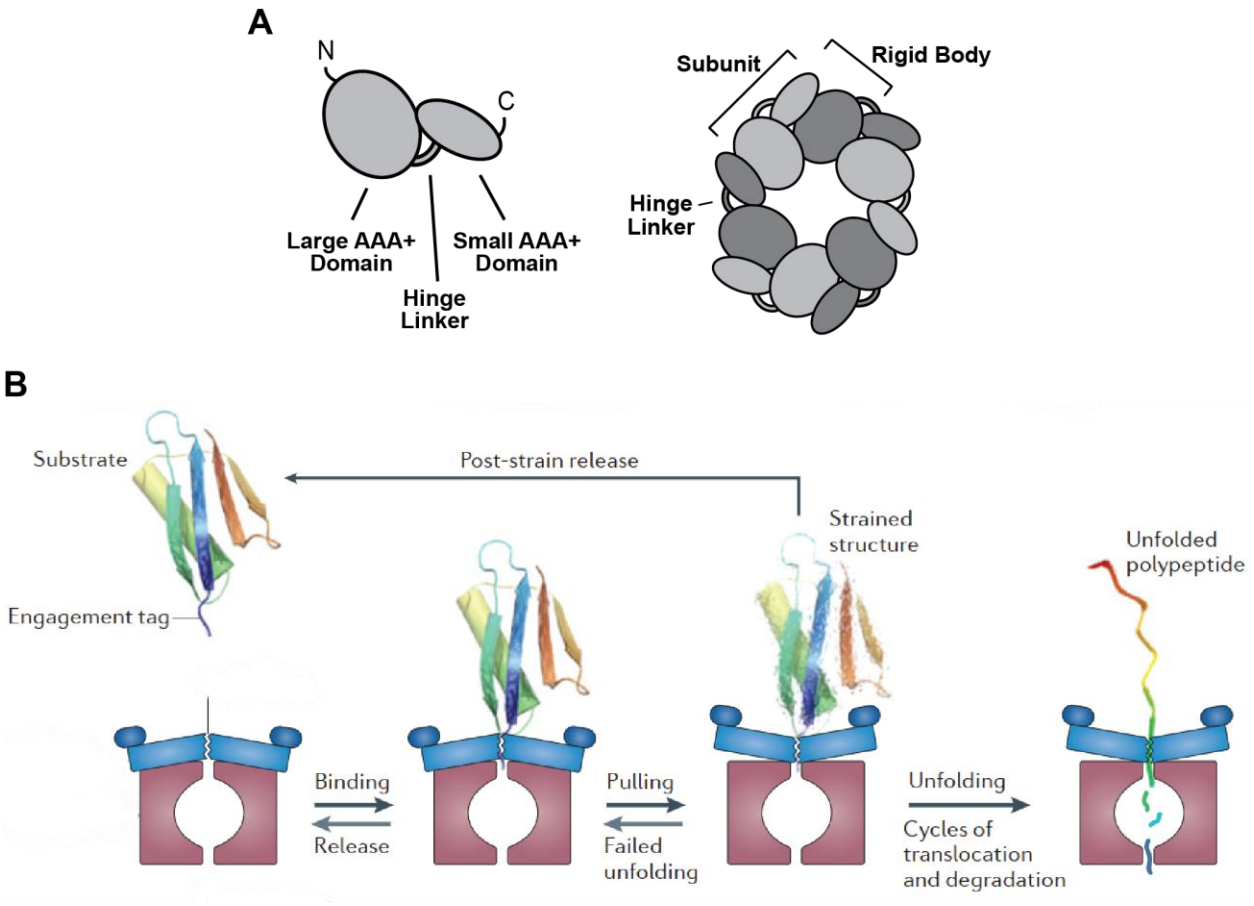


Figure 1.3 – ClpX assembly and substrate unfolding mechanism

(A) ClpX monomers assemble into closed-ring hexamers through interactions between large and small AAA+ domains of adjacent protomers. **(B)** Mechanism of protein unfolding and degradation by ClpXP. Adapted from Olivares et al., 2016.

then processively pull the unfolded polypeptide through its axial pore and into the partner protease ClpP for degradation (Figure 1.3B) (Baker and Sauer, 2012).

The ClpP protease is a homo-tetradecamer composed of two stacked heptameric rings, together forming a barrel-shaped complex with all 14 active sites sequestered within the interior of the chamber (Wang et al., 1997). Subunits are expressed with an N-terminal propeptide, which renders the protease inactive until it is properly assembled (Liu et al., 2014). Once assembled into 14-mers, $ClpP_{14}$ autocatalytically cleaves off its propeptide sequences, rendering it fully active (Maurizi et al., 1990). Ensuring regulated protein degradation through a gated and compartmentalized protease is a common strategy among AAA+ proteases, and is evolutionarily conserved between the bacterial HslV and eukaryotic and archaeal 20S proteasomal peptidases (Bochtler et al., 1997, 2000; Tomko and Hochstrasser, 2013).

Although ClpX and ClpP have not been co-crystalized, crystal structures of ClpX and ClpP have been solved separately. ClpX crystal structures show a variety of nucleotide binding states and conformations (Figure 1.4A) (Glynn et al., 2009; Stinson et al., 2013). Subunits in these crystal structures exhibit large rotations relative to each other, with some subunits rotated into conformations that preclude nucleotide binding. However, several loop regions that are predicted to contact substrates and ClpP were too flexible to be resolved in these structures, leaving many aspects of ClpX function unclear. Crystal structures of ClpP have facilitated mechanistic understanding of how the axial pores at either end of the barrel are gated (Liu et al., 2014). These gates control interaction with ClpX and ClpA, and are the site where unfolded proteins are spooled into the proteolytic inner chamber (Figure 1.4B) (Grimaud et al., 1998). The axial gate of ClpP

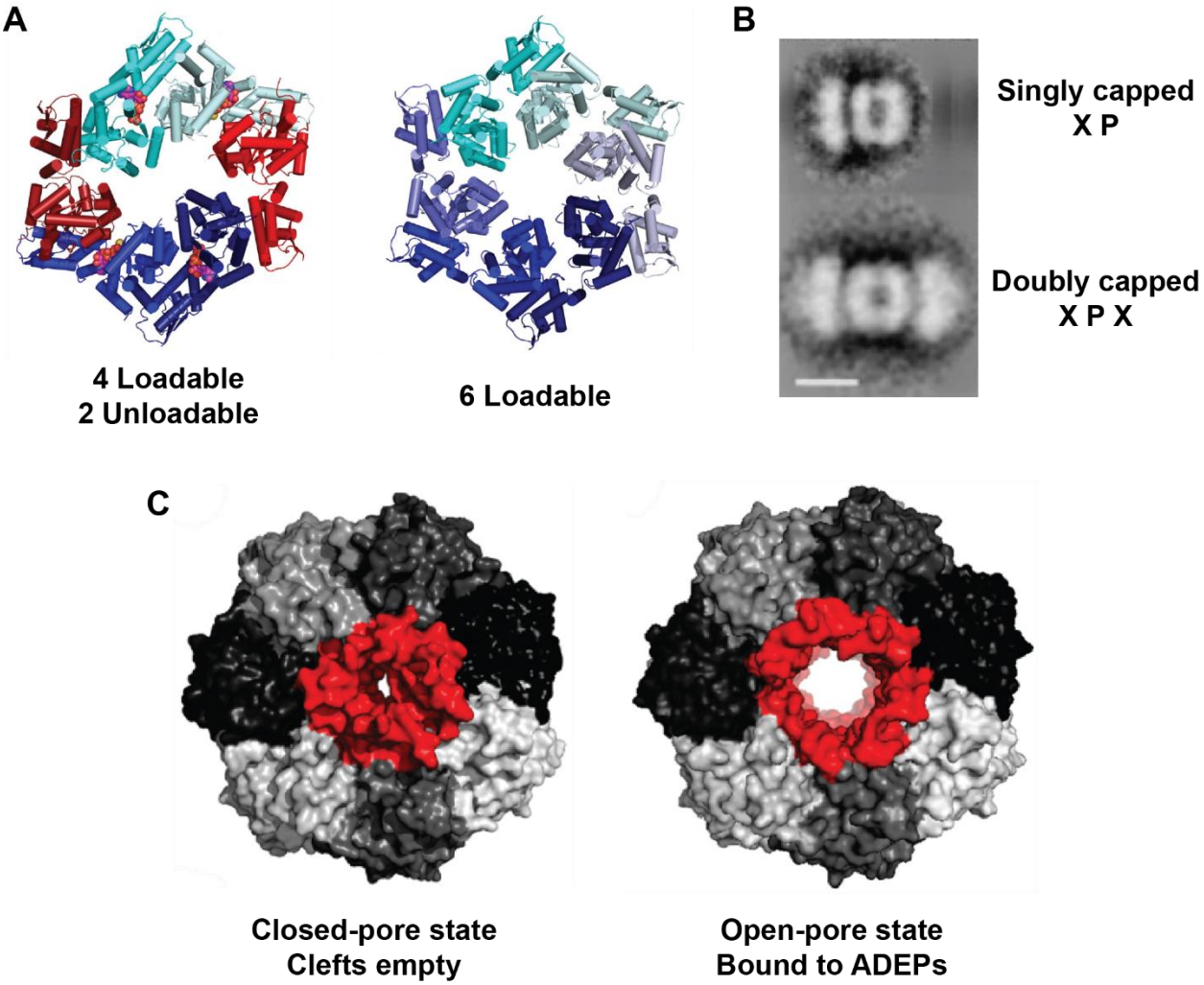


Figure 1.4 – Conformational variability in ClpX and ClpP

(A) Comparison of crystal structures of ClpX: (left) four subunit interfaces are positioned to bind nucleotide, and two subunits are not in a binding competent state; (right) all six subunits are in a conformation competent to bind nucleotide. Adapted from Stinson et al., 2013. **(B)** Negative-stain electron-microscopy 2D class averages of singly capped (top) and doubly capped (bottom) ClpXP complexes. Adapted from Grimaud et al., 1998. **(C)** Crystal structures of ClpP in (left) pore-closed (Bewley et al., 2006) and (right) pore-open (Li et al., 2010) states. Adapted from Baker and Sauer, 2006.

adopts a closed state in the absence of binding partners, with the pore narrowed to prevent promiscuous degradation of non-substrate proteins (Figure 1.4C) (Lee et al., 2010b). Binding of ClpX or ClpA to ClpP opens the axial gates to permit translocation of unfolded substrates from the AAA+ motor (Effantin et al., 2010; Lee et al., 2010b). The open state of the axial pore has been visualized in crystal structures of ClpP bound to ADEPs, a class of small-molecule antibiotics that mimic ClpX binding contacts and allow diffusion of large peptides and unfolded proteins through the gates (Brötz-Oesterhelt et al., 2005; Lee et al., 2010a; Li et al., 2010).

ClpX interacts with ClpP through a multivalent series of contacts between the two enzymes. Loops at the lower surface of the ClpX ring (termed IGF loops for a conserved Ile-Gly-Phe tripeptide they contain) dock into a series of hydrophobic clefts surrounding the axial gate of ClpP (Joshi et al., 2004; Kim et al., 2001; Martin et al., 2007). The IGF loops only adopt conformations competent to bind ClpP when ClpX is bound to ATP (Joshi et al., 2004). ClpX-ClpP binding affinity is very sensitive to IGF loop mutations and to binding of ADP by ClpX, suggesting that interactions between individual IGF loops and ClpP are highly dynamic (Amor et al., 2016, 2019; Kim et al., 2001; Martin et al., 2007). ClpP binding is also mediated through interactions between a series of loops within the lower pore of ClpX, termed the pore-2 loops, and a series of N-terminal loops that surround the axial pore of ClpP (Martin et al., 2007). ClpA binds to ClpP via a similar mechanism, with a series of IGF-related IGL loops docking into the ClpP hydrophobic clefts (Effantin et al., 2010). Other AAA+ motors contain C-terminal motifs that mediate binding to and activation of a partner protease. For example, the Rpt1-6 proteasomal motor and Cdc48/p97/VCP interact with their partner proteases using a C-terminal φYX (hydrophobic-Tyr-any residue) motif

(Barthelme and Sauer, 2012; Smith et al., 2007). Likewise, the conserved hydrophobic C-terminus of HslU activates peptidase activity in HslV (Seong et al., 2002).

With the help of the SspB adapter protein, ClpX can receive ssrA-tagged proteins directly targeted for degradation from the ribosome-rescue pathway. SspB interacts with the ClpX N-terminal domains and presents substrates to ClpX in a manner that facilitates productive binding and handoff (Levchenko et al., 2000). Similarly, ClpAP interfaces with the ClpS adapter, which delivers and enhances degradation of N-end rule substrates (Erbse et al., 2006). Many proteins in the 19S proteasomal regulatory particle can be considered an extension of this adapter protein strategy, with a complex suite of adapters enhancing ubiquitin recognition, removal, and recycling (Bard et al., 2018).

ClpX first interacts with substrate proteins at the apical surface of the central pore with a series of positively-charged loops termed the RKH loops (named for a conserved Arg-Lys-His motif) (Baker and Sauer, 2012). These loops act as a selectivity filter and are hypothesized to electrostatically promote interaction with the C-terminal carboxylate of degrons such as ssrA (Farrell et al., 2007). The RKH loops are unique to bacterial ClpX, and may reflect its primary role in recognition of ssrA-tagged proteins. After passing the RKH loops, the substrate is bound by pore-1 and pore-2 loops within the axial pore (Martin et al., 2008a). These loops are widely conserved among AAA+ protein remodeling motors (Sauer and Baker, 2011). Mutation of pore-1 or pore-2 loops is detrimental to substrate binding and degradation (Martin et al., 2008a, 2008b; Siddiqui et al., 2004).

Recent cryo-electron microscopy (cryo-EM) structures of many AAA+ proteases and protein remodelers including Yme1, Agf3L2, the 26S proteasome, Cdc48, PAN, Vps4, VAT, NSF, katanin, Rix7, Hsp104, and ClpB have helped clarify the native conformations of many AAA+ motors (Cooney et al., 2019; Deville et al., 2017; Ding et al., 2017; Dong et al., 2019; Gates et al., 2017; Han et al., 2017; Lee et al., 2010b; Lo et al., 2019; Majumder et al., 2019; Michalska et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017, 2019; Ripstein et al., 2017; Rizo et al., 2019; Su et al., 2017; Sun et al., 2017; Twomey et al., 2019; White et al., 2018; Yokom et al., 2016; Yu et al., 2018; Zehr et al., 2017; Zhu et al., 2018). In essentially all of these structures, the six AAA+ protomers adopt a shallow spiral assembly in which most subunits assume conformations competent to bind nucleotide (Figure 1.5A). In all substrate-bound structures, the pore-1 and pore-2 loops interact with the bound polypeptide with a two-residue periodicity (Figure 1.5B). A published cryo-EM structure of ClpXP from *L. monocytogenes* argues that ClpX adopts a symmetric rather than a spiral structure (Gatsogiannis et al., 2019). In contrast with this structure, cryo-EM structures of ClpXP presented in Chapter IV show that ClpX adopts a structure and pattern of pore loop conformations similar to other AAA+ protein remodeling motors.

Substrate grip and sustained substrate interaction

Although ClpXP readily recognizes proteins marked with an appropriate degron, not all proteins that are initially bound are successfully degraded. When pulling on the terminus of a stably folded substrate, ClpXP cycles through many ATP hydrolysis events that fail to denature the substrate (Kenniston et al., 2003). Substrates that fail to unfold may remain associated with ClpX during repeated unfolding attempts, but because ClpX–substrate interactions are dynamic, the substrate

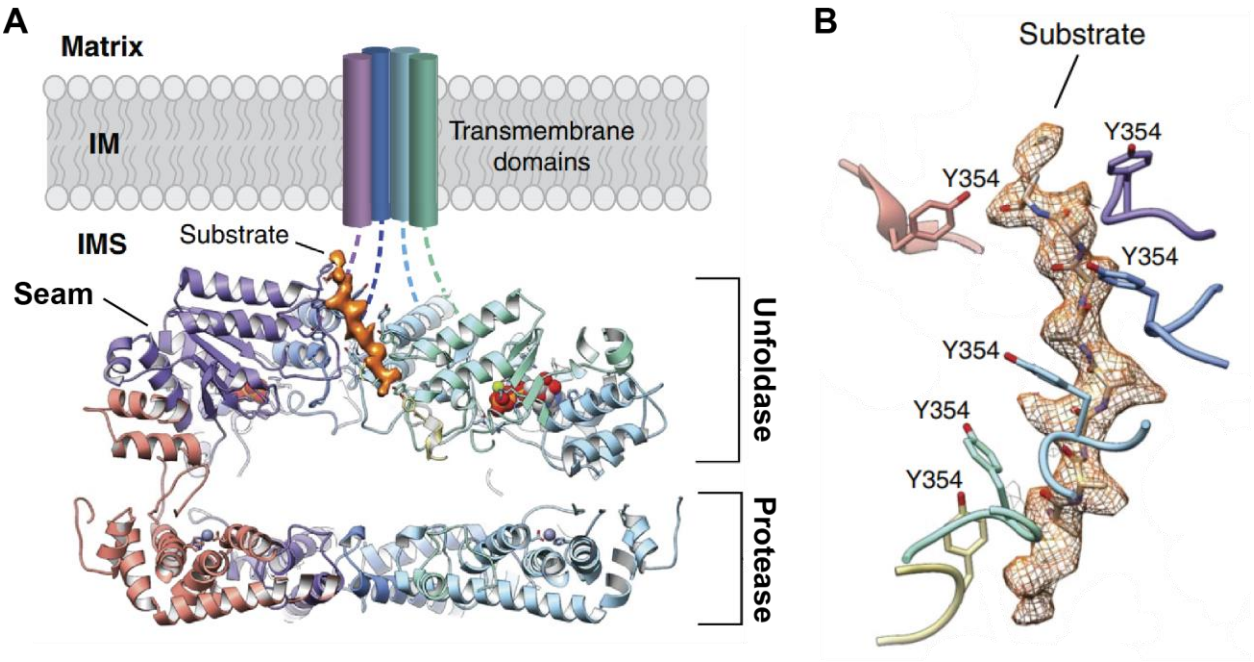


Figure 1.5 – Cryo-EM structures reveal novel features of AAA+ proteases
(A) Cryo-EM structure of the Yme1 protease bound to a substrate shows spiral subunit arrangement in the AAA+ motor. The seam between the highest- and lowest-positioned subunits is indicated. **(B)** Pore-1 loops (Tyr-354) of Yme1 intercalate with an extended substrate polypeptide with a two-residue periodicity. Two-residue periodic intercalation is a common feature of other AAA+ motors as well. Both panels are adapted from Puchades et al., 2017.

may ultimately be released back into solution (Kraut, 2013; Kraut et al., 2012; Too et al., 2013). This was elegantly demonstrated by the observation that when processing multi-domain substrates, ClpXP produces a ladder of partially degraded products corresponding to different numbers of successfully degraded domains (Kenniston et al., 2005).

The pore-1 loops of ClpX contain a conserved Tyr-Val-Gly motif, and Tyr and Val interact directly with bound substrates to produce substrate grip, defined here as the ability of ClpXP to maintain association with a substrate while transmitting the force necessary for unfolding. ClpXP–substrate grip during unfolding is analogous to a climber scaling a rock wall (Figure 1.6A). Just as the climber must supply enough grip to the wall to overcome their weight and move upwards, ClpXP must supply sufficient grip to a substrate to overcome a folded domain's free energy of folding and pull it down into the axial pore. Mutation of Tyr and Val residues impairs substrate unfolding and translocation activity (Iosefson et al., 2015a, 2015b; Martin et al., 2008b; Rodriguez-Aliaga et al., 2016; Siddiqui et al., 2004). When ClpX pore-1 loop Tyr residues are mutated to Ala in increasing numbers of subunits, the motor successively loses its ability to unfold decreasingly stable substrates (Iosefson et al., 2015a). This suggests that the pore-1 loops together exert a combinatorial effect on grip during substrate unfolding. Intriguingly, these mutations also accelerate ATP hydrolysis by ClpX, suggesting crosstalk between pore-1 loops and nucleotide binding sites (Iosefson et al., 2015a; Martin et al., 2008b). In further support of this model, the pore-1 loop in a given ClpX subunit is only conformationally positioned to interact productively with substrate when the subunit is bound to ATP (Siddiqui et al., 2004). It is therefore possible that fast ATP hydrolysis in multiple subunits could render the motor transiently incapable of binding substrate until ADP is re-exchanged for ATP.

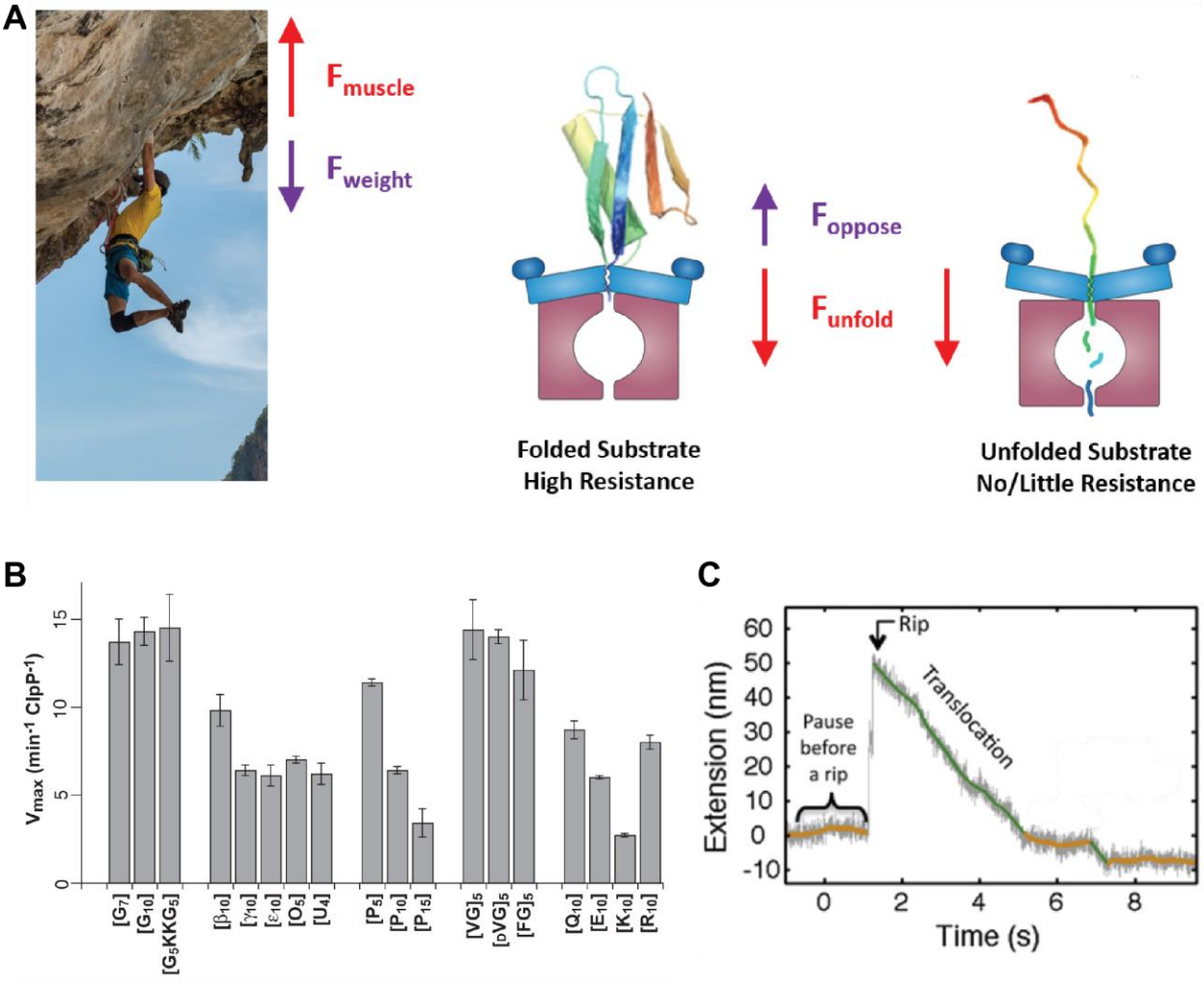


Figure 1.6 – Substrate grip during protein unfolding by ClpXP

(A) Simplified force diagrams for a climber scaling a vertical rock wall (left) and ClpXP doing work on a stably folded (middle) or unfolded (right) substrate. ClpXP images adapted from Olivares et al., 2016. **(B)** Maximal translocation rates of ssrA-tagged peptides by ClpXP. Each peptide contains a region containing the indicated set of natural or unnatural amino acids. Adapted from Barkow et al., 2009. **(C)** Single-molecule optical trap measurement of translocation of a GFP domain by a ClpXP complex. The rate of translocation, measured here as change in substrate extension, is roughly linear with respect to time. Adapted from Maillard et al., 2011.

By nature of being a general purpose unfoldase-protease, ClpXP must be able to translocate polypeptides of all compositions through its central pore. In a study of peptide translocation, rates varied by at most six-fold between peptides containing stretches of acidic, basic, or hydrophobic residues (Figure 1.6B) (Barkow et al., 2009). Peptides containing stretches of unnatural amino acids with irregular peptide bond spacing were also translocated efficiently, demonstrating that strict side chain spacing is not required for substrate translocation. These findings have been further substantiated by measurements of translocation along true substrates. In single-molecule optical trapping studies, both ClpXP and ClpAP process across unfolded polypeptides at a near constant rate, regardless of the sequence being translocated (Figure 1.6C) (Aubin-Tam et al., 2011; Cordova et al., 2014; Olivares et al., 2017; Rodriguez-Aliaga et al., 2016; Sen et al., 2013).

Although all substrate sequences are translocated similarly, not all sequences equivalently promote substrate domain unfolding. When a AAA+ unfoldase binds a substrate and attempts to unfold a domain, it must transmit unfolding force to the domain through interactions with the unstructured residues directly adjacent to the folded domain (Figure 1.6A). In nature, several viral, bacterial, and eukaryotic proteins have been characterized that cannot be unfolded and degraded because unfavorably gripped amino-acid sequences neighbor the protected domain (Daskalogianni et al., 2008; Levitskaya et al., 1997; Lin and Ghosh, 1996; Vass and Chien, 2013). These "slippery sequences" tend to be rich in Gly and Ala residues, with the remaining amino acids predominantly having small and polar side chains (Tian et al., 2005). Several studies *in vitro* have confirmed that low complexity tracts of Gly and Ala can frustrate unfolding of stably folded substrates by both ClpXP and the proteasome (Hoyt et al., 2006; Kraut, 2013; Kraut et al., 2012; Sharipo et al., 2001; Tian et al., 2005; Too et al., 2013; Zhang and Coffino, 2004). However, it is unclear whether

AAA+ proteases fail to grip these sequences because they interact poorly with small residues or whether the low complexity of the sequence inhibits unfolding (Tian et al., 2005).

Communication and coordinated action between motor subunits

All six subunits of the ClpX protein-unfolding motor can hydrolyze ATP and interact with substrates and ClpP, but several lines of evidence indicate that subunits communicate and coordinate their hydrolysis activity rather than acting independently. For example, the rate of ATP hydrolysis by ClpX is positively cooperative, indicating that ATP binding and hydrolysis in one subunit increase the likelihood of binding and/or hydrolysis in other subunits (Figure 1.7A) (Hersch et al., 2005). Evidence for intersubunit coordination also can be seen directly in the motor's mechanical activity. In single-molecule optical-trapping studies of ClpXP translocation along unfolded substrates, the motor takes steps ranging from 1–4 nm in size (Figure 1.7B) (Cordova et al., 2014; Sen et al., 2013). Because the distance from the highest to lowest pore-1 loop Tyr in solved cryo-EM structures of AAA+ proteases is approximately 2 nm (Figure 1.5B) (Puchades et al., 2017), steps as large as 4 nm only appear possible through the activity of multiple subunits in a coordinated burst of activity. The ClpAP motor also takes translocation steps larger than its 2 nm spiral rise (Olivares et al., 2017). These bursts of mechanical activity are proposed to result either from fast and coordinated ATP-hydrolysis events in multiple subunits or from a triggered release of accumulated strain from multiple slower ATP-hydrolysis events (Cordova et al., 2014; Sen et al., 2013). There is also some evidence for coordinated mechanical action between nonadjacent subunits, as ClpX hexamers in which ATP can only be hydrolyzed in two of six subunits positioned on opposite sides of the ring still take translocation steps as large as 3 nm, despite longer average dwell times between steps (Figure 1.7C) (Cordova et al., 2014).

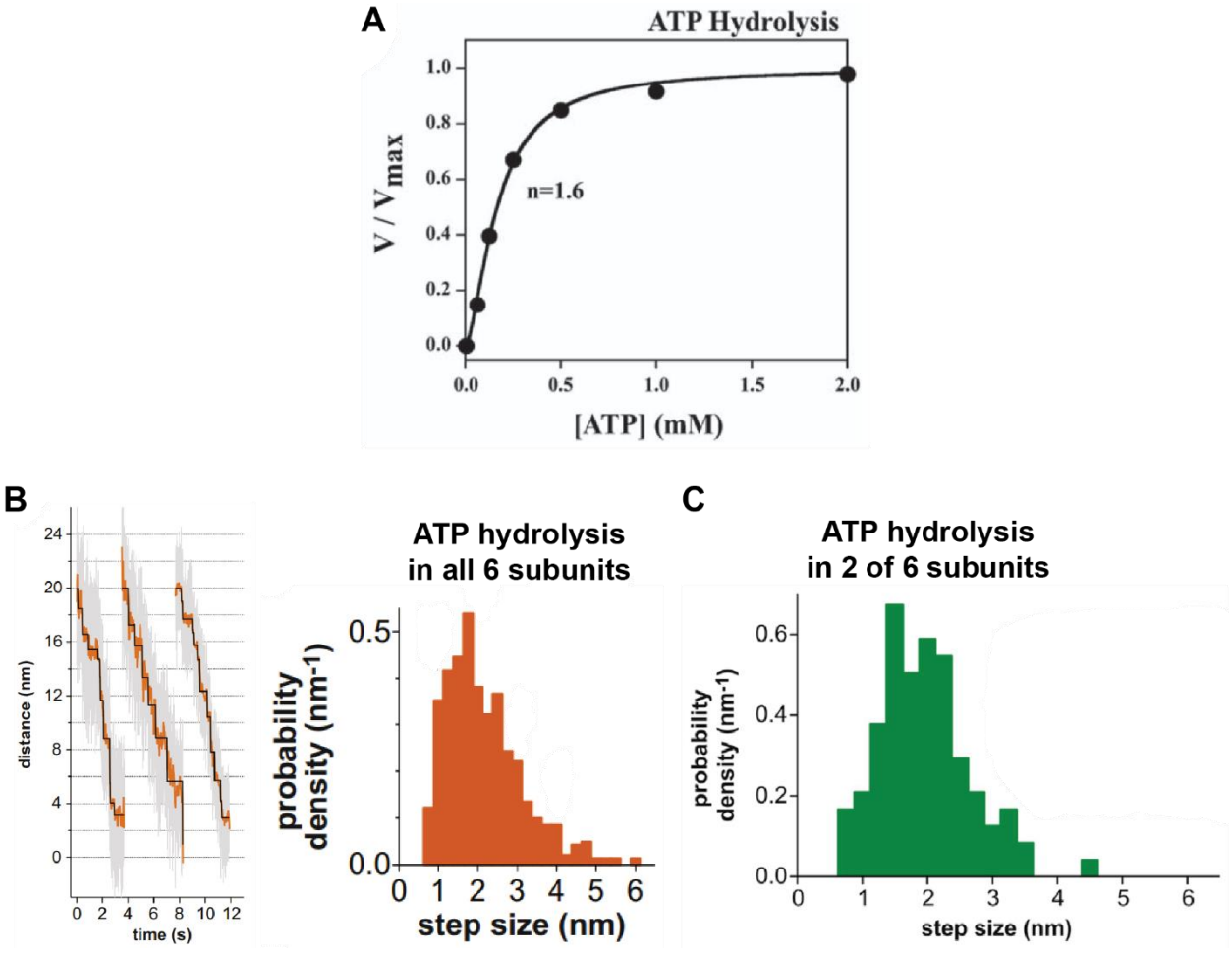


Figure 1.7 – ClpX subunits communicate and work cooperatively

(A) ATP hydrolysis by ClpX is positively cooperative (Hill coefficient n > 1). Adapted from Hersch et al., 2005. **(B)** Single-molecule optical trap measurements of individual translocation steps (left) and the distribution of translocation step sizes (right) by wild-type ClpX$^{\Delta N}$P. Adapted from Cordova et al., 2014. **(C)** Distribution of translocation step sizes by ClpX$^{\Delta N}$P with ATP hydrolysis activity in two of six subunits, oriented *para* to each other. Adapted from Cordova et al., 2014.

How do subunits communicate and coordinate their activity into bursts of mechanical activity? Like nearly all AAA+ enzymes, ClpX hexamers assemble through extensive contacts between protomers, and two subunits contribute to each nucleotide-binding site (Hanson and Whiteheart, 2005). Each subunit can sense its own nucleotide binding state through well-characterized AAA+ nucleotide binding site motifs, including Walker A and B loops and the sensor-II helix (Figure 1.8). Trans interactions are formed by the Arg finger, a conserved arginine which extends from one subunit toward the γ phosphate of ATP bound in the neighboring subunit (Erzberger and Berger, 2006). As a result, subunits can directly read out the nucleotide binding state of their neighbors. Several AAA+ proteases, though not ClpXP, also contain an intersubunit signaling motif, a short helix near the nucleotide binding site that coordinates ATP hydrolysis between neighboring subunits (Augustin et al., 2009).

ClpX subunits may also communicate at longer ranges by sensing the global conformational state of the hexameric ring. After hexameric assembly, these interaction interfaces behave as rigid bodies, retaining nearly invariant conformation (Glynn et al., 2012). As a result, conformational dynamics within the ClpX hexamer are achieved through rotation of rigid bodies relative to each other about the hinge-linkers. Increasing or decreasing linker length by a single amino acid severely impairs substrate degradation while permitting ATP hydrolysis at normal or elevated rates (Glynn et al., 2012). Given this sensitivity, the hinge-linkers are promising candidates to transmit strain accumulated from ATP hydrolysis events and mediate long-distance communication (Glynn et al., 2009).

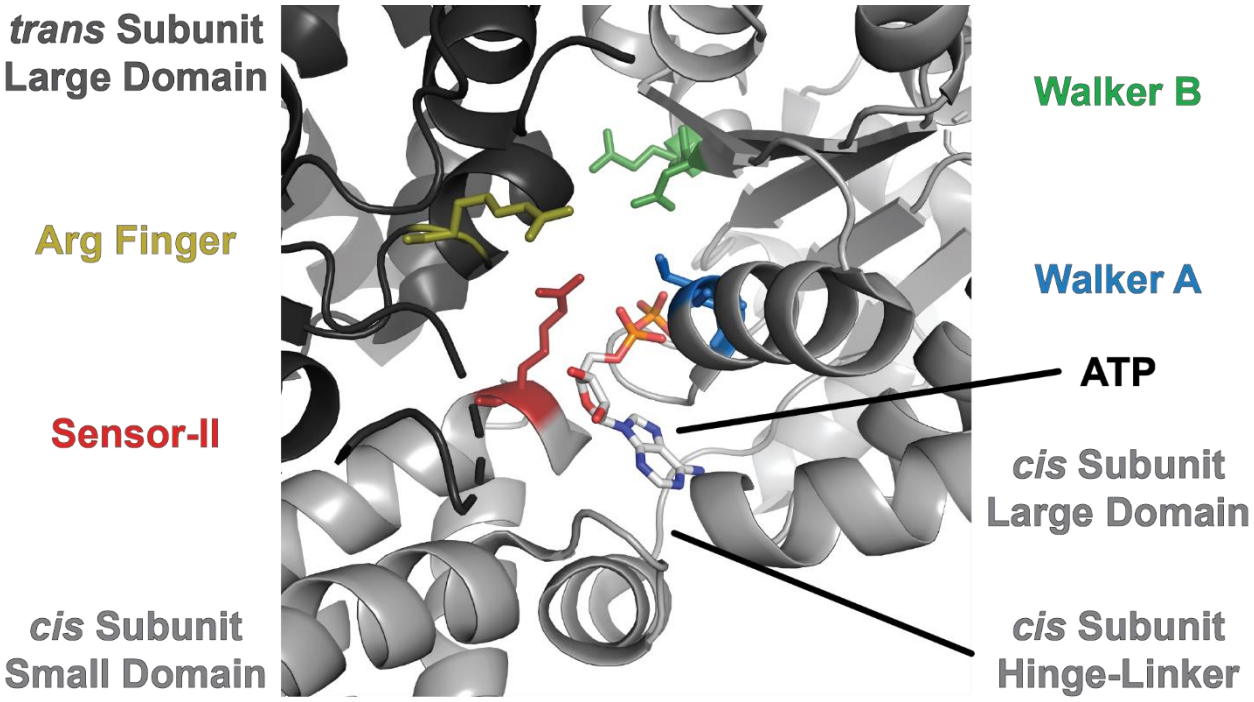


Figure 1.8 – ClpX nucleotide binding interface resides between subunits
Important AAA+ motifs that make contact with bound ATP in the *cis* (Walker A, Walker B, and sensor-II) and *trans* (Arg finger) ClpX protomers. Image generated using the molecular model in PDB 3HWS (Glynn et al., 2009).

Structural models for substrate unfolding and translocation

Recent cryo-EM structures of AAA+ proteases and other protein-remodeling motors in native conformations have demonstrated surprisingly conserved global structure (Cooney et al., 2019; Deville et al., 2017; Ding et al., 2017; Dong et al., 2019; Gates et al., 2017; Han et al., 2017; Lee et al., 2010b; Lo et al., 2019; Majumder et al., 2019; Michalska et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017, 2019; Ripstein et al., 2017; Rizo et al., 2019; Su et al., 2017; Sun et al., 2017; Twomey et al., 2019; White et al., 2018; Yokom et al., 2016; Yu et al., 2018; Zehr et al., 2017; Zhu et al., 2018). The similarity in these structures suggests that AAA+ proteases may employ common mechanistic principles for substrate translocation and unfolding.

In one model for processive substrate translocation, termed the sequential model, the six subunits in the hexamer each sequentially hydrolyze ATP and undergo a conformational change that shifts the subunit downward, pulling the bound substrate further down into the pore (Figure 1.9A) (Enemark and Joshua-Tor, 2006; Puchades et al., 2017). By contrast, a separate model termed the stochastic model posits that any ATP-bound subunit can hydrolyze ATP and translocate the substrate downward (Figure 1.9B) (Cordova et al., 2014; Martin et al., 2005; Sen et al., 2013). Elements of these models can be combined to produce more complex models as well. For example, a sequential model with stochastic elements is possible if the motor normally hydrolyzes ATP cyclically in neighboring subunits, but occasionally resets stochastically through thermally driven motions. Similarly, a stochastic model could have sequential elements if ATP hydrolysis can occur in any subunit but hydrolysis bursts propagate in a coordinated manner between neighboring subunits within the ring.

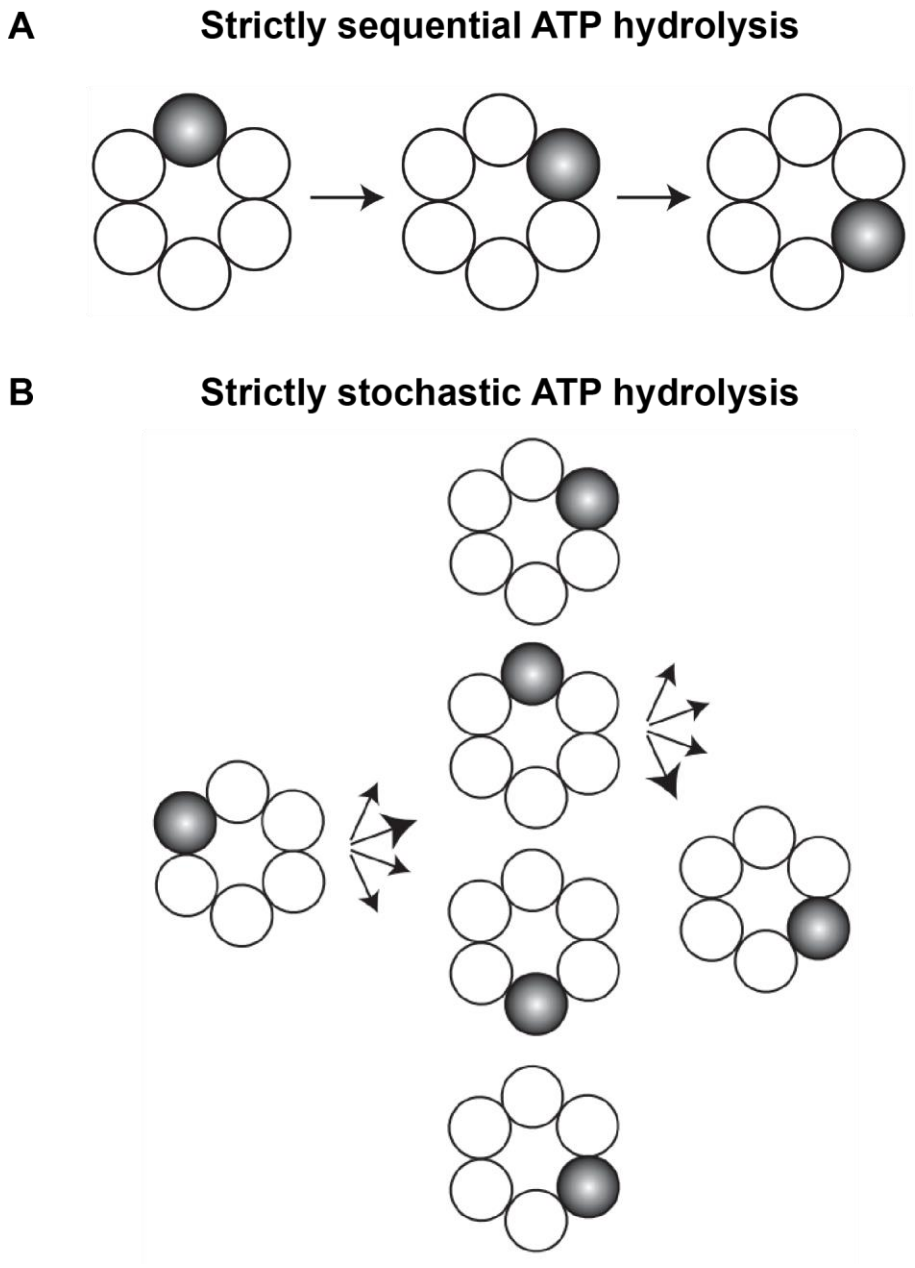


Figure 1.9 – Models for substrate translocation by ClpX
Simplified schemes for translocation by **(A)** a purely sequential hydrolysis model, and **(B)** a purely stochastic hydrolysis model. In both panels, circles represent ClpX subunits, and a darkened subunit indicates an ATP hydrolysis event. Figure adapted from Martin et al., 2005.

In dissecting the mechanism of substrate translocation by ClpXP, both the sequential and stochastic models present strengths and weaknesses. From a structural perspective, the asymmetry in almost all cryo-EM structures of AAA+ machines suggests a sequential mechanism. Every subunit in the AAA+ hexamer has a defined vertical position within the spiral hierarchy, providing a basis for different subunits to be differentially optimized for ATP hydrolysis and subunit motion. Conformational changes driven by hydrolysis in one subunit could then sequentially reconfigure each subunit into a new position, driving processive rotary motion. This model is also favored by the general assembly structure of AAA+ machines, as *trans* interactions between the nucleotide binding site of one subunit and the Arg finger of its neighbor theoretically make each subunit most sensitive to the nucleotide hydrolysis state of its direct neighbor. However, a substantial body of biochemical data suggest that ClpX cannot operate by a purely sequential mechanism. Most strikingly, eliminating ATP hydrolysis in one subunit within each hexamer, which a sequential mechanism predicts should cause the motor to stall, does not substantially impair unfolding or translocation activity (Martin et al., 2005). Likewise, hexamers with only two of six subunits competent to hydrolyze ATP take similar-sized translocation steps as wild-type ClpX (Cordova et al., 2014).

Other AAA+ motors may also operate outside of the strict constraints of the sequential model. For example, eliminating ATP hydrolysis abolishes motor activity only in three of the six subunits of the proteasomal Rpt1-6 motor (Beckwith et al., 2013). In Yta10/Yta12 (a AAA+ motor which normally functions sequentially), various mutations break strict coupling of hydrolysis without impairing mechanical activity, demonstrating that sequential hydrolysis is not required for mechanical function (Augustin et al., 2009). Despite the recent wealth of structural information on

many AAA+ protein remodelers, the mechanistic details of substrate grip, unfolding, and translocation by ClpX and other AAA+ proteases remain enigmatic.

Thesis overview

In the following chapters of this thesis, I describe a series of biochemical and structural experiments that elucidate mechanistic details of intersubunit communication, substrate grip, and substrate translocation by ClpXP. In Chapter II, I explore the role of the ClpX hinge-linkers in facilitating intersubunit communication. I find that the hinge-linkers mediate communication both by enforcing closed-ring topology within a strained ring and by allowing neighboring subunits to transmit information about their conformational state to each other. In Chapter III, I identify the determinants of substrate grip during unfolding and identify the features that allow some amino-acid sequences to protect adjacent folded domains from degradation. I demonstrate that the ClpX pore-1 loops contribute asymmetrically to substrate grip, and that grip is primarily mediated through van der Waal's interactions between the pore-1 loop and hydrophobic or aromatic substrate residues. In Chapter IV, I present a collaborative project that solved a series of substrate-bound cryo-EM structures of ClpXP. I explore and validate previously unseen elements of ClpXP through biochemical characterization and identify new roles for the RKH loops, IGF loops, and several residues that interact with bound nucleotide. In Chapter V, I present results from crosslinking of a single ClpX-IGF loop to the surface of ClpP. This crosslinked species prevents rotation of ClpX relative to ClpP, and biochemical characterization elucidates new details about the mechanism of substrate unfolding and translocation. In Chapter VI, I contextualize these findings into the current knowledge of ClpXP mechanism, and present outstanding questions that will need to be answered to advance the field toward a more complete understanding of AAA+

protease function. Finally, in two Appendices, I present unpublished experiments related to the studies of substrate grip in Chapter III. These include measurements of substrate-dissociation kinetics during failed unfolding and a computational approach that locates amino-acid sequences that are predicted to impair grip by AAA+ proteases and impart protection on the folded domains they flank within several bacterial and eukaryotic proteomes.

REFERENCES

Adam, Z., Adamska, I., Nakabayashi, K., Ostersetzer, O., Haussuhl, K., Manuell, A., Zheng, B., Vallon, O., Rodermel, S.R., Shinozaki, K., et al. (2001). Chloroplast and Mitochondrial Proteases in Arabidopsis. A Proposed Nomenclature. Plant Physiol. *125*, 1912–1918.

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.

Augustin, S., Gerdes, F., Lee, S., Tsai, F.T.F., Langer, T., and Tatsuta, T. (2009). An Intersubunit Signaling Network Coordinates ATP Hydrolysis by m-AAA Proteases. Mol. Cell *35*, 574–585.

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

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

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.

Barkow, S.R., Levchenko, I., Baker, T.A., and Sauer, R.T. (2009). Polypeptide Translocation by the AAA+ ClpXP Protease Machine. Chem. Biol. *16*, 605–612.

Bar-Nun, S., and Glickman, M.H. (2012). Proteasomal AAA-ATPases: Structure and function. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1823*, 67–82.

Barthelme, D., and Sauer, R.T. (2012). Identification of the Cdc48•20S Proteasome as an Ancient AAA+ Proteolytic Machine. Science *337*, 843–846.

Barthelme, D., Jauregui, R., and Sauer, R.T. (2015). An ALS disease mutation in Cdc48/p97 impairs 20S proteasome binding and proteolytic communication. Protein Sci. *24*, 1521–1527.

Beckwith, R., Estrin, E., Worden, E.J., and Martin, A. (2013). Reconstitution of the 26S proteasome reveals functional asymmetries in its AAA+ unfoldase. Nat. Struct. Mol. Biol. *20*, 1164–1172.

Belle, A., Tanay, A., Bitincka, L., Shamir, R., and O'Shea, E.K. (2006). Quantification of protein half-lives in the budding yeast proteome. Proc. Natl. Acad. Sci. *103*, 13004–13009.

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*, 113–128.

Bochtler, M., Ditzel, L., Groll, M., and Huber, R. (1997). Crystal structure of heat shock locus V (HslV) from Escherichia coli. Proc. Natl. Acad. Sci. *94*, 6070–6074.

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

Bolon, D.N., Grant, R.A., Baker, T.A., and Sauer, R.T. (2004). Nucleotide-Dependent Substrate Handoff from the SspB Adaptor to the AAA+ ClpXP Protease. Mol. Cell *16*, 343–350.

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*, 1082–1087.

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

Burton, R.E., Baker, T.A., and Sauer, R.T. (2005). Nucleotide-dependent substrate recognition by the AAA+ HslUV protease. Nat. Struct. Mol. Biol. *12*, 245–251.

Cambridge, S.B., Gnad, F., Nguyen, C., Bermejo, J.L., Krüger, M., and Mann, M. (2011). Systems-wide Proteomic Analysis in Mammalian Cells Reveals Conserved, Functional Protein Turnover. J. Proteome Res. *10*, 5275–5284.

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*, 502–505.

Cordova, J.C., Olivares, A.O., Shin, Y., Stinson, B.M., Calmat, S., Schmitz, K.R., Aubin-Tam, M.-E., Baker, T.A., Lang, M.J., and Sauer, R.T. (2014). Stochastic but Highly Coordinated Protein Unfolding and Translocation by the ClpXP Proteolytic Machine. Cell *158*, 647–658.

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*, 30090–30100.

Del Dotto, V., Fogazza, M., Carelli, V., Rugolo, M., and Zanna, C. (2018). Eight human OPA1 isoforms, long and short: What are they for? Biochim. Biophys. Acta BBA - Bioenerg. *1859*, 263–269.

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*, e1701726.

Ding, Z., Fu, Z., Xu, C., Wang, Y., Wang, Y., Li, J., Kong, L., Chen, J., Li, N., Zhang, R., et al. (2017). High-resolution cryo-EM structure of the proteasome in complex with ADP-AlFx. Cell Res. *27*, 373–385.

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.

Dulebohn, D., Choy, J., Sundermeier, T., Okan, N., and Karzai, A.W. (2007). Trans-Translation: The tmRNA-Mediated Surveillance Mechanism for Ribosome Rescue, Directed Protein Degradation, and Nonstop mRNA Decay. Biochemistry *46*, 4681–4693.

van Dyck, L., Dembowski, M., Neupert, W., and Langer, T. (1998). Mcx1p, a ClpX homologue in mitochondria of Saccharomyces cerevisiae. FEBS Lett. *438*, 250–254.

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*, 14834–14840.

Enemark, E.J., and Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. Nature *442*, 270–275.

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*, 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*, 161–166.

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*, 671–683.

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*, 273–279.

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.

Giudice, E., and Gillet, R. (2013). The task force that rescues stalled ribosomes in bacteria. Trends Biochem. Sci. *38*, 403–411.

Glynn, S.E. (2017). Multifunctional Mitochondrial AAA Proteases. Front. Mol. Biosci. *4*.

Glynn, S.E., Martin, A., Nager, A.R., Baker, T.A., and Sauer, R.T. (2009). Structures of Asymmetric ClpX Hexamers Reveal Nucleotide-Dependent Motions in a AAA+ Protein-Unfolding Machine. Cell *139*, 744–756.

Glynn, S.E., Nager, A.R., Baker, T.A., and Sauer, R.T. (2012). Dynamic and static components power unfolding in topologically closed rings of a AAA+ proteolytic machine. Nat. Struct. Mol. Biol. *19*, 616–622.

Grimaud, R., Kessel, M., Beuron, F., Steven, A.C., and Maurizi, M.R. (1998). Enzymatic and Structural Similarities between theEscherichia coli ATP-dependent Proteases, ClpXP and ClpAP. J. Biol. Chem. *273*, 12476–12481.

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.

Hanson, P.I., and Whiteheart, S.W. (2005). AAA+ proteins: have engine, will work. Nat. Rev. Mol. Cell Biol. *6*, 519–529.

Hari, S.B., and Sauer, R.T. (2016). The AAA+ FtsH Protease Degrades an ssrA-Tagged Model Protein in the Inner Membrane of Escherichia coli. Biochemistry *55*, 5649–5652.

Hersch, G.L., Burton, R.E., Bolon, D.N., Baker, T.A., and Sauer, R.T. (2005). Asymmetric Interactions of ATP with the AAA+ ClpX6 Unfoldase: Allosteric Control of a Protein Machine. Cell *121*, 1017–1027.

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*, 1720–1729.

Imkamp, F., Ziemski, M., and Weber-Ban, E. (2015). Pupylation-dependent and -independent proteasomal degradation in mycobacteria. Biomol. Concepts *6*, 285–301.

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

Iosefson, O., Olivares, A.O., Baker, T.A., and Sauer, R.T. (2015b). Dissection of Axial-Pore Loop Function during Unfolding and Translocation by a AAA+ Proteolytic Machine. Cell Rep. *12*, 1032–1041.

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.

Kardon, J.R., Yien, Y.Y., Huston, N.C., Branco, D.S., Hildick-Smith, G.J., Rhee, K.Y., Paw, B.H., and Baker, T.A. (2015). Mitochondrial ClpX Activates a Key Enzyme for Heme Biosynthesis and Erythropoiesis. Cell *161*, 858–867.

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*, 449–455.

Keiler, K.C., and Feaga, H.A. (2014). Resolving Nonstop Translation Complexes Is a Matter of Life or Death. J. Bacteriol. *196*, 2123–2130.

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*, 990–993.

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.

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. *102*, 1390–1395.

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/Hsp100 ATPases and the ClpP peptidase. Nat. Struct. Biol. *8*, 230–233.

Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K., and Inokuchi, H. (1994). A tRNA-like structure is present in 10Sa RNA, a small stable RNA from Escherichia coli. Proc. Natl. Acad. Sci. *91*, 9223–9227.

Koren, I., Timms, R.T., Kula, T., Xu, Q., Li, M.Z., and Elledge, S.J. (2018). The Eukaryotic Proteome Is Shaped by E3 Ubiquitin Ligases Targeting C-Terminal Degrons. Cell *173*, 1622-1635.e14.

Kraut, D.A. (2013). Slippery Substrates Impair ATP-dependent Protease Function by Slowing Unfolding. J. Biol. Chem. *288*, 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*, 1444–1453.

Langklotz, S., Baumann, U., and Narberhaus, F. (2012). Structure and function of the bacterial AAA protease FtsH. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1823*, 40–48.

Lau, J., Hernandez-Alicea, L., Vass, R.H., and Chien, P. (2015). A Phosphosignaling Adaptor Primes the AAA+ Protease ClpXP to Drive Cell Cycle-Regulated Proteolysis. Mol. Cell *59*, 104–116.

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*, 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*, 707–718.

Levchenko, I., Seidel, M., Sauer, R.T., and Baker, T.A. (2000). A Specificity-Enhancing Factor for the ClpXP Degradation Machine. Science *289*, 2354–2356.

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. *94*, 12616–12621.

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*, 959–969.

Lien, H.-Y., Yu, C.-H., Liou, C.-M., and Wu, W.F. (2009). Regulation of clpQ+Y+ (hslV+U+) Gene Expression in Escherichia coli. Open Microbiol. J. *3*, 29–39.

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*, 2248–2254.

Liu, K., Ologbenla, A., and Houry, W.A. (2014). Dynamics of the ClpP serine protease: A model for self-compartmentalized proteases. Crit. Rev. Biochem. Mol. Biol. *49*, 400–412.

Lo, Y.-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*, 1–12.

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*, 459–469.

Majumder, P., Rudack, T., Beck, F., Danev, R., Pfeifer, G., Nagy, I., and Baumeister, W. (2019). Cryo-EM structures of the archaeal PAN-proteasome reveal an around-the-ring ATPase cycle. Proc. Natl. Acad. Sci. *116*, 534–539.

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

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.

Martin, A., Baker, T.A., and Sauer, R.T. (2008a). Diverse Pore Loops of the AAA+ ClpX Machine Mediate Unassisted and Adaptor-Dependent Recognition of ssrA-Tagged Substrates. Mol. Cell *29*, 441–450.

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

Maurizi, M.R., Clark, W.P., Katayama, Y., Rudikoff, S., Pumphrey, J., Bowers, B., and Gottesman, S. (1990). Sequence and structure of Clp P, the proteolytic component of the ATP-dependent Clp protease of Escherichia coli. J. Biol. Chem. *265*, 12536–12545.

Michalska, K., Zhang, K., March, Z.M., Hatzos-Skintges, C., Pintilie, G., Bigelow, L., Castellano, L.M., Miles, L.J., Jackrel, M.E., Chuang, E., et al. (2019). Structure of

Calcarisporiella thermophila Hsp104 Disaggregase that Antagonizes Diverse Proteotoxic Misfolding Events. Structure *27*, 449-463.e7.

Monroe, N., Han, H., Shen, P.S., Sundquist, W.I., and Hill, C.P. (2017). Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. ELife *6*, e24487.

Oh, E., Akopian, D., and Rape, M. (2018). Principles of Ubiquitin-Dependent Signaling. Annu. Rev. Cell Dev. Biol. *34*, 137–162.

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*, 33–44.

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.

Ostersetzer, O., Kato, Y., Adam, Z., and Sakamoto, W. (2007). Multiple Intracellular Locations of Lon Protease in Arabidopsis: Evidence for the Localization of AtLon4 to Chloroplasts. Plant Cell Physiol. *48*, 881–885.

Pearce, M.J., Mintseris, J., Ferreyra, J., Gygi, S.P., and Darwin, K.H. (2008). Ubiquitin-Like Protein Involved in the Proteasome Pathway of Mycobacterium tuberculosis. Science *322*, 1104–1107.

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*, eaav0725.

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.

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*, 1073-1085.e6.

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.

Rizo, A.N., Lin, J., 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*, 1–12.

Rodriguez-Aliaga, P., Ramirez, L., Kim, F., Bustamante, C., and Martin, A. (2016). Substrate-translocating loops regulate mechanochemical coupling and power production in AAA+ protease ClpXP. Nat. Struct. Mol. Biol. *23*, 974–981.

Röttgers, K., Zufall, N., Guiard, B., and Voos, W. (2002). The ClpB Homolog Hsp78 Is Required for the Efficient Degradation of Proteins in the Mitochondrial Matrix. J. Biol. Chem. *277*, 45829–45837.

Sauer, R.T., and Baker, T.A. (2011). AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annu. Rev. Biochem. *80*, 587–612.

Schmidt, R., Bukau, B., and Mogk, A. (2009). Principles of general and regulatory proteolysis by AAA+ proteases in Escherichia coli. Res. Microbiol. *160*, 629–636.

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*, 636–646.

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*, 25976–25982.

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*, 137–142.

Siddiqui, S.M., Sauer, R.T., and Baker, T.A. (2004). Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. Genes Dev. *18*, 369–374.

Smith, D.M., Chang, S.-C., Park, S., Finley, D., Cheng, Y., and Goldberg, A.L. (2007). Docking of the Proteasomal ATPases' Carboxyl Termini in the 20S Proteasome's α Ring Opens the Gate for Substrate Entry. Mol. Cell *27*, 731–744.

Smith, S.C., Joshi, K.K., Zik, J.J., Trinh, K., Kamajaya, A., Chien, P., and Ryan, K.R. (2014). Cell cycle-dependent adaptor complex for ClpXP-mediated proteolysis directly integrates phosphorylation and second messenger signals. Proc. Natl. Acad. Sci. *111*, 14229–14234.

Snider, J., Thibault, G., and Houry, W.A. (2008). The AAA+ superfamily of functionally diverse proteins. Genome Biol. *9*, 216.

Stinson, B.M., Nager, A.R., Glynn, S.E., Schmitz, K.R., Baker, T.A., and Sauer, R.T. (2013). Nucleotide Binding and Conformational Switching in the Hexameric Ring of a AAA+ Machine. Cell *153*, 628–639.

Su, M., Guo, E.Z., Ding, X., Li, Y., Tarrasch, J.T., Brooks, C.L., Xu, Z., and Skiniotis, G. (2017). Mechanism of Vps4 hexamer function revealed by cryo-EM. Sci. Adv. *3*, e1700325.

Sun, S., Li, L., Yang, F., Wang, X., Fan, F., Yang, M., Chen, C., Li, X., Wang, H.-W., and Sui, S.-F. (2017). Cryo-EM structures of the ATP-bound Vps4 E233Q hexamer and its complex with Vta1 at near-atomic resolution. Nat. Commun. *8*, 1–13.

Sundar, S., McGinness, K.E., Baker, T.A., and Sauer, R.T. (2010). Multiple Sequence Signals Direct Recognition and Degradation of Protein Substrates by the AAA+ Protease HslUV. J. Mol. Biol. *403*, 420–429.

Tian, L., Holmgren, R.A., and Matouschek, A. (2005). A conserved processing mechanism regulates the activity of transcription factors Cubitus interruptus and NF-κB. Nat. Struct. Mol. Biol. *12*, 1045–1053.

Tomko, R.J., and Hochstrasser, M. (2013). Molecular Architecture and Assembly of the Eukaryotic Proteasome. Annu. Rev. Biochem. *82*, 415–445.

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*, 13243–13257.

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*, 48–53.

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*, eaax1033.

Ushida, C., Himeno, H., Watanabe, T., and Muto, A. (1994). tRNA-like structures in 10Sa RNAs of Mycoplasma capricolum and Bacillus subtilis. Nucleic Acids Res. *22*, 3392–3396.

Van Melderen, L., and Aertsen, A. (2009). Regulation and quality control by Lon-dependent proteolysis. Res. Microbiol. *160*, 645–651.

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

Vass, R.H., Nascembeni, J., and Chien, P. (2017). The Essential Role of ClpXP in Caulobacter crescentus Requires Species Constrained Substrate Specificity. Front. Mol. Biosci. *4*.

Wah, D.A., Levchenko, I., Baker, T.A., and Sauer, R.T. (2002). Characterization of a Specificity Factor for an AAA+ ATPase: Assembly of SspB Dimers with ssrA-Tagged Proteins and the ClpX Hexamer. Chem. Biol. *9*, 1237–1245.

Wang, J., Hartling, J.A., and Flanagan, J.M. (1997). The Structure of ClpP at 2.3 Å Resolution Suggests a Model for ATP-Dependent Proteolysis. Cell *91*, 447–456.

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.

Wickner, S., Gottesman, S., Skowyra, D., Hoskins, J., McKenney, K., and Maurizi, M.R. (1994). A molecular chaperone, ClpA, functions like DnaK and DnaJ. Proc. Natl. Acad. Sci. *91*, 12218–12222.

Ye, Y., Tang, W.K., Zhang, T., and Xia, D. (2017). A Mighty "Protein Extractor" of the Cell: Structure and Function of the p97/CDC48 ATPase. Front. Mol. Biosci. *4*.

Yien, Y.Y., Ducamp, S., van der Vorm, L.N., Kardon, J.R., Manceau, H., Kannengiesser, C., Bergonia, H.A., Kafina, M.D., Karim, Z., Gouya, L., et al. (2017). Mutation in human CLPX elevates levels of δ-aminolevulinate synthase and protoporphyrin IX to promote erythropoietic protoporphyria. Proc. Natl. Acad. Sci. U. S. A. *114*, E8045–E8052.

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*, 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. *115*, E9560–E9569.

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*, 717–725.

Zhang, M., and Coffino, P. (2004). Repeat Sequence of Epstein-Barr Virus-encoded Nuclear Antigen 1 Protein Interrupts Proteasome Substrate Processing. J. Biol. Chem. *279*, 8635–8641.

Zheng, N., and Shabek, N. (2017). Ubiquitin Ligases: Structure, Function, and Regulation. Annu. Rev. Biochem. *86*, 129–157.

Zhu, Y., Wang, W.L., Yu, D., Ouyang, Q., Lu, Y., and Mao, Y. (2018). Structural mechanism for nucleotide-driven remodeling of the AAA-ATPase unfoldase in the activated human 26S proteasome. Nat. Commun. *9*, 1–12.

Chapter II

Hinge-linker elements in the AAA+ protein unfoldase ClpX mediate intersubunit communication, assembly, and mechanical activity

This chapter is published:

Bell, Tristan A., Baker, Tania A. and Sauer, Robert T. (2018) Hinge-linker elements in the AAA+ protein unfoldase ClpX mediate intersubunit communication, assembly, and mechanical activity. Biochemistry *57*, 6787–6796.

T.A. Bell performed experiments. T.A. Bell, T.A. Baker, and R.T.S. designed experiments and analyzed data. T.A. Bell and R.T.S. wrote the manuscript.

ABSTRACT

The ClpXP protease plays important roles in protein homeostasis and quality control. ClpX is a ring-shaped AAA+ homohexamer that unfolds target proteins and translocates them into the ClpP peptidase for degradation. AAA+ modules in each ClpX subunit – consisting of a large AAA+ domain, a short hinge-linker, and a small AAA+ domain – mediate the mechanical activities of the ring hexamer. Here, we investigate the roles of these hinge-linkers in ClpX function. Deleting one hinge-linker in a single-chain ClpX pseudohexamer dramatically decreases unfolding and degradation activity, in part by compromising formation of closed rings, protein-substrate binding, and ClpP binding. Covalently re-closing the broken hinge-linker interface rescues activity. Deleting one hinge-linker from a single-chain dimer or trimer prevents assembly of stable hexamers. Mutationally disrupting a hinge-linker preserves closed ring assembly but reduces ATP-hydrolysis cooperativity and degradation activity. These results indicate that hinge-linker length and flexibility are optimized for efficient substrate unfolding and support a model in which the hinge-linkers of ClpX facilitate efficient degradation both by maintaining proper ring geometry and facilitating subunit-subunit communication. This model informs our understanding of ClpX as well as the larger AAA+ family of motor proteins, which play diverse roles in converting chemical into mechanical energy in all cells.

INTRODUCTION

All organisms require the ability to convert chemical energy into mechanical work. Members of the AAA+ protein family (ATPases associated with various cellular activities) use ATP hydrolysis to power many cellular mechanical processes, including protein unfolding and disaggregation (Erzberger and Berger, 2006). Typical family members assemble into ring- or spiral-shaped structures, often hexamers, which are stabilized by interactions between the large AAA+ domain of one subunit and the small AAA+ domain of its neighbor (Hanson and Whiteheart, 2005). A hinge-linker connects the large and small AAA+ domains of each subunit (Figure 2.1A).

ClpX is a hexameric AAA+ protein unfoldase that is present in most bacteria and some eukaryotic organelles. It unfolds specific proteins, including those that are mistranslated or damaged, and feeds the unfolded polypeptide into an internal chamber of the associated ClpP peptidase for proteolysis (Sauer and Baker, 2011). Recognition of appropriate substrates is mediated by sequence motifs (degrons) that target proteins to ClpX either directly or with the help of adapter proteins. In *Escherichia coli*, the 11-residue ssrA degron is appended to the C-terminus of partially synthesized proteins on stalled ribosomes, and ClpXP or other cellular proteases degrade the resulting ssrA-tagged polypeptides (Choy et al., 2007; Gottesman et al., 1998; Hari and Sauer, 2016; Keiler et al., 1996). The ssrA tag initially binds to loops within the axial pore of the hexameric ClpX ring (Martin et al., 2008; Siddiqui et al., 2004). Conformational changes in the pore – driven by ATP binding, hydrolysis, and product release – then pull on the tag, eventually resulting in unfolding and step-wise translocation into ClpP for degradation (Aubin-Tam et al., 2011; Cordova et al., 2014; Maillard et al., 2011; Sen et al., 2013).

In crystal structures, the hexameric ClpX ring is stabilized by relatively invariant rigid-body interactions between the large and small AAA+ domains of neighboring subunits (Glynn et al., 2009; Stinson et al., 2013) (Figure 2.1A). Moreover, covalent crosslinks across these rigid-body interfaces do not compromise ClpXP degradation activity, indicating that these contacts are maintained during the conformational rearrangements in the hexamer that drive substrate unfolding and translocation (Glynn et al., 2012). Consequently, the enzyme motions responsible for mechanical work are proposed to arise from changes in the orientations of neighboring rigid bodies, which are mediated in turn by structural changes in the four-residue linkers that connect the large and small AAA+ domains of each subunit and function as hinges between successive rigid bodies (Glynn et al., 2009, 2012; Stinson et al., 2013).

Mutant ClpX rings with only a single subunit capable of hydrolyzing ATP support low levels of ClpP degradation (Martin et al., 2005), indicating that unfolding and translocation do not require sequential or concerted ATP hydrolysis. Although ClpX hydrolyzes many molecules of ATP while attempting to unfold a stable protein substrate, cooperative unfolding ultimately occurs as a consequence of a power stroke driven by a single hydrolysis event (Aubin-Tam et al., 2011; Cordova et al., 2014; Kenniston et al., 2003; Maillard et al., 2011; Olivares et al., 2017; Sen et al., 2013). Robust unfolding requires coordination between the axial-pore loops of multiple ClpX subunits (Iosefson et al., 2015a, 2015b; Rodriguez-Aliaga et al., 2016). Moreover, ATP hydrolysis by ClpX is positively cooperative, implying some means of subunit-subunit communication (Nager et al., 2011). Evidence of intersubunit communication is also observed in the enzyme's mechanical activity. Although the smallest translocation steps taken by ClpXP are ~1 nm in length, kinetic bursts of ATP hydrolysis appear to be responsible for steps of approximately 2, 3, and 4

nm (Cordova et al., 2014; Sen et al., 2013). Such communication could arise from contacts between neighboring large AAA+ domains and/or through the hinge-linkers that control the orientation of neighboring rigid bodies.

The hinge-linkers of ClpX form part of the ATP binding site. In *E. coli* ClpX, a conserved hinge-linker residue (L317) contacts nucleotide in crystal structures (Glynn et al., 2009). The sequence of the rest of the four-residue hinge-linker is poorly conserved but its length is important, as insertion or deletion of one residue in all six hinge-linkers of a hexamer severely impairs unfolding activity but not ATP hydrolysis (Glynn et al., 2012). To better understand the roles of the ClpX hinge-linkers in intersubunit communication during substrate unfolding, we have used circular permutation to delete one or more hinge-linkers and engineered linker-disruption and ATP-hydrolysis mutations to interrogate how the hinge-linkers affect ClpX function. Our results support a model in which the hinge-linkers facilitate communication between subunits both by enforcing closed-ring topology and maintaining proper subunit-subunit geometry within the ring during ATP hydrolysis.

RESULTS

Mutating or deleting a single hinge-linker impairs intersubunit communication

To determine whether a ClpX hexamer requires six hinge-linkers for robust function, we used circular permutation to create a single-chain pseudohexamer missing one hinge-linker (Figure 2.1B). Prior studies show that single-chain pseudohexamers of $ClpX^{\Delta N}$, in which the wild-type N-terminal domain is deleted, have activities similar to wild-type ClpX hexamers in supporting degradation of ssrA-tagged substrates by ClpP (Martin et al., 2005). As shown in Figure 2.1B,

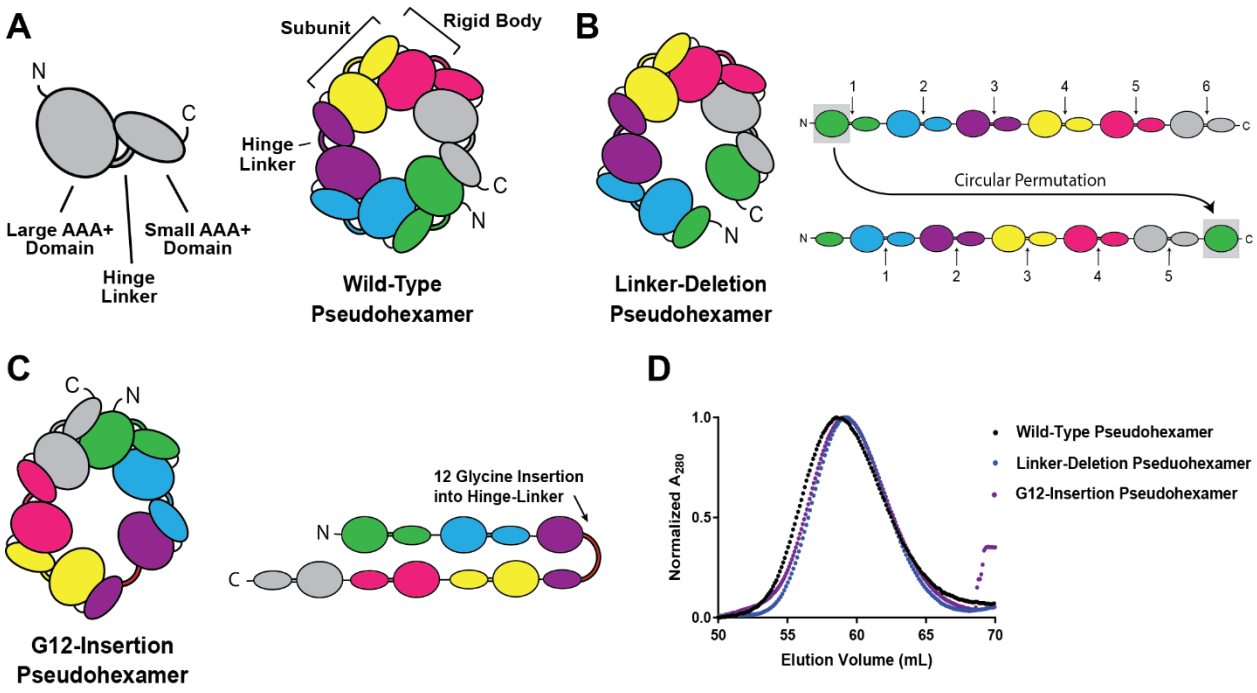


Figure 2.1 – Hinge-linker deletion and mutation in a single-chain ClpX$^{\Delta N}$ hexamer
(A) Cartoon representation of a single subunit of ClpX$^{\Delta N}$ (left) and an assembled single-chain hexamer (right). **(B)** Cartoon representation of the linker-deletion pseudohexamer, which was constructed through circular permutation by moving the N-terminal large AAA+ domain of wild-type pseudohexamer to the C-terminus. **(C)** Cartoon representation of the G12-insertion pseudohexamer, which was constructed by adding 12 glycine residues into one hinge-linker of wild-type pseudohexamer. **(D)** Size exclusion chromatograms of single-chain pseudohexamer variants.

moving the large domain of the N-terminal subunit of a single-chain pseudohexamer to the end of the C-terminal small domain results in a linker-deletion variant missing one hinge-linker. To compare the effects of deleting versus mutationally disrupting a hinge-linker, we also constructed a non-permuted pseudohexamer in which 12 glycine residues were inserted into a single hinge-linker (G12 insertion, Figure 2.1C). During purification, both the linker-deletion and G12-insertion variants eluted from a size-exclusion column at positions expected for hexamers (Figure 2.1D), establishing that these hinge-linker mutations do not cause multimeric aggregation.

ATP hydrolysis by ClpX is positively cooperative (Hersch et al., 2005). We measured the ATP dependence of hydrolysis for the parental pseudohexamer and the linker-deletion and G12-insertion variants and fit the resulting data to the Hill form of the Michaelis-Menten equation (Figure 2.2A, Table 2.1). Compared to the parental pseudohexamer, both linker-disruption mutants had ~3-fold higher V_{max} values and ~5-7-fold higher apparent K_M values. The Hill constant was ~2.0 for the parent, ~1.6 for the linker-deletion variant, and ~1.1 for the G12-insertion enzyme (Figure 2.2A, Table 2.1). Hence, the hinge-linker disruptions affect many aspects of steady-state ATP hydrolysis, including communication between subunits that results in positive cooperativity. Addition of ClpP repressed ATP hydrolysis by all three variants, although this effect was only observed in the linker-deletion variant at high concentrations of ClpP (Figures 2.2B, 2.3). Addition of an Arc-st11-ssrA protein substrate and ClpP stimulated ATP hydrolysis by the parental and G12-insertion enzymes but had little effect on the linker-deletion variant, even at high ClpP and substrate concentrations (Figures 2.2B, 2.3).

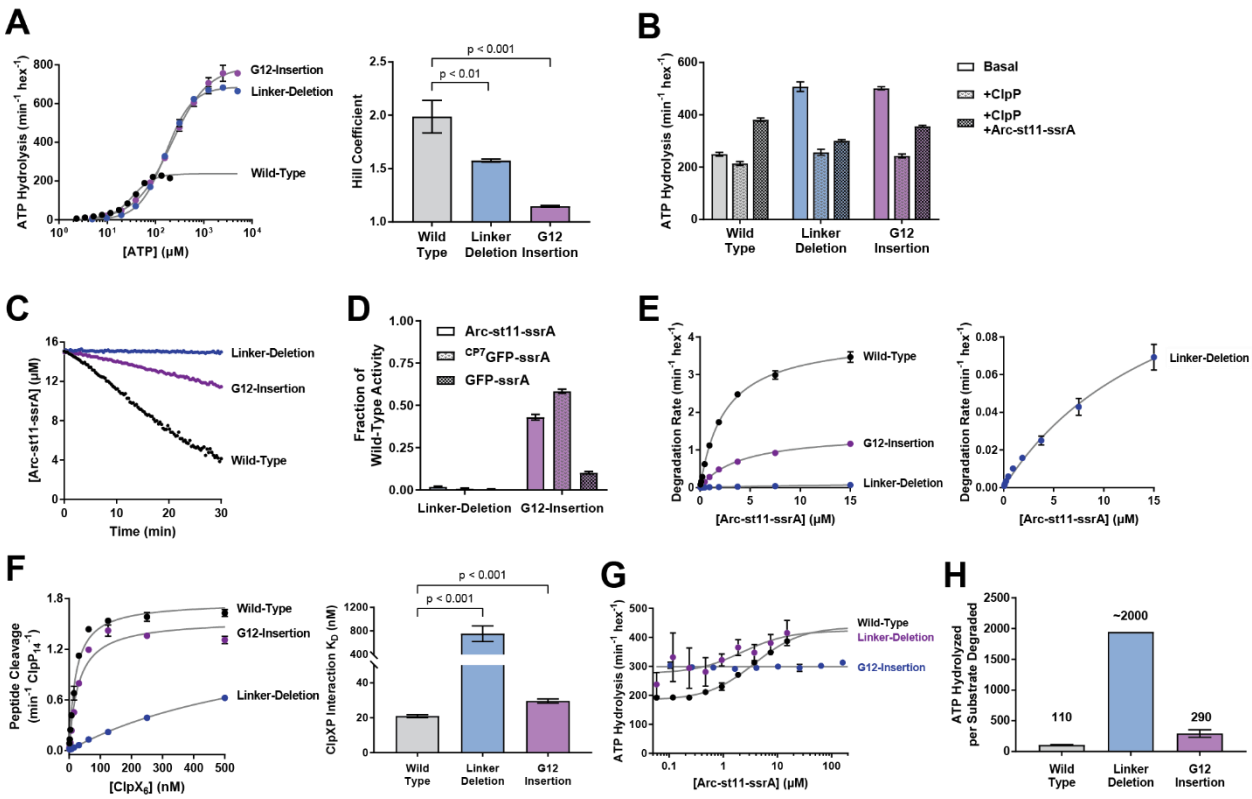


Figure 2.2 – Hinge-linker deletion and mutation impair intersubunit communication
Unless noted, experimental values are averages of three independent replicates ± SD. Concentrations are given as follows: ClpX, pseudohexamer units; ClpP, 14-mer units; substrates, monomer units. **(A) Left** – Hill fits of the concentration dependence of ATP-hydrolysis rates. **Right** – Hill values with the significance of differences from wild type evaluated by a Student's two-tailed t-test for the hinge-linker deletion (t = 4.6, dof = 4, p = 0.0098) or G12 insertion (t = 9.5, dof = 4, p = 0.0007). **(B)** Rates of ATP hydrolysis by pseudohexamer variants alone (0.03 μM), in the presence of ClpP, or in the presence of ClpP and Arc-st11-ssrA. Linker-deletion pseudohexamer was tested at higher concentrations of ClpP and Arc-st11-ssrA because of its reduced affinity for both molecules (see panels **E** and **F**). Wild-type and G12-insertion pseudohexamers: 0.09 μM ClpP, 15 μM Arc-st11-ssrA. Linker-deletion pseudohexamer: 6.7 μM ClpP, 160 μM Arc-st11-ssrA. **(C)** Kinetics of degradation of fluorescent Arc-st11-ssrA (15 μM) by single-chain pseudohexamers (0.1 μM) and ClpP (0.3 μM). Data are representative of three replicates. **(D)** Degradation of Arc-st11-ssrA, ^{CP7}GFP-ssrA, or GFP-ssrA (20 μM) by pseudohexamer variants (0.1 μM ClpX, 0.3 μM ClpP). Rates are plotted as a fraction of the wild-type rate. **(E) Left** – Michaelis-Menten plots of rates of Arc-st11-ssrA degradation by pseudohexamer variants (0.1 μM ClpX, 0.3 μM ClpP). **Right** – Same plot as left, rescaled to show fit of linker-deletion curve. **(F) Left** – pseudohexamer stimulation of ClpP (50 nM) cleavage of a fluorogenic peptide (15 μM). **Right** – p values determined by Student's two-tailed t-test of K_D

differences between the wild type pseudohexamer and the hinge-linker deletion (t = 9.6, dof = 4, p = 0.0006) or G12 insertion (t = 11, dof = 4, p = 0.0004). **(G)** Changes in ATP hydrolysis in response to increasing concentrations of Arc-st11-ssrA substrate. Wild-type and G12-insertion pseudohexamers: 0.1 μM ClpX, 0.3 μM ClpP, fit to an initial value and a hyperbolic increase to a saturated value. Linker-deletion pseudohexamer: 0.1 μM ClpX, 10 μM ClpP, fit to a constant linear value. **(H)** Energetic cost of ClpXP degradation of Arc-st11-ssrA supported by different pseudohexamers. Due to uncertainty inherent in the V_{max} determined for linker-deletion pseudohexamer, the calculated energetic cost for this variant is approximate.

We assayed the ability of the parental pseudohexamer, the linker-deletion variant, and the G12-insertion variant to support ClpP degradation of 15 μM Arc-st11-ssrA (Figure 2.2C). The G12-insertion variant supported ClpP degradation at ~25% of the parental rate, whereas the linker-deletion variant supported degradation at less than 1% of the parental rate. We also measured proteolysis activity against ^{CP7}GFP-ssrA and GFP-ssrA, which are incrementally more difficult to unfold and degrade (Nager et al., 2011). The G12-insertion variant supported ClpP degradation of these substrates at lower rates than the wild-type pseudohexamer, but the linker-deletion variant did not appreciably degrade either of these more-stable substrates (Figure 2.2D) Thus, deletion of a single hinge-linker dramatically reduces degradation activity, whereas mutational insertion diminishes activity far less severely across a range of substrate stabilities.

In principle, poor binding to the ssrA-tagged substrate or to ClpP could be responsible for the reduced degradation activities of the hinge-linker variants. To test substrate binding, we assayed degradation rates as a function of the concentration of Arc-st11-ssrA (Figure 2.2E, Table 2.1). The parental enzyme and the G12-insertion variant had K_M values for the protein substrate within ~2-fold, whereas K_M for the linker-deletion variant increased significantly. To assay ClpP binding, we

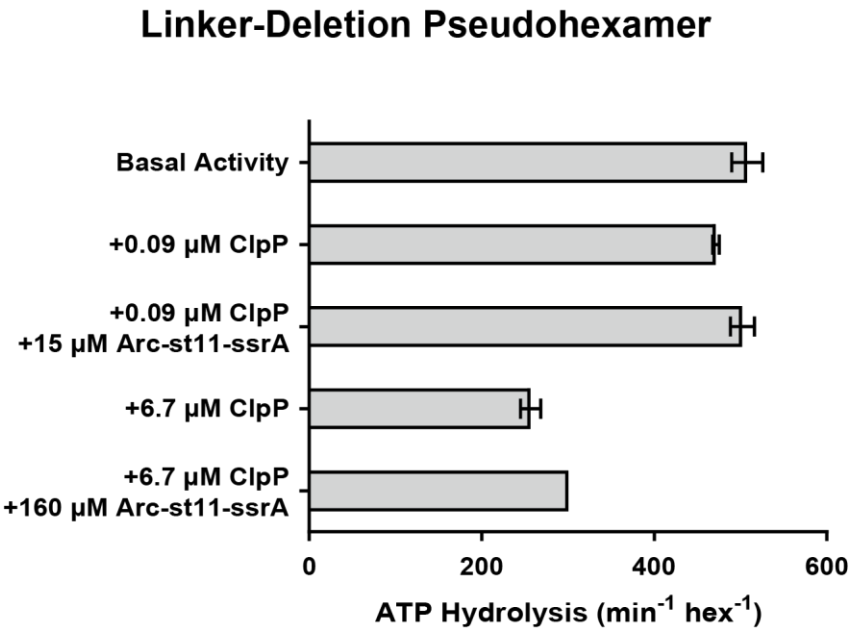


Figure 2.3

ATP hydrolysis by linker-deletion pseudohexamer (0.03 μM) under the conditions used in Fig. 2B for wild-type and G12-insertion pseudohexamers (0.09 μM $ClpP_{14}$, 15 μM Arc-st11-ssrA monomeric) and for linker-deletion pseudohexamer (6.7 μM $ClpP_{14}$, 160 μM Arc-st11-ssrA monomeric). Experimental values are averages of three independent replicates ± SD.

took advantage of the fact that ClpX activates ClpP degradation of a small peptide substrate by opening the axial pore (Grimaud et al., 1998). The parental and G12-insertion enzymes bound ClpP with low nanomolar affinities in this assay, whereas the linker-deletion variant bound with an affinity of ~750 nM (Figure 2.2F, Table 2.1). Thus, the presence of all six hinge-linkers appears to be required for robust interaction between ClpX and ClpP.

Because the linker-deletion pseudohexamer appeared to have low affinity for ClpP, we were concerned that our measurements could underestimate degradation rates because they were performed at concentrations of ClpP below the K_D measured by pore opening. To test this model directly, we compared rates of degradation of 15 μM Arc-st11-ssrA by the linker-deletion pseudohexamer in the presence of different concentrations of ClpP (Figure 2.4A). The apparent affinity was substantially higher in this assay, and the linker-deletion pseudohexamer was largely saturated in degradation assays performed in the presence of 300 nM $ClpP_{14}$ (Figure 2.4B). Although we do not know why the pore-opening and degradation assays give different affinities, one possibility is that the presence of a protein substrate strengthens the interaction between ClpX and ClpP.

We also assayed rates of ATP hydrolysis in the presence of ClpP and increasing concentrations of Arc-st11-ssrA (Figure 2.2G, Table 2.1), which stimulated ATP hydrolysis by the parental enzyme and G12-insertion variant but had no significant effect on the linker-deletion variant. The maximal rate of ATP hydrolysis at saturating Arc-st11-ssrA divided by V_{max} for substrate degradation provides an estimate of the ATP cost of degradation (Figure 2.2H). Degradation of one Arc-st11-ssrA protein required hydrolysis of ~110 ATPs by the parental enzyme and ~290 ATPs by the

ATP hydrolysis (Fig. 2A)			
	V_{max} / E_{TOT} (min^{-1} hex^{-1})	***K_M* (μM)**	**n**
Wild-Type	240 ± 10	33 ± 1	2.0 ± 0.2
Linker-Deletion	670 ± 20	160 ± 10	1.6 ± 0.1
G12-Insertion	790 ± 30	220 ± 10	1.1 ± 0.1
Arc-st11-ssrA degradation (Fig. 2E)			
	V_{max} / E_{TOT} (min^{-1} hex^{-1})	***K_M* (μM)**	
Wild-Type	4.1 ± 0.2	2.5 ± 0.2	
Linker-Deletion	~0.2 *	~20 *	
G12-Insertion	1.5 ± 0.1	4.1 ± 0.1	
ClpP binding (Fig. 2F)			
	V_{max} / E_{TOT} (min^{-1} $ClpP_{14}^{-1}$)	***K_D* (nM)**	
Wild-Type	1.8 ± 0.1	21 ± 1	
Linker-Deletion	1.6 ± 0.1	750 ± 130	
G12-Insertion	1.6 ± 0.1	30 ± 1	
ATP hydrolysis dependence on protein substrate concentration (Fig. 2G)			
	V_{init} / E_{TOT} (min^{-1} hex^{-1})	**V_{max} / E_{TOT} (min^{-1} hex^{-1})**	***K_{app}* (μM)**
Wild-Type	180 ± 10	440 ± 10	3.8 ± 0.8
Linker-Deletion	-	300 ± 10	-
G12-Insertion	190 ± 100	440 ± 90	2.2 ± 4
Arc-st11-ssrA degradation (Fig. 6B)			
	V_{max} / E_{TOT} (min^{-1} hex^{-1})	***K_M* (μM)**	
G12-Insertion	1.4 ± 0.1	6.9 ± 0.5	
G12-Ins, FAB_{EQ}	0.085 ± 0.006	3.0 ± 0.7	
G12-Ins, DEF_{EQ}	0.19 ± 0.02	4.5 ± 2	
ATP Hydrolysis dependence on protein substrate concentration (Fig. 6C)			
	V_{init} / E_{TOT} (min^{-1} hex^{-1})	**V_{max} / E_{TOT} (min^{-1} hex^{-1})**	***K_{app}* (μM)**
G12-Insertion	160 ± 10	300 ± 10	4.6 ± 0.4
G12-Ins, FAB_{EQ}	160 ± 10	200 ± 10	2.8 ± 4
G12-Ins, DEF_{EQ}	-	86 ± 1	-

Table 2.1 – Fitted parameters from biochemical analysis of single-chain $ClpX^{\Delta N}$ hexamers
Values are reported as average ± SD. The reported error was calculated by fitting independent replicates to curves, and then measuring the mean and variance of the resultant fitted parameters. Values labeled * are approximate because degradation rates for substrate concentrations above K_M could not be tested.

G12-insertion variant. Although weak interaction with substrate precluded precise determination of the V_{max} for substrate degradation by the linker-deletion variant, extrapolating the fitted curve suggests a bulk energy cost of ~2000 ATPs, substantially higher than either the parental or G12-insertion variants. These results support a model in which hinge-linker deletion or disruption reduce the mechanical efficiency of ClpX to different extents, either because many power strokes fail or because ATP hydrolysis and power strokes are uncoupled.

Hinge-linkers facilitate formation of closed rings

ClpX functions as a topologically closed ring (Glynn et al., 2012). In negative-stain electron microscopy, the parental pseudohexamer and G12-insertion variant exclusively showed ring structures in 2D class averages (Figures 2.5A, 2.5B). By contrast, the linker-deletion variant (Figure 2.5C) displayed a mixture of C-shaped open structures (~55%), twisted shapes that did not classify discreetly (~5%), and rings (~40%). The last value should be viewed as an upper bound, as some particles that appear as rings in projection could be open in three dimensions. This mixture of open- and closed-ring structures suggests that a stable closed-ring conformation is unlikely to be the dominant conformation of the linker-deletion variant.

Multiple hinge-linker deletions prevent stable hexamer assembly

To test whether deletion of multiple hinge-linkers increases the severity of the mechanical defects observed for the linker-deletion variant, we constructed a circularly permuted $ClpX^{\Delta N}$ monomer (no hinge-linker), a circularly permuted single-chain dimer (one hinge-linker; similar to a variant characterized in Stinson et al., 2013), and a circularly permuted single-chain trimer (two hinge-linkers) as shown in Figure 2.6A. In size-exclusion chromatography experiments in the presence

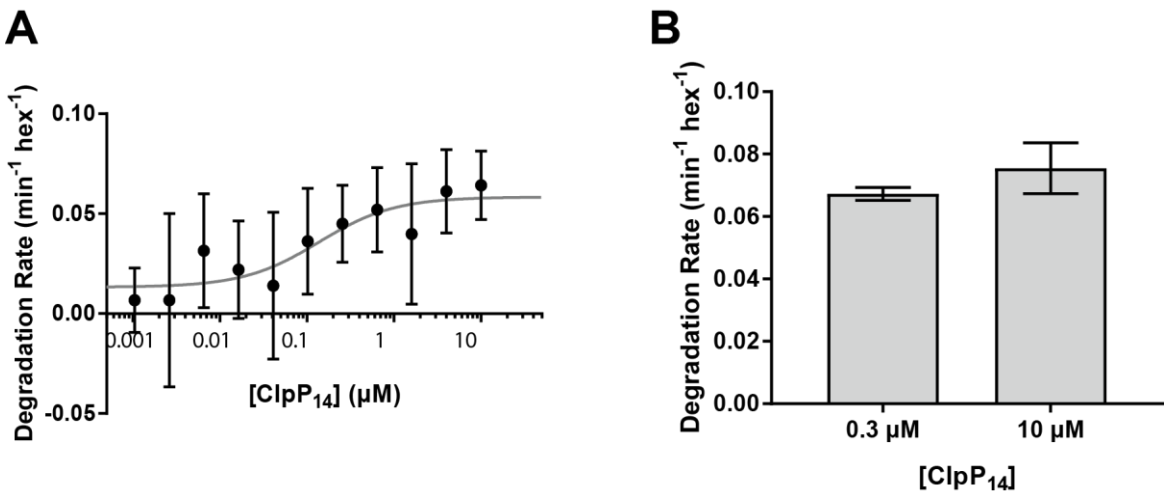


Figure 2.4 – Effect of ClpP concentration of Arc-st11-ssrA degradation rate
(A) Degradation of 15 μM Arc-st11-ssrA (monomeric) by 0.1 μM linker-deletion pseudohexamer in the presence of increasing amounts of $ClpP_{14}$. Experimental values are averages of three independent replicates ± SD. **(B)** Comparison of degradation rate of 15 μM Arc-st11-ssrA (monomeric) by 0.1 μM linker-deletion pseudohexamer in the presence of 0.3 μM or 10 μM $ClpP_{14}$. Experimental values are averages of three sets of three independent replicates ± SEM.

of ADP, which stabilizes assembly of ClpX$^{\Delta N}$ hexamers, these purified linker-deletion variants eluted at positions expected for monomers, dimers, and trimers (Figure 2.6B). Because non-permuted single-chain ClpX$^{\Delta N}$ dimers and trimers assemble into pseudohexamers (Martin et al., 2005), the hinge-linker deletions clearly compromise stable hexamer formation. Rates of basal ATP hydrolysis increased in rough proportion to the number of hinge-linkers in each variant (Figure 2.6C). None of the variants tested exhibited significant increases in hydrolysis rate in response to Arc-st11-ssrA substrate, and only linker-deletion single-chain hexamer exhibited hydrolysis repression in the presence of ClpP. The severe hydrolysis defect of the monomeric linker-deletion variant is not surprising, as side chains from neighboring subunits are required to form the wild-type active site for ATP hydrolysis (Glynn et al., 2009).

We used a FRET-based Arc-Gcn4-ssrA unfolding assay (Baytshtok et al., 2015) to measure substrate denaturation by different linker-deletion variants in the absence of ClpP (Figure 2.6D). The linker-deletion dimer and trimer displayed no detectable unfolding, whereas the linker-deletion hexamer displayed less than 1% of the activity of wild-type ClpX$^{\Delta N}$. In the presence of ClpP, low levels of degradation activity were observed for the linker-deletion dimer and trimer (Figure 2.6E), suggesting that ClpP can promote unfolding by serving as a scaffold for hexamer assembly. Interestingly, the ClpP degradation activities of the linker-deletion dimer (three missing hinge-linkers per hexamer) and trimer (two missing hinge-linkers per hexamer) were roughly 30% and 50%, respectively, of that of the circularly permuted hexamer (one missing hinge-linker per hexamer). Hence, although deletion of a single hinge-linker causes a ~100-fold decrease in degradation activity, additional deletions cause degradation defects roughly proportional to the number of missing hinge-linkers (Figure 2.6F). To exclude the possibility that the low degradation

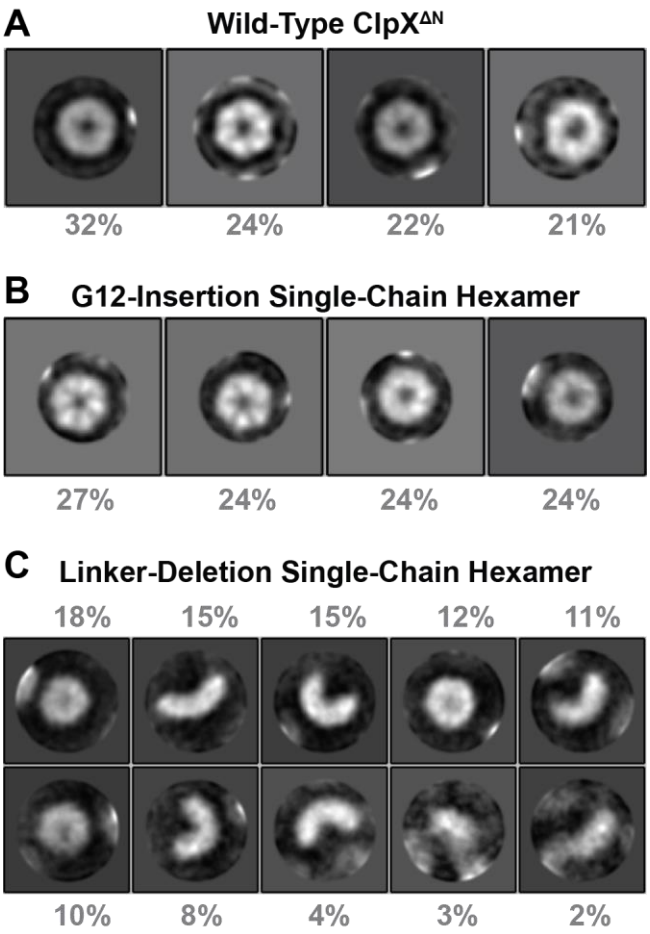


Figure 2.5 – Hinge-linker deletion disrupts ClpX ring structure
2D class averages of negative-stain EM images of **(A)** wild-type ClpX$^{\Delta N}$ hexamers; **(B)** G12-insertion pseudohexamers; and **(C)** linker-deletion pseudohexamers. The percent of particles assigned to each class are indicated.

activities of the linker-deletion variants might result from ATP-independent degradation, we tested degradation in the presence of ATPγS, an ATP analog hydrolyzed very slowly by ClpX (Burton et al., 2003). Under these conditions, the linker-deletion variants did not support detectable levels of ClpP-mediated degradation, confirming that degradation results from ClpX mechanical activity (Figure 2.6E).

Re-closing a broken hinge-linker interface improves mechanical activity

Because the linker-deletion pseudohexamer degrades substrates slowly and inefficiently and often adopts open-ring conformations, we asked whether covalently re-closing the ring would restore mechanical activity. To permit crosslinking, we engineered a variant with Cys-Leu and Leu-Cys residues at the N- and C-terminal ends of the pseudohexamer, respectively (dual-Cys $ClpX^{\Delta N}$, Figure 2.7A), and saturated the reaction with ClpP with the aim of bringing the terminal Cys residues into proximity for crosslinking. We mixed dual-Cys $ClpX^{\Delta N}$ with 1,11-bismaleimido-triethyleneglycol (BM-PEG_3, a bifunctional maleimide crosslinker with an 18 Å flexible spacer) in the presence of a 100-fold excess of a Cys-free $ClpP_{14}$ variant and ATPγS (Figure 2.7A). To control for nonspecific effects of sulfhydryl modification, a parallel reaction was treated with N-propyl-maleimide (NPM), a monofunctional maleimide crosslinker (Figure 2.7A). After crosslinking, a supershifted ClpX band was observed only in the BM-PEG_3 reaction, consistent with covalent circularization of the ClpX pseudohexamer (Stinson et al., 2013) (Figure 2.7B). Compared to the NPM-modified enzyme, dual-Cys $ClpX^{\Delta N}$ treated with BM-PEG_3 supported substantially faster ClpP degradation of Arc-st11-ssrA (Figure 2.7C) and ^{CP7}GFP-ssrA (Figure 2.7D). The hinge-linkers therefore contribute to robust degradation by enforcing a closed-ring topology.

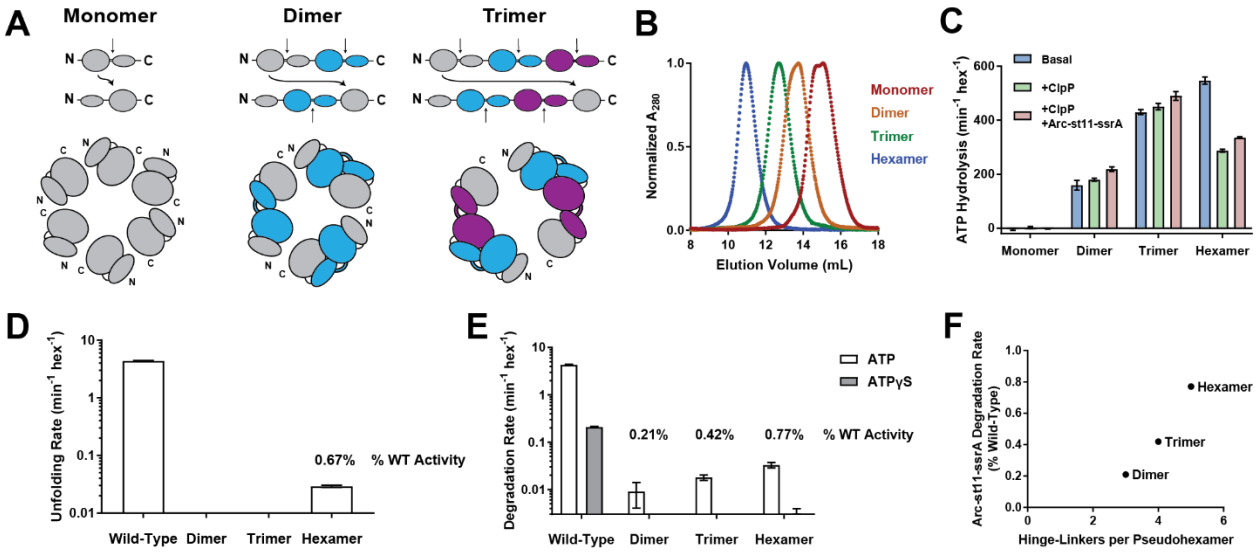


Figure 2.6 – Smaller hinge-linker variants fail to form hexamers

Unless noted experimental values are averages of three independent replicates ± SD. Concentrations are given as follows: ClpX, pseudohexamer units; ClpP, 14-mer units; substrates, monomer units. **(A)** Construction of circularly permuted variants that delete hinge-linkers from ClpX$^{\Delta N}$ monomers, pseudodimers, and pseudotrimers. **(B)** Superdex 200 size-exclusion chromatography of linker-deletion variants, performed in the presence of 1 mM ADP, which stabilizes native ClpX$^{\Delta N}$ hexamers. **(C)** Rates of ATP hydrolysis by hinge-linker variants alone (0.1 μM), in the presence of ClpP (6.7 μM), or in the presence of ClpP (6.7 μM) and Arc-st11-ssrA (160 μM). **(D)** Rates of unfolding of Arc-Gcn4-ssrA (20 μM) by linker-deletion variants (0.3 μM) in the absence of ClpP plotted on a logscale. **(E)** Rates of degradation of Arc-st11-ssrA (15 μM) by linker-deletion variants (0.3 μM) in the presence of ClpP (0.9 μM) and ATP or ATPγS (10 mM) plotted on a logscale. **(F)** Arc-st11-ssrA degradation activity of linker-deletion variants (taken from panel **E**) plotted as a function of the number of hinge-linkers per hexamer equivalent.

Functional asymmetry relative to the position of a hinge-linker mutation

To better understand how the hinge-linkers transmit information between ClpX subunits, we constructed single-chain hexamers with Walker-B ATP-hydrolysis mutations in three subunits counter-clockwise (FAB_{EQ}) or clockwise (DEF_{EQ}) from a G12 hinge-linker insertion (when viewed from the substrate binding face; Figure 2.8A). V_{max} for ClpP degradation of Arc-st11-ssrA was reduced for both variants compared to the G12-insertion parent, but the DEF_{EQ} variant was about twice as active as the FAB_{EQ} variant (Figure 2.8B, Table 2.1). However, the maximal rate of ATP hydrolysis was higher for FAB_{EQ} than for DEF_{EQ} in the presence of ClpP and Arc-st11-ssrA (Figure 2.8C, Table 2.1). Indeed, the G12 parent hydrolyzed an average of 210 ATPs per Arc-st11-ssrA degraded, whereas the cost was 450 ATPs for DEF_{EQ} and 2400 ATPs for FAB_{EQ} (Figure 2.8D). Thus, the hinge-linker insertion is less deleterious when it is clockwise from two wild-type subunits than when it is counter-clockwise.

DISCUSSION

The mechanical functions of the ClpX ring depend upon conformational rearrangements driven by ATP binding and hydrolysis. Hinge-linkers that connect the large and small AAA+ domains of each subunit in the hexamer appear to be the primary sites of flexibility that allow conformational motion within the ring (Glynn et al., 2009, 2012). However, it was unclear previously how the hinge-linkers contribute to ATP hydrolysis and substrate unfolding by ClpX. The present work identifies two distinct roles for these elements in facilitating intersubunit communication and robust unfolding activity.

Several experiments reveal a major role of the hinge-linkers in maintaining closed hexameric rings. As assayed by electron microscopy, deleting one hinge-linker in a hexamer allows a majority of

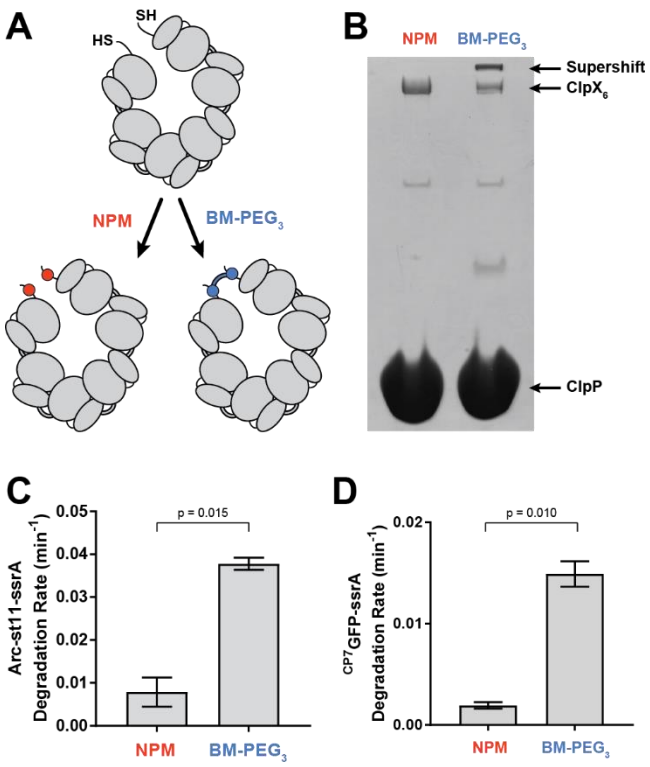


Figure 2.7 – Crosslinking a pseudohexamer with a hinge-linker deletion improves degradation

(A) Cartoon of crosslinking and control reactions. **(B)** 10% SDS-PAGE analysis after the modification reactions depicted in panel A. The supershifted species corresponds to a circularized hexamer. Rates of Arc-st11-ssrA (20 μM monomer) **(C)** and ^{CP7}GFP-ssrA (20 μM monomer) **(D)** degradation by the NPM- and BM-PEG$_3$-modified linker-deletion variants. Values are averages ± SEM of two independent sets of three replicates. Significance of differences determined by Student's two-tailed t-test are shown in panel **C** (t = 8.1, dof = 2) and panel D (t = 10, dof = 2).

molecules to adopt open-ring conformations. The ability to form and maintain a topologically closed ring also correlates with robust ClpX function. Deletion of one hinge-linker prevents ClpXP degradation of stable protein substrates and dramatically slows proteolysis of less stable substrates, and deletion of multiple hinge-linkers abrogates hexamer assembly and mechanical function. Insertion of a flexible 12-glycine sequence into a single hinge-linker of a hexamer resulted in less severe functional defects than those associated with deleting a hinge-linker. Similarly, the proteolysis defects of the single hinge-linker deletion were partially rescued by a chemical crosslink that restored covalent connectivity.

The hinge-linkers play a second important role facilitating communication and regulating activity within the closed ClpX ring. We find that insertion of 12 glycines into a single hinge-linker increases the maximal velocity of ATP hydrolysis, reduces the positive cooperativity of hydrolysis with respect to ATP concentration, and substantially increases the average number of ATP hydrolysis events required for degradation of one substrate. Interestingly, inserting 12 glycines into a hinge-linker within the context of the closed ring decreases ATP hydrolysis cooperativity to a greater extent than the open-ring hinge-linker deletion, suggesting that hinge-linker length may be optimized for intersubunit communication within a closed-ring topology. We conclude that the wild-type hinge-linkers contribute to regulation of basal ATP hydrolysis rate and efficiently couple ATP hydrolysis with the mechanical processes needed for protein unfolding and translocation. Previous studies reinforce this conclusion, as single-residue insertions or deletions in all six ClpX hinge-linkers also reduce the rate of degradation and increase the ATP cost (Glynn et al., 2012). Because disrupting a single hinge-linker is phenotypically similar to disrupting all six, the hinge-linkers may mediate long-range communication between subunits within the ring. A mechanism

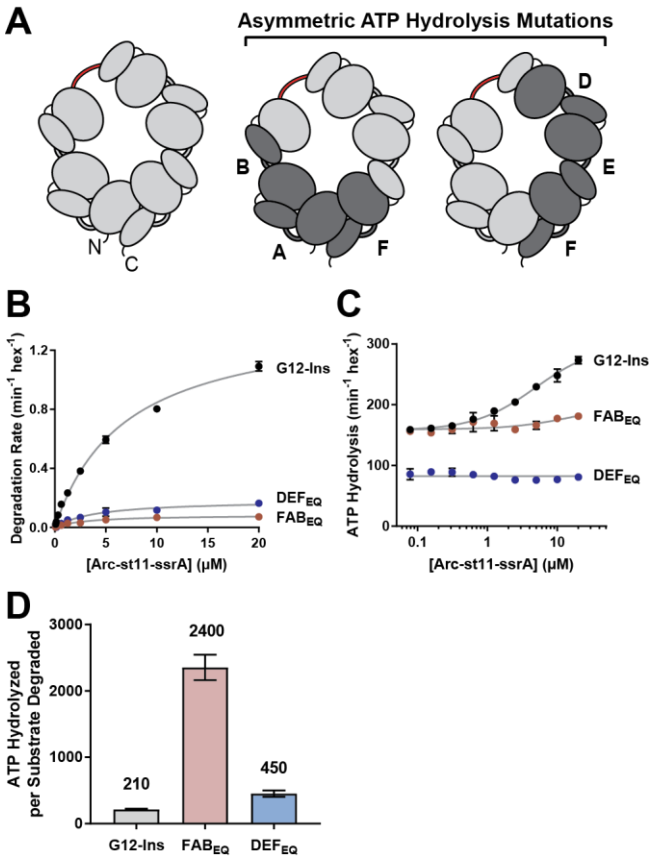


Figure 2.8 – Hinge-linker mutation reveals functional asymmetry

Concentrations are given as follows: ClpX, pseudohexamer units; ClpP, 14-mer units; substrates, monomer units. **(A)** Cartoons of pseudohexamer variants with the substrate-binding face up. The G12 insertion is red; subunits with wild-type active-site residues for ATP hydrolysis are light gray; and subunits carrying the Walker-B E185Q ATP-hydrolysis mutation are dark gray. **(B)** Michaelis-Menten analysis of Arc-st11-ssrA degradation (0.1 μM pseudohexamer variants, 0.3 μM ClpP). Values are averages of three independent replicates ± SD. **(C)** Changes in ATP hydrolysis as a function of Arc-st11-ssrA concentration were fit as described in the legend of panel 2G (0.1 μM pseudohexamer variants, 0.3 μM ClpP). Values are averages of three independent replicates ± SD. **(D)** Energetic cost of Arc-st11-ssrA degradation. Values are averages ± SEM of two independent sets of three replicates.

in which conformational changes in one subunit are sensed by all other subunits via tension transmitted through the hinge-linkers is consistent with this model of intersubunit communication.

Compared to the parent or G12 insertion, the single hinge-linker deletion weakened ClpX affinity for ClpP and protein substrates, suggesting that a stable closed ring is needed for optimal binding. Structures of ClpP reveal that its ClpX-interacting surface is planar (Wang et al., 1997). In an open-ring structure, some of the ClpX loops that normally interact with clefts on the ClpP surface are unlikely to make productive contacts (Amor et al., 2016; Kim et al., 2001; Martin et al., 2007). In support of this model, a previous study demonstrates that deleting one of the six loops that interact with ClpP from a single-chain ClpX pseudohexamer weakens ClpP interaction ~50-fold (Martin et al., 2007), whereas our hinge-linker deletion variant has a ~35-fold binding defect as measured by ClpP pore opening. Furthermore, ClpP may function as a scaffold for proper ClpX geometry, enforcing a ring-hexamer conformation. Linker-deletion single-chain dimers and trimers, which do not detectably unfold substrates on their own, support low levels of ClpP-mediated substrate degradation. Additionally, both linker-deletion and G12-insertion single-chain hexamer variants exhibit high levels of ATP hydrolysis on their own, but this activity is repressed to near-parental levels when sufficient ClpP is added. It seems likely that ClpX must adopt a ring-hexamer topology to productively interact with ClpP, and that the weakened affinity of the linker-deletion variant for ClpP reflects the entropic penalty of constraining it to a closed-ring conformation.

Loss of coordinated contacts from multiple subunits could also explain the substrate-binding defect of the single hinge-linker deletion variant. Residues in the axial-pore-1 loops of ClpX are important

for binding of ssrA-tagged substrates, and mutating increasing numbers of these pore loops increases K_M for substrate degradation (Iosefson et al., 2015a, 2015b; Martin et al., 2008; Rodriguez-Aliaga et al., 2016; Siddiqui et al., 2004). Similar to the observed ClpP binding defect, substrate binding may be impaired in the linker-deletion pseudohexamer because the open-ring conformation prevents pore-1 loops on either side of the broken interface from coordinating grip and unfolding activity during substrate processing.

We find that placing Walker-B ATP hydrolysis mutations on one side or the other of a disrupted hinge-linker results in an approximate 5-fold difference in the average number of ATPs required to degrade a single substrate. This result suggests that communication between subunits in the ClpX hexamer occurs with some directional bias. Single-molecule studies indicate that kinetic bursts of power strokes from multiple ClpX subunits contribute to mechanical activity (Aubin-Tam et al., 2011; Cordova et al., 2014; Maillard et al., 2011; Olivares et al., 2017; Sen et al., 2013). Thus, asymmetry relative to a disrupted hinge-linker could arise if conformational changes are communicated more efficiently through a wild-type hinge-linker on its clockwise side, allowing coordinated action of more subunits during each unfolding power stroke. However, axial pore contacts from multiple subunits also contribute to substrate grip and thus to unfolding efficiency (Iosefson et al., 2015a, 2015b; Rodriguez-Aliaga et al., 2016), and the observed asymmetry could also arise from directional effects on grip.

Many recent structures of ATP-dependent protein translocation motors show the six subunits of each hexamer in a spiral staircase topology (Gates et al., 2017; Lander et al., 2012; Monroe et al., 2017; Puchades et al., 2017). The asymmetric nature of these structures has led to proposed

mechanisms in which subunits hydrolyze ATP sequentially, facilitating hand-over-hand substrate translocation (Gates et al., 2017; Monroe et al., 2017; Puchades et al., 2017). A recent study using cryo EM single-particle analysis identified classes in which ATP was bound to different sets of protomers within the Rpt_{1-6} motor, which could represent intermediates in a rotary hydrolysis and translocation mechanism (Eisele et al., 2018). However, ClpX has been shown to hydrolyze ATP and translocate substrates via a probabilistic rather than sequential mechanism (Cordova et al., 2014; Martin et al., 2005), and several other AAA+ unfoldases are also unlikely to be constrained to a sequential mode of action (Baytshtok et al., 2017; Beckwith et al., 2013). Our observation of functional asymmetry relative to a hinge-linker disruption is consistent with a hybrid model in which ClpX operates most efficiently when bursts of ATP hydrolysis initiate stochastically, but subsequent hydrolysis events are biased toward a sequential, unidirectional mode of action.

The principles identified for hinge-linkers in ClpX may be broadly applicable to other protein unfolding motors. The AAA+ protein unfoldases HslU, FtsH, Yme1, PAN, ClpA, ClpB, Hsp104, Cdc48/p97, and the Rpt_{1-6} ring of the proteasome all share conserved large and small domains connected by a hinge-linker. The sequence of the ClpX hinge-linker region is poorly conserved in proteobacteria but its length is invariant (Glynn et al., 2012). In annotated crystal structures of AAA+ unfoldases from multiple clades, the hinge-linkers are also 4-5 amino acids in length. This possible evolutionary constraint on hinge-linker length suggests that defined length is an important determinant of function for most, if not all, AAA+ protein remodeling machines.

Conclusions

In conclusion, we have characterized the structural and functional impacts of deleting or disrupting

a single hinge-linker element in the ClpX protein unfolding motor. Deleting one hinge-linker prevents ring-hexamer assembly and causes defects in substrate unfolding, ATP hydrolysis cooperativity, and ClpP binding. Mutational insertion of 12 glycines into a hinge-linker preserves closed-ring structure, but impairs substrate unfolding and ATP hydrolysis cooperativity. We further validated the significance of the closed ring structure by demonstrating that covalently reclosing a broken hinge-linker interface partially restores substrate unfolding activity. Finally, we report an asymmetry in ClpX degradation efficiency when a hinge-linker mutation is placed on either side of hydrolysis-defective subunits, suggesting biased directional effects on grip or intersubunit communication through the hinge-linkers.

METHODS

Protein expression and purification

$ClpX^{\Delta N}$ (residues 62-424) constructs were expressed and purified as described (Martin et al., 2005). Linker-deletion $ClpX^{\Delta N}$ was produced by circularly permuting covalently tethered $ClpX^{\Delta N}$ constructs using Gibson Assembly (New England Biolabs). The G12-insertion was introduced into one hinge-linker of a $ClpX^{\Delta N}$ pseudohexamer (Martin et al., 2005) after residue N315 using PCR mutagenesis. Arc-st11-ssrA, Arc-Gcn4-st11-ssrA, GFP-ssrA, and ^{CP7}GFP-ssrA substrates were purified as described (Baytshtok et al., 2015; Kenniston et al., 2003; Nager et al., 2011).

Linkerless ClpX assembly assays

The oligomeric state of linker-deletion constructs was determined by analytical size-exclusion chromatography at 4°C. Protein samples were diluted to 3 μM (hexamer equivalents) and 200 μL was injected onto a Superdex-200 10/300 column (GE Healthcare) equilibrated in buffer A (25

mM HEPES-KOH, pH 7.5, 5 mM $MgCl_2$, 200 mM KCl, 10% glycerol) supplemented with 1 mM ADP.

$ClpX^{\Delta N}$ hexamers were analyzed using negative-stain electron microscopy performed by Nicki Watson (W.M. Keck Microscopy Facility, Whitehead Institute, MIT). Freshly-ionized 200-mesh Cu/carbon grids were floated on 10 µL drops of protein sample for ~60 s. The grids were then blotted dry, washed with 1% aqueous uranyl acetate, and blotted dry again. Grids were imaged at 100 kV on a Technai Spirit Transmission Electron Microscope (FEI). Electron micrographs were class averaged using RELION 2.0 software (Scheres, 2012). For the wild-type, G12-insertion, and linker-deletion variants imaged, 641, 792, and 1287 particles, respectively, were manually selected at 100 pixel diameter from micrographs taken at 49,000-fold (wild-type and linker-deletion) or 98,000-fold (G12-insertion) magnification. The length-to-pixel ratio was increased two-fold for the higher-magnification G12-insertion micrographs to facilitate consistent downstream analysis. For the wild-type and G12-insertion constructs, in which the particles appear as compact rings, an 80-pixel opaque circular mask was applied to each particle, and particles were averaged into four classes. For the linker-deletion construct, a 100-pixel mask was applied, and particles were averaged into 10 classes to accommodate the less compact and more heterogeneous nature of these particles.

Biochemical assays

Activity assays were carried out at 30°C in buffer PD (25 mM HEPES-KOH, pH 7.5, 35 mM $MgCl_2$, 200 mM KCl, 10% glycerol). Kinetics were analyzed using Prism 7 (GraphPad). ATP hydrolysis was measured using a coupled-NADH oxidation assay as described (Martin et al.,

2005). ClpP binding assays were performed using the protocol and RseA fluorogenic decapeptide substrate described (Lee et al., 2010). Solution unfolding was assayed using a FRET-based unfolding assay as described (Baytshtok et al., 2015).

GFP-ssrA and ^{CP7}GFP-ssrA degradation assays were performed by measuring loss of GFP fluorescence (excitation 467 nm; emission 511 nm) using a SpectraMax M5 plate reader (Molecular Devices). For fluorescence-based Arc-st11-ssrA degradation assays, purified Arc-st11-ssrA was labeled with a 10-fold molar excess of NHS-Fluorescein (ThermoFisher) in Buffer B (25 mM HEPES-KOH, pH 7.5, 300 mM KCl, 10% glycerol, 1 mM dithiothreitol [DTT]) for 45 min at ambient temperature. The reaction was quenched with an excess of Tris-HCl (pH 7.6), and the labeled protein was separated from free dye and desalted into Buffer A over a Superdex-75 16/600 column (GE Healthcare). Because Arc-st11-ssrA has six lysine residues, multiple labeling with fluorescein caused label molecules on the same substrate to mutually quench fluorescence emission. Substrate degradation by ClpXP was monitored by increased fluorescence (excitation 480 nm; emission 525 nm) using a SpectraMax M5 plate reader (Molecular Devices). Endpoint fluorescence was measured after adding an excess of trypsin to the reaction, and the amount of Arc-st11-ssrA degraded was calculated from the initial and final fluorescence values. Because each preparation of substrate was labeled to a different extent, we observed ~30% batch-to-batch variability in apparent degradation rate by the same prep of ClpXP. To prevent invalid comparison of rates between batches, we included appropriate controls in each experiment and reported Arc-st11-ssrA degradation activity relative to a wild-type ClpX$^{\Delta N}$ construct where practical.

For degradation efficiency experiments, ATP hydrolysis and substrate degradation activity were

measured as a function of substrate concentration, and the resulting responses were fit to hyperbolic functions to obtain V_{max} values for the rates of ATP hydrolysis and substrate degradation at saturating substrate. In cases where the ATP hydrolysis rate was unresponsive to increasing concentrations of substrate, a constant linear fit was used. To calculate mechanical efficiency, V_{max} for ATP hydrolysis was divided by V_{max} for substrate degradation.

Crosslinking

A variant of linker-deletion $ClpX^{\Delta N}$ pseudohexamer with cysteine residues on either side of the deleted hinge-linker interface was mixed with a 100-fold excess of a Cys-free *E. coli* ClpP variant (C91V/C113A). The protein mixture was desalted into PD buffer using G25 resin (GE Healthcare) to remove DTT. The desalted protein was separated into two pools, and 2 mM ATPγS was added to each pool. Either 1,11-bismaleimido-triethyleneglycol (Thermo Scientific) or equal maleimide equivalents of N-propyl-maleimide (Sigma) was then added. The maleimide agents were added in 1:1 stoichiometry with sulfhydryl groups on $ClpX^{\Delta N}$, and DMSO cosolvent was added to 12% of the reaction volume. The reactions were allowed to proceed for 1 hour at 30°C, then quenched with 1 mM DTT for 15 min at 30°C. The reactions were again desalted into PD buffer over PD-10 columns (GE Healthcare), and concentrated with centrifugal concentrators (Millipore-Sigma). Reactions were normalized by total protein using a Bradford assay prior to measuring degradation activity (Thermo Scientific). To gauge crosslinking efficiency, equal volumes of each reaction were separated on a 10% bis-tris SDS-polyacrylamide gel run at 120V in MES buffer, stained with Coomassie Brilliant Blue, and imaged.

ACKNOWLEDGMENTS

This work was supported by US National Institutes of Health (NIH) grant GM-101988 (R.T.S.). T.A. Bell was supported in part by US NIH grant 5T32GM-007287. T.A. Baker is an employee of the Howard Hughes Medical Institute. We thank Nicki Watson and Xue Fei for help with microscopy and downstream analysis; Steven Glynn, Karl Schmitz, Sanjay Hari, and Juhee Park for helpful discussions; and members of the Sauer and Baker labs for providing feedback on the manuscript.

REFERENCES

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.

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.

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. *112*, 5377–5382.

Baytshtok, V., Chen, J., Glynn, S.E., Nager, A.R., Grant, R.A., Baker, T.A., and Sauer, R.T. (2017). Covalently linked HslU hexamers support a probabilistic mechanism that links ATP hydrolysis to protein unfolding and translocation. J. Biol. Chem. *292*, 5695–5704.

Beckwith, R., Estrin, E., Worden, E.J., and Martin, A. (2013). Reconstitution of the 26S proteasome reveals functional asymmetries in its AAA+ unfoldase. Nat. Struct. Mol. Biol. *20*,

1164–1172.

Burton, R.E., Baker, T.A., and Sauer, R.T. (2003). Energy-dependent degradation: Linkage between ClpX-catalyzed nucleotide hydrolysis and protein-substrate processing. Protein Sci. *12*, 893–902.

Choy, J.S., Aung, L.L., and Karzai, A.W. (2007). Lon Protease Degrades Transfer-Messenger RNA-Tagged Proteins. J. Bacteriol. *189*, 6564–6571.

Cordova, J.C., Olivares, A.O., Shin, Y., Stinson, B.M., Calmat, S., Schmitz, K.R., Aubin-Tam, M.-E., Baker, T.A., Lang, M.J., and Sauer, R.T. (2014). Stochastic but Highly Coordinated Protein Unfolding and Translocation by the ClpXP Proteolytic Machine. Cell *158*, 647–658.

Eisele, M.R., Reed, R.G., Rudack, T., Schweitzer, A., Beck, F., Nagy, I., Pfeifer, G., Plitzko, J.M., Baumeister, W., Tomko, R.J., et al. (2018). Expanded Coverage of the 26S Proteasome Conformational Landscape Reveals Mechanisms of Peptidase Gating. Cell Rep. *24*, 1301-1315.e5.

Erzberger, J.P., and Berger, J.M. (2006). Evolutionary Relationships and Structural Mechanisms of Aaa+ Proteins. Annu. Rev. Biophys. Biomol. Struct. *35*, 93–114.

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*, 273–279.

Glynn, S.E., Martin, A., Nager, A.R., Baker, T.A., and Sauer, R.T. (2009). Structures of Asymmetric ClpX Hexamers Reveal Nucleotide-Dependent Motions in a AAA+ Protein-

Unfolding Machine. Cell *139*, 744–756.

Glynn, S.E., Nager, A.R., Baker, T.A., and Sauer, R.T. (2012). Dynamic and static components power unfolding in topologically closed rings of a AAA+ proteolytic machine. Nat. Struct. Mol. Biol. *19*, 616–622.

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*, 1338–1347.

Grimaud, R., Kessel, M., Beuron, F., Steven, A.C., and Maurizi, M.R. (1998). Enzymatic and Structural Similarities between theEscherichia coli ATP-dependent Proteases, ClpXP and ClpAP. J. Biol. Chem. *273*, 12476–12481.

Hanson, P.I., and Whiteheart, S.W. (2005). AAA+ proteins: have engine, will work. Nat. Rev. Mol. Cell Biol. *6*, 519–529.

Hari, S.B., and Sauer, R.T. (2016). The AAA+ FtsH Protease Degrades an ssrA-Tagged Model Protein in the Inner Membrane of Escherichia coli. Biochemistry *55*, 5649–5652.

Hersch, G.L., Burton, R.E., Bolon, D.N., Baker, T.A., and Sauer, R.T. (2005). Asymmetric Interactions of ATP with the AAA+ ClpX6 Unfoldase: Allosteric Control of a Protein Machine. Cell *121*, 1017–1027.

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

Iosefson, O., Olivares, A.O., Baker, T.A., and Sauer, R.T. (2015b). Dissection of Axial-Pore

Loop Function during Unfolding and Translocation by a AAA+ Proteolytic Machine. Cell Rep. *12*, 1032–1041.

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*, 990–993.

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.-I., Levchenko, I., Fraczkowska, K., Woodruff, R.V., Sauer, R.T., and Baker, T.A. (2001). Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. Nat. Struct. Biol. *8*, 230–233.

Lander, G.C., Estrin, E., Matyskiela, M.E., Bashore, C., Nogales, E., and Martin, A. (2012). Complete subunit architecture of the proteasome regulatory particle. Nature *482*, 186–191.

Lee, M.E., Baker, T.A., and Sauer, R.T. (2010). Control of Substrate Gating and Translocation into ClpP by Channel Residues and ClpX Binding. J. Mol. Biol. *399*, 707–718.

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*, 459–469.

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

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.

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*, 1147–1151.

Monroe, N., Han, H., Shen, P.S., Sundquist, W.I., and Hill, C.P. (2017). Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. ELife *6*, e24487.

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*, 4–16.

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.

Rodriguez-Aliaga, P., Ramirez, L., Kim, F., Bustamante, C., and Martin, A. (2016). Substrate-translocating loops regulate mechanochemical coupling and power production in AAA+ protease ClpXP. Nat. Struct. Mol. Biol. *23*, 974–981.

Sauer, R.T., and Baker, T.A. (2011). AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annu. Rev. Biochem. *80*, 587–612.

Scheres, S.H.W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. *180*, 519–530.

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*, 636–646.

Siddiqui, S.M., Sauer, R.T., and Baker, T.A. (2004). Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. Genes Dev. *18*, 369–374.

Stinson, B.M., Nager, A.R., Glynn, S.E., Schmitz, K.R., Baker, T.A., and Sauer, R.T. (2013). Nucleotide Binding and Conformational Switching in the Hexameric Ring of a AAA+ Machine. Cell *153*, 628–639.

Wang, J., Hartling, J.A., and Flanagan, J.M. (1997). The Structure of ClpP at 2.3 Å Resolution Suggests a Model for ATP-Dependent Proteolysis. Cell *91*, 447–456.

Chapter III

Interactions between a subset of substrate side chains and AAA+ motor pore loops determine grip during protein unfolding

This chapter is published:

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.

T.A. Bell conceived and performed all experiments. All authors contributed to data analysis and/or writing the paper.

ABSTRACT

Most AAA+ remodeling motors denature proteins by pulling on the peptide termini of folded substrates, but it is not well-understood how motors produce grip when resisting a folded domain. Here, at single amino-acid resolution, we identify the determinants of grip by measuring how substrate tail sequences alter the unfolding activity of the unfoldase-protease ClpXP. The seven amino acids abutting a stable substrate domain are key, with residues 2-6 forming a core that contributes most significantly to grip. ClpX grips large hydrophobic and aromatic side chains strongly and small, polar, or charged side chains weakly. Multiple side chains interact with pore loops synergistically to strengthen grip. In combination with recent structures, our results support a mechanism in which unfolding grip is primarily mediated by non-specific van der Waal's interactions between core side chains of the substrate tail and a subset of YVG loops at the top of the ClpX axial pore.

INTRODUCTION

Cells maintain homeostasis by balancing protein synthesis and degradation with growth. When nutrients are available, new proteins are constantly synthesized, whereas damaged, misfolded, or unneeded proteins are degraded. Regulated degradation typically requires a protein-unfolding motor of the AAA+ family (ATPases associated with various cellular activities) that associates with a self-compartmentalized protease (e.g., ClpX with ClpP, the 19S regulatory particle with the 20S proteasome) or is genetically tethered to a protease (e.g., Lon, FtsH, Yme1) (Glynn, 2017; Olivares et al., 2016; Sauer and Baker, 2011). When challenged with degrading a folded substrate, AAA+ motors couple ATP hydrolysis to mechanical motion that overcomes the resistance of the folded domain. Despite broad consensus on the overall mechanism of protein unfolding, it is largely unknown how interactions between a AAA+ motor and its substrate produce grip, the ability for the motor to maintain hold of the substrate while applying an unfolding force.

ClpX is a ring-shaped AAA+ homohexamer that functions autonomously in protein remodeling in bacteria and eukaryotic organelles and also associates with ClpP tetradecamers to form the ATP-dependent ClpXP protease (Baker and Sauer, 2012). Substrates are targeted to ClpX or ClpXP by N- or C-terminal peptide tails (also called degradation tags or degrons), which initially bind in the ClpX axial pore. Proteins marked with a sequence-defined degron or post-translational modification can be recruited to the AAA+ protease directly or with assistance from auxiliary adaptors (Sauer and Baker, 2011; Trentini et al., 2016). For example, during rescue of stalled ribosomes in *Escherichia coli*, the 11-residue ssrA tag is appended to the C-terminus of abortive protein products (Keiler et al., 1996), allowing ClpXP to recognize and degrade the attached protein (Farrell et al., 2007; Gottesman et al., 1998).

The ssrA tag and other degron tails interact with ClpX loops that line the axial pore. A Tyr-Val-Gly (YVG) sequence in the pore-1 loop is critical for substrate binding, unfolding, and translocation (Iosefson et al., 2015; Martin et al., 2008a, 2008b; Siddiqui et al., 2004). Other AAA+ unfolding motors contain related pore-1 loops and mutation of these loops typically abolishes function (Hinnerwisch et al., 2005; Park et al., 2005; Schlieker et al., 2004; Yamada-Inagawa et al., 2003). In several AAA+ unfolding motors, the pore-1 loops adopt a spiral staircase conformation within the pore, which facilitates multivalent interaction with the bound peptide tail (Dong et al., 2019; Gates et al., 2017; Majumder et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017; White et al., 2018). ATP-dependent conformational changes are thought to draw the tail of a substrate into the pore until a folded domain too large to transit the pore impedes progress. For ClpXP, repeated cycles of ATP hydrolysis are then required to unfold the substrate and to translocate the polypeptide through the pore and into ClpP for degradation (Aubin-Tam et al., 2011; Cordova et al., 2014; Kenniston et al., 2003; Maillard et al., 2011; Sen et al., 2013).

How the amino acids in the bound substrate tail contribute to grip during unfolding remains poorly understood. When degrading unfolded substrates, the rate of substrate translocation through ClpXP is largely insensitive to amino-acid charge, size, or peptide-bond spacing (Barkow et al., 2009). In contrast, when directly abutting a folded domain, sequences rich in glycine can result in failed unfolding by ClpXP or the 26S proteasome, leading to release of partially processed intermediates (Daskalogianni et al., 2008; Hoyt et al., 2006; Kraut, 2013; Levitskaya et al., 1997; Lin and Ghosh, 1996; Sharipo et al., 2001; Too et al., 2013; Vass and Chien, 2013). Abortive unfolding caused by

Gly-rich motifs occurs as a result of slower domain unfolding rather than rapid substrate dissociation, suggesting that these motifs bind normally but are gripped poorly during unfolding (Kraut, 2013; Kraut et al., 2012). These results suggest that AAA+ motors struggle to efficiently grip sequences with very small side chains. Alternatively, sequence complexity rather than composition may dictate grip strength (Tian et al., 2005).

Here, we use green fluorescent protein (GFP) reporter substrates to interrogate the contributions of individual amino acids in the peptide tail to grip strength by ClpXP. In degradation assays performed *in vivo* and *in vitro*, we observe that substrate grip by ClpX is primarily mediated by interactions with a block of five amino acids, located two to six residues from the native GFP domain. Through systematic mutation, we characterize the ability of each amino acid to promote ClpX grip, and find that aromatic and large hydrophobic residues are gripped well, whereas charged and polar residues impair grip. Finally, we analyze synergistic contributions of multiple residues to unfolding, and show that contacts with more than one side chain lead to stronger grip and faster substrate unfolding. Our results provide unprecedented detail into the mechanism by which AAA+ motors grip terminal substrate tails during protein unfolding.

RESULTS

Substrate design and degradation assays

To probe grip during substrate unfolding, we used *Aequorea victoria* GFP, as its native structure is highly kinetically stable, unfolding is rate limiting for ClpXP degradation of GFP-ssrA, and the pathway of mechanical unfolding of GFP-ssrA by ClpXP is well characterized (Kim et al., 2000; Maillard et al., 2011; Nager et al., 2011). In our substrates, we truncated GFP at Ile-229, the last

amino acid that makes extensive native contacts in multiple crystal structures (Ormö et al., 1996; Yang et al., 1996), and added a 12-residue cassette of variable sequence followed by a partial ssrA degron to allow recognition by ClpXP (Figure 3.1A). Given the length of the axial pore (~35 Å; Glynn et al., 2009), we reasoned that ClpX should only interact with residues within the cassette region during GFP unfolding.

For studies of intracellular degradation, we used an *E. coli* B strain, which lacks the AAA+ Lon protease; deleted the chromosomal copies of *clpP*, *clpX*, and *clpA*, as ClpAP can also degrade ssrA-tagged substrates (Farrell et al., 2007; Gottesman et al., 1998); and placed genes encoding $ClpX^{\Delta N}$ and ClpP on a plasmid under arabinose-inducible control (Figure 3.1B; Guzman et al., 1995). Despite lacking a family-specific N-terminal domain, $ClpX^{\Delta N}$ supports ClpP-degradation of ssrA-tagged substrates as well as wild-type ClpX but does not interact with many other cellular substrates and adaptors (Flynn et al., 2003; Martin et al., 2005; Singh et al., 2001; Wojtyra et al., 2003). GFP substrates were cloned under transcriptional control of a constitutive ProD promoter (Figure 3.1B; Davis et al., 2011). To control for ClpXP-independent degradation, we used a pBAD plasmid isogenic to the *clpP*/$clpX^{\Delta N}$ vector but lacking these genes (Figure 3.1B). To determine the extent of intracellular GFP degradation, we measured GFP fluorescence after arabinose induction and growth for 35 min.

We first used this system to characterize substrates with different high or low complexity sequences in the 12-residue cassette. Substrate expression levels were sensitive to the cassette sequence, possibly because of effects on mRNA stability or translation, varying by as much as 7-fold (Figure 3.1C). To control for differences in expression when measuring degradation of

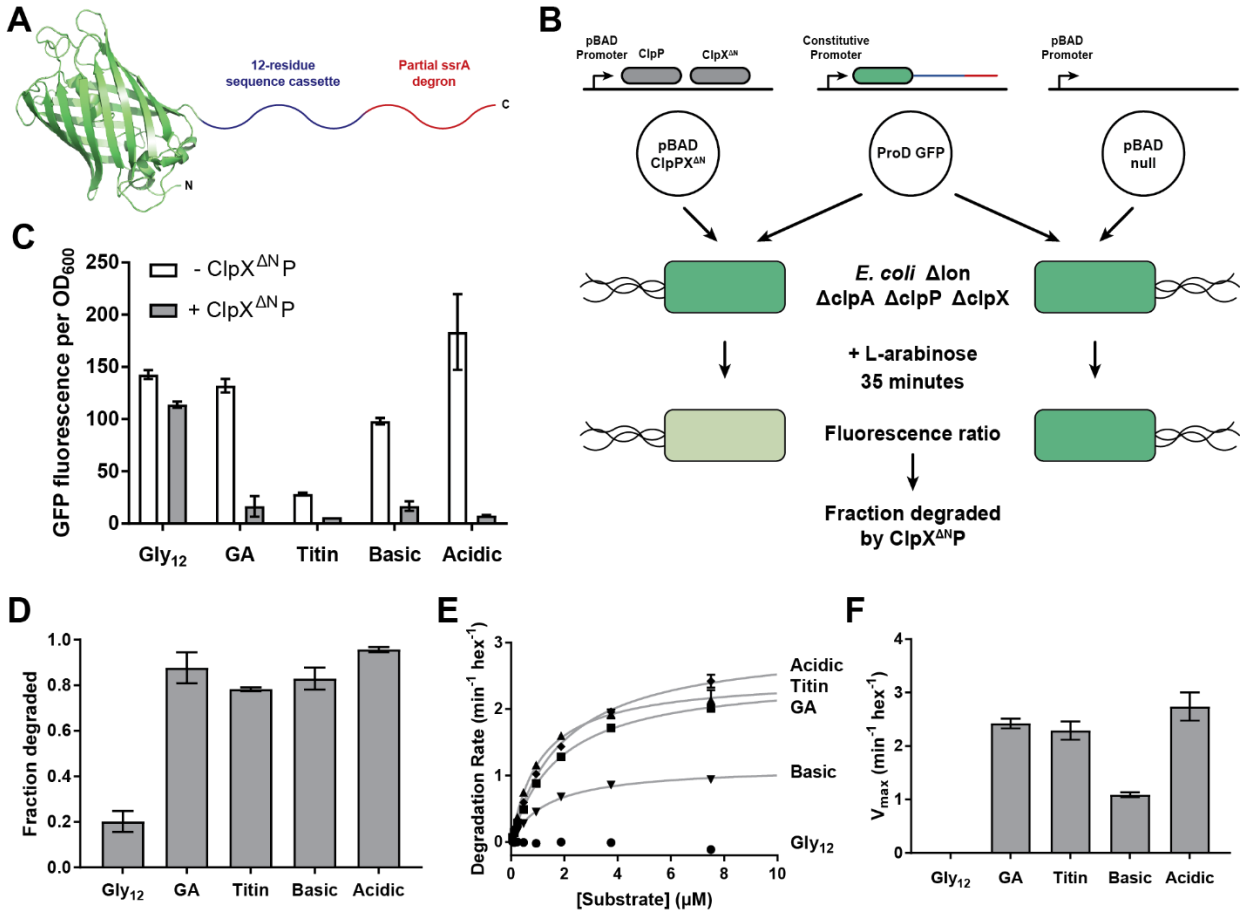


Figure 3.1 – Effects of cassette sequence on GFP unfolding and degradation

(A) Starting at the N terminus, substrates contained residues 1–229 of *A. victoria* GFP (PDB 1GFL, ref. 43), a cassette with 12 variable residues, and a partial ssrA degron. **(B)** Method for measuring intracellular degradation of substrates by $ClpX^{\Delta N}$/ClpP. **(C)** Cellular fluorescence depends upon $ClpX^{\Delta N}$/ClpP expression and cassette sequence (listed in Table 3.1). **(D)** Fraction intracellular degradation for substrates bearing different cassettes. **(E)** Fits of the substrate dependence of degradation *in vitro* to a hyperbolic Michaelis-Menten equation. **(F)** V_{max} values for different substrates. In panels, C–F, values represent averages (± S.D.) of three biological replicates.

different substrates, we normalized measurements in the strain expressing ClpX$^{\Delta N}$P to a strain lacking it (Figure 3.1D, Table 3.1). In these experiments, low-complexity tails rich in acidic or basic residues promoted GFP degradation at levels comparable to a high-complexity sequence derived from the human titin^{I27} domain or a sequence of interspersed glycines and alanines (called GA). By contrast, a cassette sequence of twelve glycines (called Gly_{12}) resulted in poor degradation.

We purified N-terminally His_6-tagged variants of these substrates and performed Michaelis-Menten analysis of steady-state ClpX$^{\Delta N}$P degradation *in vitro* (Figure 3.1E). Degradation with the Gly_{12} cassette was too slow to measure, but the GA and acidic cassettes promoted degradation with V_{max} values similar to the titin sequence (Figure 3.1F, Table 3.2). The basic cassette sequence resulted in an intermediate rate of maximal degradation.

The differences between our results *in vivo* and *in vitro* suggest that the endpoint assay *in vivo* has an upper limit and cannot differentiate rates once the pool of cellular GFP has been degraded. About 20% of the Gly_{12} substrate appeared to be degraded *in vivo*, whereas no degradation was seen *in vitro*. Maturation of the GFP chromophore lags protein folding (Reid and Flynn, 1997), and thus degradation of immature non-fluorescent GFP would not be detected in our cellular assay. It is possible that solutes or macromolecular crowding in the cell enhance ClpXP activity or make GFP easier to unfold. Nevertheless, 12 consecutive glycines inhibit ClpXP unfolding/degradation of GFP both *in vivo* and *in vitro*, whereas other low-complexity sequences do not.

A small stretch of tail residues mediates unfolding grip

We substituted amino acids in the Gly_{12} cassette to identify residues/positions that might improve ClpX grip and thus rates of unfolding and degradation. We first positioned a three-residue Leu-Tyr-Val (LYV) sequence in a sliding window across an otherwise poly-Gly cassette (Figure 3.2A). This tripeptide sequence was selected because its residues are large and hydrophobic, unlike the surrounding Gly residues. Placing the LYV sequence at positions 2-4, 4-6, or 6-8 (numbered relative to the last residue of the folded domain) improved GFP degradation to levels similar to the GA substrate, whereas this tripeptide at positions 8-10 and 10-12 had no substantial effect relative to the Gly_{12} parent (Figure 3.2A, Table 3.1). These results suggest that ClpX grips side chains within the first eight residues of the substrate tail during unfolding.

To examine the contributions of individual residues to grip, we constructed another panel of substrates in which a single Tyr residue was placed at each of the first eight tail positions in otherwise all-glycine cassettes (Figure 3.2B). We then tested degradation *in vivo*. Substrates with a single Tyr at positions 3, 4, and 5 were efficiently degraded, a Tyr at position 2 supported an intermediate level of degradation, and Tyr side chains at other cassette positions supported degradation similar to the Gly_{12} parent (Figure 3.2B, Table 3.1). Thus, four residues appear to contribute the most important grip contacts during unfolding, with tail positions 3-5 being most significant. When we measured degradation of purified substrates *in vitro* (Figure 3.2C, Table 3.2), single Tyr side chains at positions 2-6 facilitated GFP degradation, with the experimental V_{max} values forming a roughly normal distribution centered around position 4. Again, Tyr side chains at positions 3-5 were most important, Tyr residues at the flanking 2 and 6 positions had small effects, and Tyr side chains at positions 1, 7, or 8 had no discernable effect. Importantly, changing

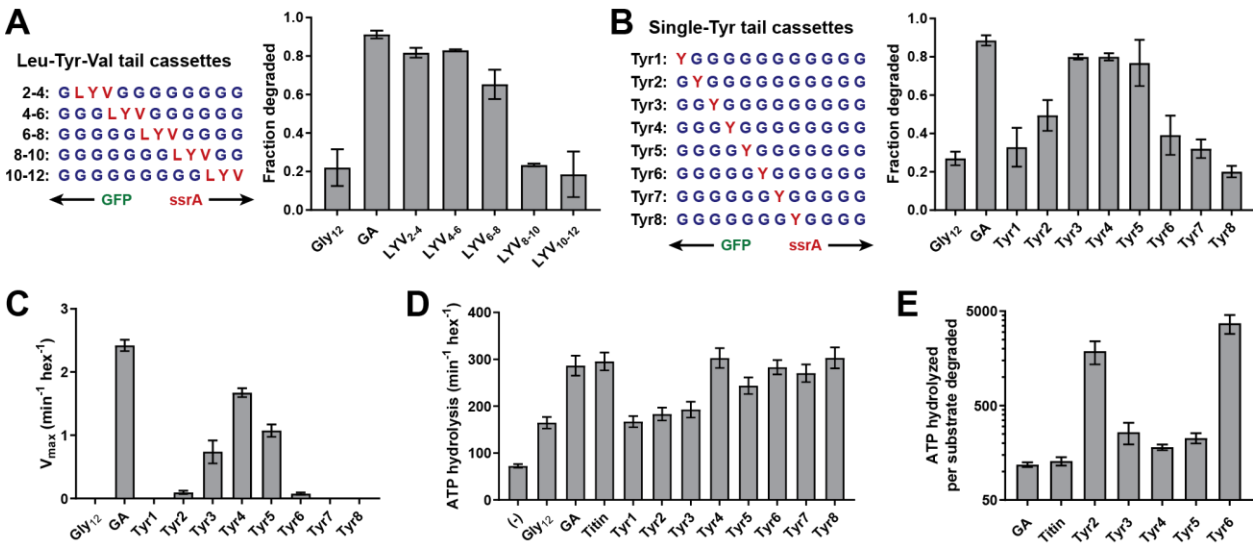


Figure 3.2 – A small subset of tail residues mediate grip during GFP unfolding

(A) Fraction intracellular degradation for substrates with tails containing LYV tripeptides in otherwise all-glycine cassettes. Gly_{12} and GA substrates were included as internal controls. **(B)** Fraction intracellular degradation for substrates with tails containing one tyrosine (Y) in otherwise all-glycine cassettes. Gly_{12} and GA substrates were included as internal controls. **(C)** V_{max} values from Michaelis-Menten analysis of degradation of purified substrates with single-tyrosine cassettes. **(D)** Rates of ATP hydrolysis by $ClpX^{\Delta N}$ (0.1 μM hexamer) in the presence of ClpP (0.3 μM 14-mer) in the absence (–) or presence of different substrates (15 μM monomer). **(E)** ATP cost of degrading substrates with single-tyrosine cassettes. Note that the Y-axis is logarithmic. In all panels, values represent averages (± S.D.) of three biological replicates.

the position of the Tyr side chain altered the maximal rate of unfolding/degradation without substantially affecting K_M for degradation or the ability of substrate to stimulate ATP hydrolysis (Figures 3.2D, 3.3, 3.4, 3.7). As a result, substrates that were degraded slowly also exhibited a high ATP cost for degradation (Figure 3.2E). In combination, these results support a model in which ClpX preferentially grips the side chains of residues at positions 3-5 during GFP unfolding. Moreover, gripping a single Tyr side chain at one of these positions is sufficient for robust unfolding and degradation of GFP.

Side-chain grip preferences

Next, we exploited this system to determine how different types of side chains affect ClpX grip. We constructed substrates in which each of the remaining 18 natural amino acids was placed at position 4 of a cassette with glycines at the other 11 positions. These substrates exhibited a wide range of susceptibility to ClpXP degradation in *E. coli* (Figure 3.5A, Table 3.1). In general, tails containing an aromatic or large/branched hydrophobic side chain (Tyr, Phe, Val, Ile, Leu, or Met) promoted the most efficient unfolding and degradation, whereas small and/or polar side chains were least efficient. The inhibitory effects of polarity and charge on grip were most obvious for side chains with similar shapes. For example, Val was one of the best side chains for grip, whereas the isosteric Thr side chain was very poor (Figure 3.5B). Similarly, a polar Gln side chain resulted in better grip than an isosteric but negatively charged Glu side chain (Figure 3.5B).

We also determined steady-state kinetic parameters for degradation of a subset of purified substrates *in vitro* (Figure 3.5C, Table 3.2). These results largely mirrored results *in vivo*, with mid-

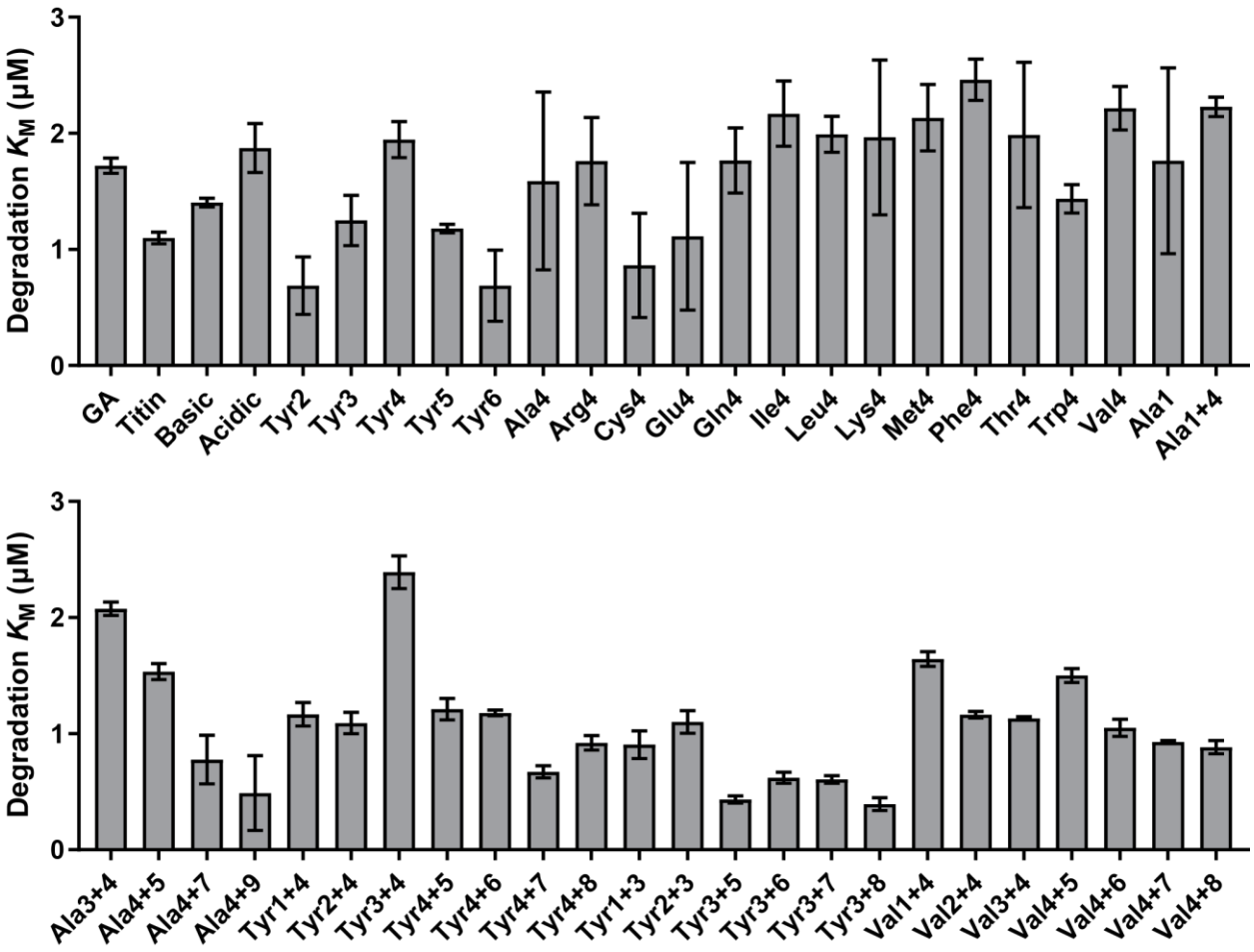


Figure 3.3 – Comparison of K_M values for substrates tested *in vitro*

Comparison of fitted values for K_M for substrate degradation. Values are the average of three biological replicates ± S.D. None of the substrates exhibited a substantial increase in K_M, indicating that differences in degradation rates result from differences in grip rather than in initial substrate recognition.

sized or large hydrophobic and aromatic residues promoting the fastest rates of degradation (Figure 3.5D).

Again, Val supported much better degradation than Thr, and Gln promoted significantly faster degradation than Glu in degradation assays *in vitro* (Figure 3.5E). Further, Ser failed to support GFP degradation while both Ala and Cys facilitated low-level degradation (Figure 3.5E). The Ala-4, Ser-4, Cys-4, Thr-4, Val-4, Glu-4, and Gln-4 substrates at concentrations of 15 μM stimulated the rate of ClpX ATP hydrolysis ~3-4 fold compared to the absence of substrate (Figures 3.3, 3.4). Thus, each substrate binds ClpX well at this concentration, supporting a model in which the large differences in maximal degradation arise from poor grip caused, at least in part, by differences in side-chain charge and polarity. The maximal degradation rates for these substrates (Figure 3.5E) were inversely correlated with their energetic efficiencies of degradation (Figure 3.5F), indicating that poor grip results in non-productive ATP hydrolysis.

Synergistic side-chain interactions promote GFP unfolding

Placing a single alanine at cassette position 4 with glycines at the remaining positions resulted in only marginally better degradation than the Gly_{12} substrate (Figure 3.5A, 3.5C). By contrast, the GA cassette – with alanines at positions 1, 3, 7, 9, and 12 – supported efficient degradation (Figures 3.1D, 3.1F), despite the fact that positions 1, 7, 9, and 12 do not seem to be important determinants of grip (Figure 3.6A). This discrepancy suggested that synergistic interactions between the ClpX pore and multiple side chains might allow substantially better grip. To test this model, we constructed a panel of substrates with one alanine at position 4 and a second alanine at position 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, or 12 (Figure 3.6B). In our cellular assay, alanines at cassette positions

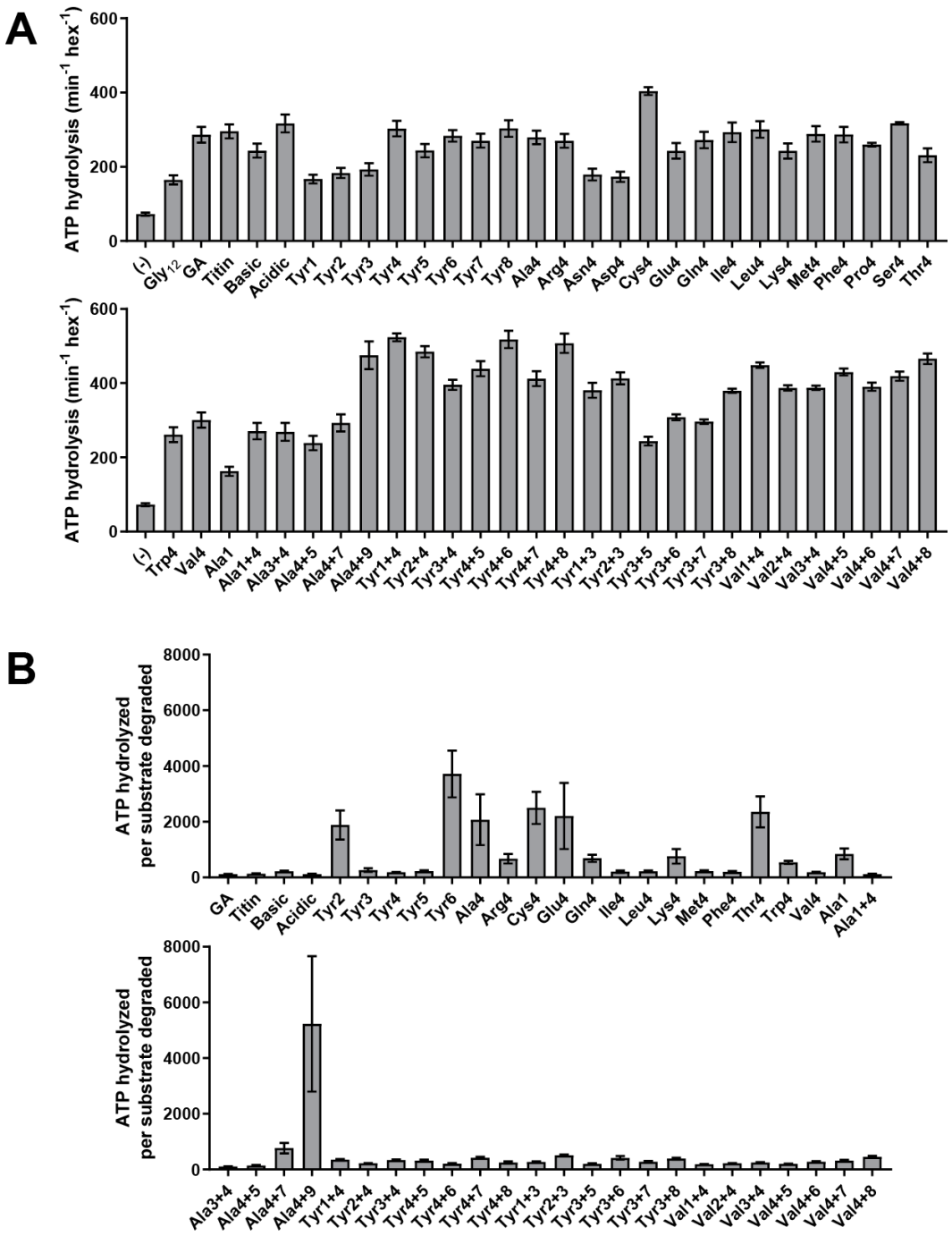


Figure 3.4 – Stimulation of ClpXP ATP hydrolysis by purified substrates
(A) Rates of ATP hydrolysis by $ClpX^{\Delta N}$ (0.1 μM hexamer) in the presence of ClpP (0.3 μM 14-mer) in the absence (–) or presence of different substrates (15 μM monomer). **(B)** ATP cost of degrading substrates. In both panels, values represent averages (± S.D.) of three biological replicates.

1/4, 2/4, 3/4, and 4/5 supported robust degradation, alanines at positions 4/6 and 4/7 facilitated moderate degradation, and alanines at positions 4/8, 4/9, 4/10, 4/11 and 4/12 were little better than the single alanine at position 4 (Figure 3.6B). A single Ala at position 1 supported slightly better degradation *in vivo* than a single Ala at position 4, but the Ala-1/4 substrate was degraded more efficiently (Figure 3.6C). This difference was more pronounced in assays of degradation *in vitro* (Figure 3.6D, Table 3.2). Indeed, V_{max} for degradation of the Ala-1/4 substrate (2.3 ± 0.2 min^{-1}) was more that 10-fold greater than V_{max} for the Ala-1 substrate (0.19 ± 0.04 min^{-1}) or Ala-4 substrate (0.13 ± 0.06 min^{-1}). The non-additivity of these V_{max} values provides direct evidence for synergy in grip.

Because the pore loops of ClpX and other AAA+ motors interact with every other substrate residue in cryo-EM structures (Dong et al., 2019; Gates et al., 2017; Majumder et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017; White et al., 2018), we investigated whether a similar spacing of large side chains enhances grip. We designed panels of substrates with either Tyr or Val fixed at tail position 4 and a second residue of the same type at positions 1, 2, 3, 5, 6, 7, or 8 in an otherwise all-Gly cassette (Figures 3.6E, 3.6F). Among the Tyr substrates, the Tyr-2/4, Tyr-4/6, and Tyr-4/8 substrates promoted faster GFP degradation than the parental Tyr-4 substrate, whereas the other substrates exhibited similar or slower degradation (Table 3.2; Figure 3.6E). It is noteworthy that although the effect of multiple residues was irrelevant for the Ala-4/8 substrate (Figure 3.6B), the Tyr-4/8 substrate was gripped well, suggesting that spacing of Tyr in multiples of two may allow ClpX to grip substrate in a preferred conformation. Among the Val substrates, Val-1/4 and Val-4/5 facilitated slightly faster GFP degradation than the parental Val-4 substrate, Val-2/4 and Val-3/4 were degraded at similar rates to Val-4, and Val-4/6, Val-4/7, and

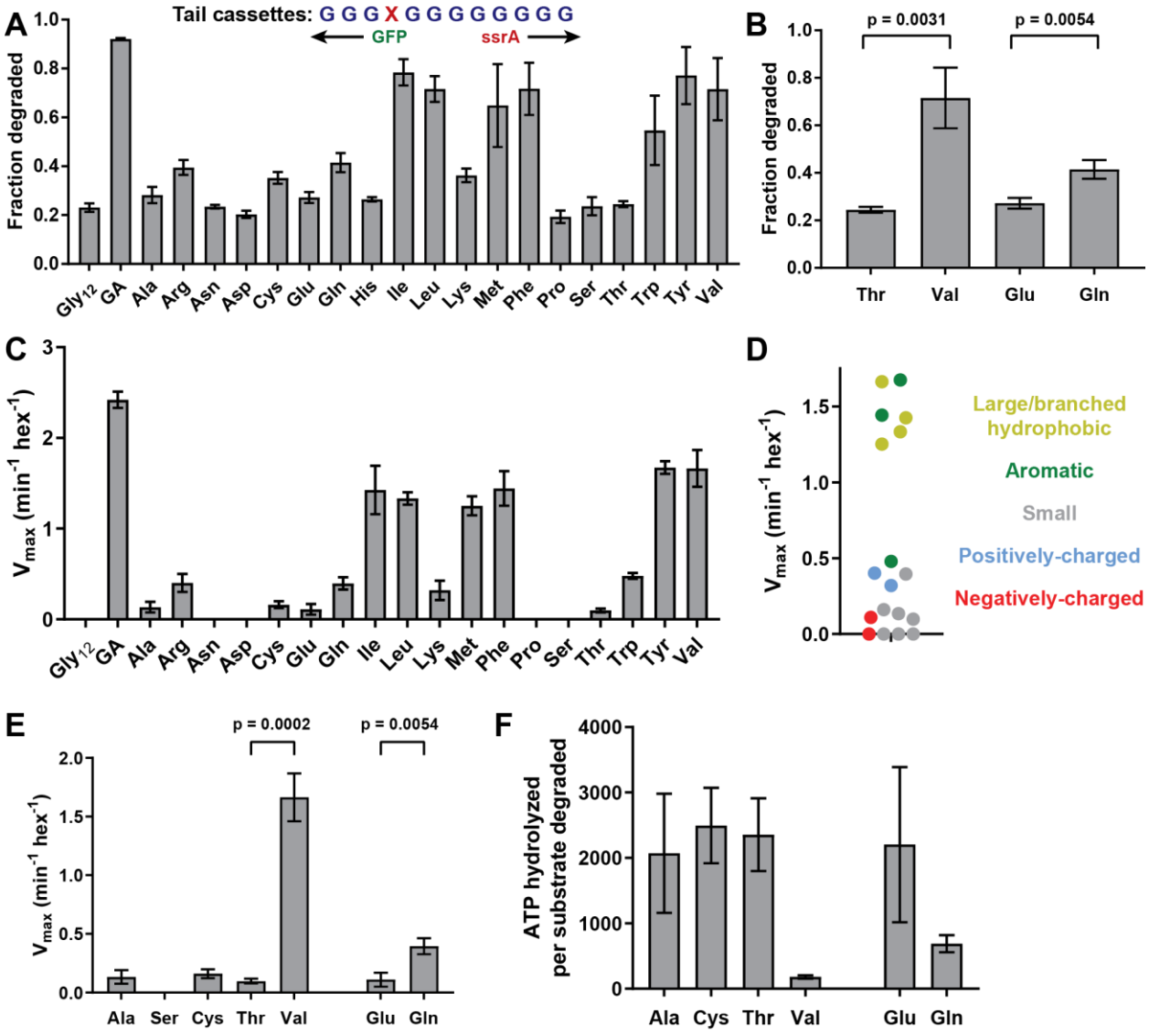


Figure 3.5 – Side-chain grip effects at tail-position 4

(A) In substrates with otherwise all-glycine cassettes, fraction intracellular degradation depends on side-chain identity at tail-position 4. **(B)** Comparison of degradation *in vivo* for substrates with Thr or Val at tail-position 4 or Glu or Gln at tail-position 4 (Student's two-tailed t-test significance; Val/Thr: $t = 6.37$, df = 4; Glu/Gln: $t = 5.47$, df = 4). **(C)** V_{max} values from Michaelis-Menten analysis of degradation of purified substrates. **(D)** Effects of position-4 residues, color-coded by side-chain properties, on V_{max}. **(E)** Comparison of degradation *in vitro* between substrates with Ala, Ser, Cys, Thr, or Val at tail-position 4 or Glu or Gln at tail-position 4 (Student's two-tailed t-test significance; Val/Thr: $t = 13.3$, df = 4; Glu/Gln: $t = 5.49$, df = 4). **(F)** ATP cost of degrading substrates with Ala, Cys, Thr, Val, Glu, or Gln at tail-position 4. With the exception of panel A, where Gly_{12} and GA values represent averages (± S.D.) of nine biological replicates, all values represent three biological replicates.

Val-4/8 were degraded slightly slower (Figure 3.6F; Table 3.2). As the pattern of degradation rates for the branched Val residue (Figure 3.6F) is more similar to Ala (Figures 3.6B, 3.6D) than the aromatic Tyr (Figure 3.6E), it is possible that the ClpX pore-1 loops interact with aromatic residues somewhat differently from non-aromatic residues.

We tested an additional panel of substrates with a Tyr residue at position 3 and a second Tyr at positions 1, 2, 4, 5, 6, 7, or 8 (Figure 3.7). Unlike the other Tyr substrates, these substrates were generally degraded at rates slightly higher than those of the parental Tyr-3 substrate irrespective of spacing (Table 3.2; Figure 3.7). Thus, binding Tyr residues separated by multiples of two residues does not always enhance grip.

DISCUSSION

To unfold target proteins, ClpX and other AAA+ protein remodeling machines use cycles of ATP binding and hydrolysis to pull on the degron tail of a substrate, thereby transmitting force to the native domain, but how these machines interact with individual tail residues during unfolding was poorly understood. Here, we identify and quantify the abilities of different tail residues to promote substrate grip during unfolding by ClpXP. Our experiments are enabled by the observation that placing 12 Gly residues between native GFP and a degron eliminates ClpXP degradation *in vitro* and markedly slows degradation *in vivo*. GFP unfolding/degradation was not inhibited by 12-residue sequences containing mixtures of Gly and Ala (GA cassette); Gly, Lys, and Arg (basic cassette); or Gly, Asp, and Glu (acidic cassette). Compared with these sequences, ClpXP probably grips Gly_{12} poorly because of the absence of β-carbons and distal side-chain atoms or increased

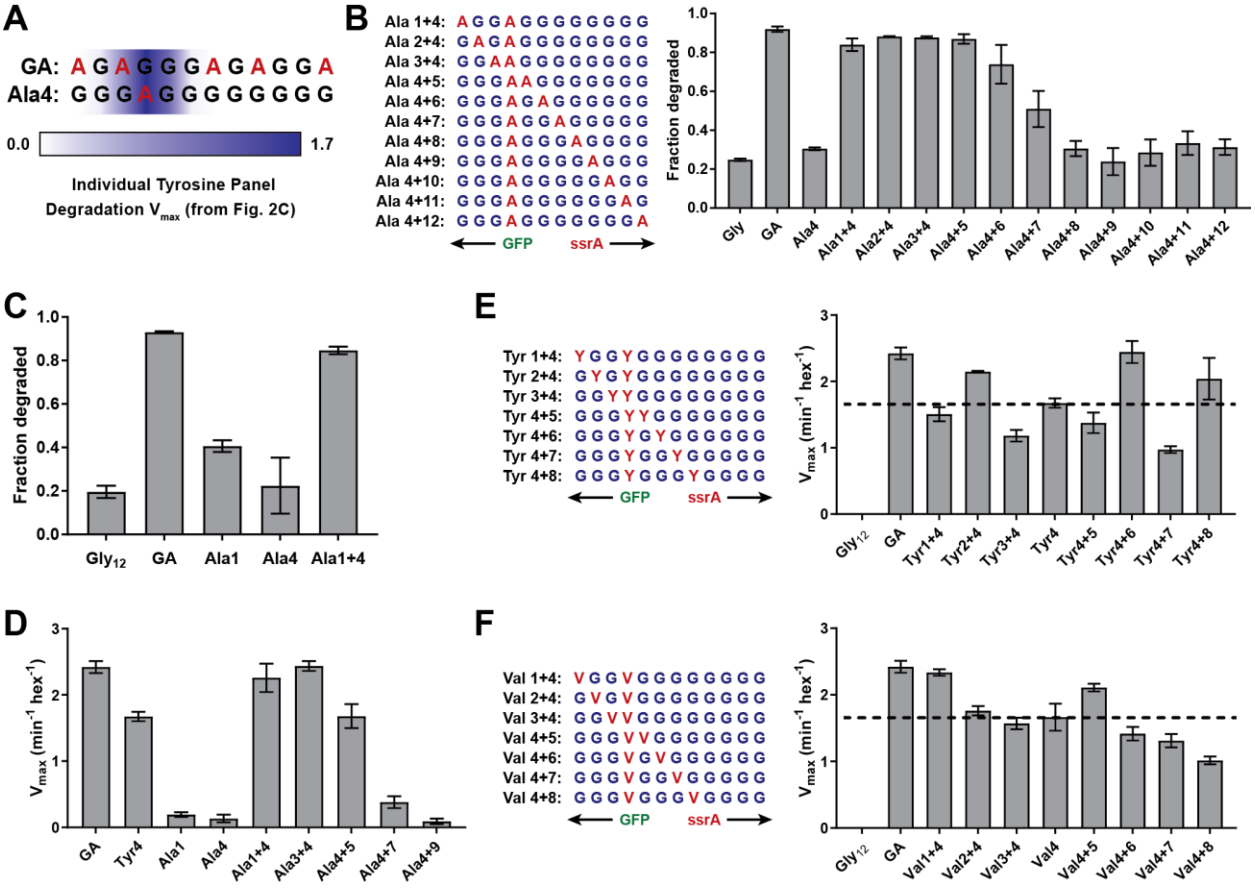


Figure 3.6 – Multiple substrate residues contribute synergistically to grip

(A) GA and Ala-4 cassette sequences. A heatmap of V_{max} values from Figure 3.2C is overlaid to show contribution of single tyrosine residues as each tail position. **(B)** Fraction intracellular degradation of substrates with one alanine at tail-position 4 and a second alanine at a variable position in otherwise all-glycine cassettes. **(C)** Comparison of intracellular degradation for a subset of substrates, including Ala-1. **(D)** V_{max} values from Michaelis-Menten analysis of degradation of purified substrates. **(E) and (F)** Michaelis-Menten V_{max} values for purified substrates with one tyrosine (E) or valine (F) at tail-position 4 and a second tyrosine (E) or valine (F) at each tail position in otherwise all-glycine cassettes. Overlaid dashed lines indicate degradation rate for the parental Tyr-4 (E) or Val-4 (F) substrates. In all panels, values represent averages (± S.D.) of three biological replicates.

backbone flexibility. We use "grip" in a functional rather than strictly physical sense, although the two concepts are undoubtedly related.

Ensemble and single-molecule experiments show that ClpXP can translocate an enormous number of different amino-acid sequences, including long Gly tracts, with only minor velocity differences (Aubin-Tam et al., 2011; Barkow et al., 2009; Cordova et al., 2014; Maillard et al., 2011; Sen et al., 2013). For example, in assays requiring ATP-dependent translocation, ClpXP degraded peptide substrates containing Gly_{10}, $[Val\text{-}Gly]_5$, or $[Phe\text{-}Gly]_5$ sequences at similar rates (Barkow et al., 2009). However, if ClpX can translocate poly-Gly sequences, then why does Gly_{12} inhibit or slow unfolding/degradation? When ClpX pulls on a native protein, Newtonian mechanics dictate that the folded domain resists with an opposing force, which would be absent during translocation of an unstructured polypeptide. Hence, when ATP hydrolysis is coupled to molecular motion during an unfolding power stroke, we imagine that ClpX's grip on the Gly_{12} sequence is insufficient to resist the opposition of the folded domain, causing an unproductive power stroke in which the pore-1 loops slip and fail to advance the substrate tail. In support of this model, we find that poor grip correlates with substantial increases in the ATP cost of degradation for the position-4 and Tyr-scan variants, an indication of slipping and futile power strokes. For example, degradation of one molecule of the Tyr-3, Tyr-4, and Tyr-5 substrates required hydrolysis of an average of ~180-240 ATPs, whereas degradation of the Tyr-2 or Tyr-6 substrates required hydrolysis of ~1900 and ~3700 ATPs, respectively. These findings corroborate a previous report of futile power strokes during unsuccessful unfolding of a difficult substrate by ClpXP (Kraut, 2013). Furthermore, substrate-tail contacts with the axial pore that stimulate ATP hydrolysis by ClpX do not fully overlap with the contacts that determine grip.

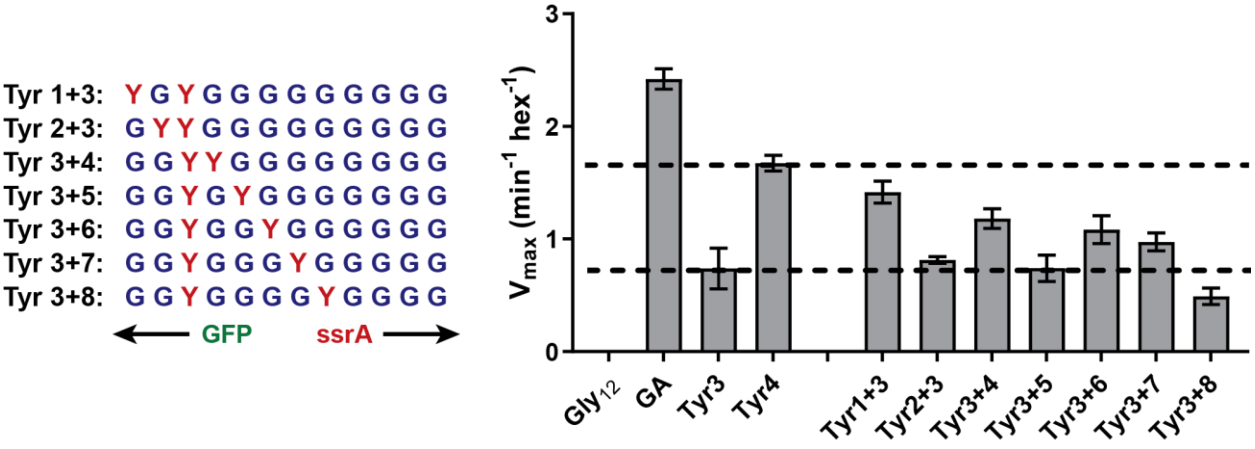


Figure 3.7 – Degradation of Dual-Tyr substrates centered at tail position 3
V_{max} values for degradation from Michaelis-Menten analysis of purified substrates with one tyrosine at tail-position 3 and a second tyrosine at a variable position in otherwise all-glycine cassettes. Relative degradation for substrate tails with a single Tyr residue at position 3 or 4 indicated by dashed lines. Values represent averages (± S.D.) of three biological replicates.

A Tyr-scan experiment shows that tail-position 4 is most important for grip, with flanking positions showing diminishing effects. We expect that amino-acid substitutions at positions 3 and 5 would show side-chain grip trends similar to those observed at position 4. This is not true at tail-position 1, where Tyr did not improve grip but Ala did, perhaps because this part of the tail interacts with different residues in ClpX or the folded GFP domain than downstream positions. A single Ala at tail-position 4 is gripped poorly but a second Ala at certain positions can improve unfolding/degradation. Contacts between the second Ala and the ClpX pore may contribute to stronger grip. Alternatively, the second Ala might affect ClpX contacts made by the first Ala by altering the substrate conformation.

In our GFP substrate with a single Tyr at tail-position 4, this side chain is likely to contact a pore-1 loop close to the folded GFP domain. In an extended chain, four residues would span ~12 Å, a distance that modeling suggests would allow interaction with either the highest or second highest pore-1 loop of ClpX (Figure 3.8A; Puchades et al., 2017; X. Fei, personal communication). This is also an area where the axial channel is most tightly constricted around substrate. The distribution of Tyr-effects at positions 2-6 could reflect interactions with different pore-1 loops or the probability that Tyr side chains at different positions contact one specific pore-1 loop. Optical trapping studies indicate that 5-8 residues are moved by a single ClpXP power stroke (Aubin-Tam et al., 2011; Maillard et al., 2011). Thus, once the Tyr side chain at position 4 is engaged by a pore-1 loop, one successful translocation event probably unfolds GFP. Indeed, although many unsuccessful power strokes and ATP hydrolysis events occur while ClpXP is attempting to unfold a stable domain, hydrolysis of a single ATP ultimately results in unfolding. We find it notable that the Tyr-scan distribution is only two residues wide at half height. Hence, one tyrosine at position

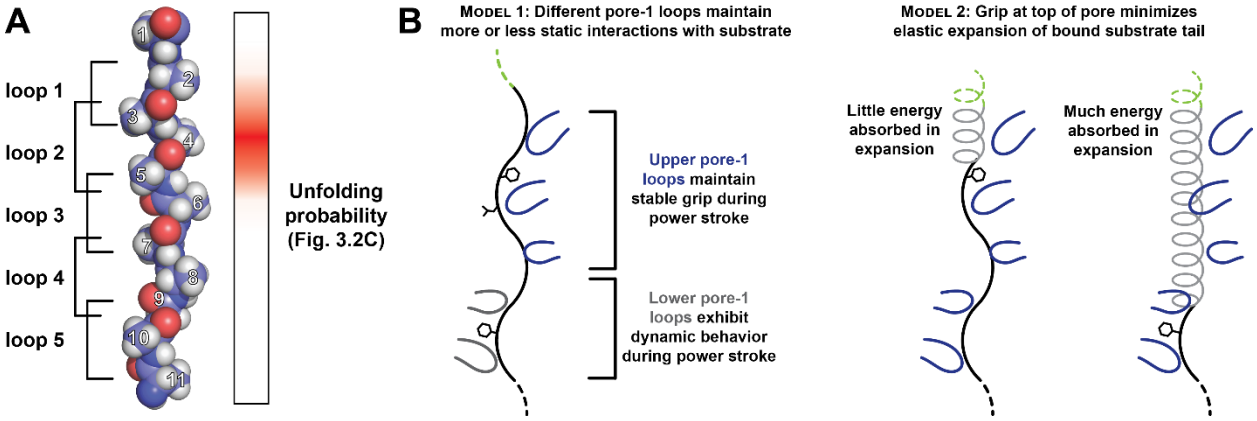


Figure 3.8 – Only a subset of pore-1 loops in ClpX appear to mediate substrate grip
(A) Model of an extended poly-alanine substrate in the axial pore of ClpX and its interactions with different pore-1 loops based on cryo-EM structures of ClpXP (X. Fei, T.A. Bell, B.M. Stinson, S. Jenni, T.A. Baker, S.C. Harrison, and R.T. Sauer, in preparation). Similar loop-substrate interactions are observed in the yeast AAA+ protease Yme1 (Puchades et al., 2017). On the right, a heatmap of V_{max} values from Figure 3.2C is shown. The substrate tail residues are numbered relative to where a folded domain would be expected to sit at the apical surface of the AAA+ ring during unfolding. Tail residues 2-6, which promote strong grip in ClpX, are positioned to interact with the three pore-1 loops at the top of the axial pore. (**B**) Two models for asymmetric contribution of pore-1 loops to substrate grip.

4 mediates robust unfolding, whereas one tyrosine at position 7 has no effect. If a position-4 side chain can contact a pore-1 loop high in the ClpX pore, then a position-7 side chain should be able to contact another pore-1 loop lower in the pore. If this model is correct, then it implies that physical contacts between substrate tail residues and the upper pore-1 loops of ClpX are far more important for grip than interaction with the lower loops.

We can imagine several different mechanisms for the asymmetry in grip between pore-1 loops in the upper and lower sections of the ClpX pore. In one model, the stronger grip of upper pore-1 loops occurs because these loops maintain relatively static interactions with substrate throughout a power stroke (Figure 3.8B, left). Several translocation models have been recently proposed for AAA+ unfoldases in which ATP-bound subunits with pore-1 loops oriented near the top of the pore move together as a rigid unit in response to ATP hydrolysis in a lower subunit (Dong et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017). Furthermore, a previous study demonstrated that pore-1 loop mutations disrupt substrate unfolding most dramatically when they are in neighboring ClpX subunits, consistent with grip mediated by a clustered subset of pore-1 loops (Iosefson et al., 2015). Alternatively, the substrate tail in the pore could absorb some unfolding force through elastic expansion, diverting part of the energy of each power stroke away from unfolding (Figure 3.8B, right). Substrate interactions with the uppermost pore-1 loops would minimize the expansion length of the substrate tail, whereas interactions with lower pore-1 loops would allow the tail to absorb more force.

In an otherwise all-Gly cassette context, we find that Val, Ile, Leu, Met, Phe, and Tyr at tail-position 4 all promote reasonable levels of grip. If we think of poly-Gly as a smooth and relatively

featureless rope, then these larger and generally non-polar side chains can be viewed as knots in the rope that afford better grip. However, grip is not a simple function of side-chain size. For example, Trp supports slower GFP unfolding than the better-gripped residues, suggesting that there may be an upper limit on the size of a side chain that can be efficiently gripped. Polar atoms, especially those close to the peptide backbone of the substrate, weaken grip. For example, Val is one of the best residues in terms of grip, whereas Thr, which differs only by substituting a hydroxyl for a methyl group, barely supports unfolding. Similarly, Ser alone does not support GFP unfolding, but removing the hydroxyl group (Ala) or substituting a less-polar thiol group (Cys) restores low-level unfolding activity. ClpXP may grip polar side chains less tightly because oxygen or nitrogen atoms bearing partial or full charges are not fully solvated when they are in productive contact with a pore-1 loop and thus incur an energetic penalty. Our finding that large hydrophobic and aromatic side chains are gripped well by ClpX is consistent with a model in which van der Waal's or hydrophobic interactions between the pore-1 loops and specific side chains in the tail are largely responsible for grip.

Several recent cryo-EM structures of AAA+ proteases and protein-remodeling motors reveal a spiral arrangement of subunits in which aromatic residues in the pore-1 loops interact with substrate side chains spaced two residues apart (Dong et al., 2019; Gates et al., 2017; Majumder et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017; White et al., 2018). Our observation that substrate tails with two tyrosines can in some cases specifically enhance grip when spaced by a multiple of two residues is consistent with these structures. However, substrates with two Val residues do not exhibit the same periodic grip enhancement as Tyr, and multiple Ala residues together promote strong grip regardless of their relative spacing. It is clear that nonspecific

interactions between the axial pore and substrate side chains are sufficient to promote strong grip independent of precise pore-1 loop intercalation. In specific cases, periodic side chain intercalation could enhance grip for aromatic side chains through the establishment of π-stacking networks with the pore-1 loop Tyr residues, possibly by optimizing the bound substrate conformation.

A recent cryo-EM structure of the AAA+ motor NSF, which disassembles SNARE complexes following vesicle fusion, contains well-resolved density for substrate side chains, revealing interactions with the pore-1 loop tyrosine (White et al., 2018). Although NSF disassembles SNARE complexes in a single round of ATP turnover in a mechanism distinct from ClpX (Ryu et al., 2015), the structural similarities between NSF and many AAA+ unfoldase-proteases suggests a common mode of substrate interaction and grip. The assembled SNARE complex is remarkably stable, and NSF likely requires strong grip to disassemble the complex. Consistent with our biochemical observations for ClpX, this structure indicates that the strongest substrate contacts are formed with pore-1 loops high in the upper ring of NSF, and that two substrate residues (Met and Leu) that are gripped well by ClpX participate in these interactions.

A previous study found that a Gly_{15} sequence placed between GFP and an ssrA tag did not slow ClpXP degradation (Barkow et al., 2009). However, their substrate contained six additional residues (Thr^1-His^2-Gly^3-Met^4-Asp^5-Glu^6) between folded GFP and the Gly_{15} sequence. As we find that Met at position 4 supports robust unfolding, it is likely that interactions with the extra sequence mediate unfolding of the Gly_{15} substrate. Our findings suggest that evolutionary placement of tail residues that are gripped well by ClpX may tune degradation of substrates that unfold non-cooperatively or that have multiple folded domains. Complete, processive unfolding

of multi-domain substrates depends critically on interactions between ClpX and the peptide-tail remnants from unfolding and degradation of the previous domain. For example, ClpXP degradation of Domain III of *C. crescentus* DnaX is inhibited by a Gly-rich sequence between Domains III and IV, which acts as a partial processing mechanism essential for DNA replication (Vass and Chien, 2013). Gly-rich tracts also inhibit unfolding/degradation of *E. coli* DHFR, although multiple alanines in the tail do not improve degradation of this substrate (Too et al., 2013). Thus, ClpX unfolding of different native substrates probably requires different degrees of grip strength, which could be mediated by more and/or better-gripped amino acids adjacent to the folded domain.

ClpXP contains just two types of subunits, whereas the 26S proteasome consists of more than 30 subunit types (Budenholzer et al., 2017). Nevertheless, our work is reminiscent of and reinforces studies of proteasomal degradation by Matouschek and colleagues. For example, they find that low-complexity sequences primarily composed of Gly, Ser, or Thr residues can inhibit proteasomal degradation (Tian et al., 2005); these residues individually are also insufficient to promote ClpXP degradation of GFP. Similarly, sequences that include Phe and Tyr residues can improve or rescue degradation by both the proteasome and ClpXP. These similarities may arise because the Rpt_{1-6} unfolding ring of the proteasome, despite containing six distinct subunits, has pore-1 loops very similar to those of ClpX. Given the structural similarities between many AAA+ protein remodeling machines, we expect that the principles underlying grip in ClpX reflect those of the broader family.

Substrate	Variable Tail Sequence	Fraction Degraded *in vivo*
Gly_{12}	GGGG GGGG GGGG	0.20 ± 0.05
GA	AGAG GGAG AGGA	0.88 ± 0.07
Titin	HLGL IEVE KPLY	0.78 ± 0.01
Basic	GKGR GKGR GKGR	0.83 ± 0.05
Acidic	GEGD GEGD GEGD	0.96 ± 0.01
LYV_{2-4}	GLYV GGGG GGGG	0.82 ± 0.03
LYV_{4-6}	GGGL YVGG GGGG	0.83 ± 0.01
LYV_{6-8}	GGGG GLYV GGGG	0.65 ± 0.08
LYV_{8-10}	GGGG GGGL YVGG	0.23 ± 0.01
LYV_{10-12}	GGGG GGGG GLYV	0.2 ± 0.1
Tyr1	YGGG GGGG GGGG	0.3 ± 0.1
Tyr2	GYGG GGGG GGGG	0.49 ± 0.08
Tyr3	GGYG GGGG GGGG	0.80 ± 0.01
Tyr4	GGGY GGGG GGGG	0.80 ± 0.02
Tyr5	GGGG YGGG GGGG	0.8 ± 0.1
Tyr6	GGGG GYGG GGGG	0.4 ± 0.1
Tyr7	GGGG GGYG GGGG	0.32 ± 0.05
Tyr8	GGGG GGGY GGGG	0.20 ± 0.03
Ala4	GGGA GGGG GGGG	0.28 ± 0.03
Arg4	GGGR GGGG GGGG	0.39 ± 0.03
Asn4	GGGN GGGG GGGG	0.23 ± 0.01
Asp4	GGGD GGGG GGGG	0.20 ± 0.02
Cys4	GGGC GGGG GGGG	0.35 ± 0.02
Glu4	GGGE GGGG GGGG	0.27 ± 0.02
Gln4	GGGQ GGGG GGGG	0.41 ± 0.04
His4	GGGH GGGG GGGG	0.26 ± 0.01
Ile4	GGGI GGGG GGGG	0.78 ± 0.05
Leu4	GGGL GGGG GGGG	0.7 ± 0.1
Lys4	GGGK GGGG GGGG	0.36 ± 0.02
Met4	GGGM GGGG GGGG	0.6 ± 0.2
Phe4	GGGF GGGG GGGG	0.7 ± 0.1
Pro4	GGGP GGGG GGGG	0.19 ± 0.03
Ser4	GGGS GGGG GGGG	0.24 ± 0.04
Thr4	GGGT GGGG GGGG	0.24 ± 0.01
Trp4	GGGW GGGG GGGG	0.5 ± 0.1
Val4	GGGV GGGG GGGG	0.7 ± 0.1
Ala1	AGGG GGGG GGGG	0.41 ± 0.03
Ala1+4	AGGA GGGG GGGG	0.84 ± 0.03

Ala2+4	GAGA GGGG GGGG	0.88 ± 0.01
Ala3+4	GGAA GGGG GGGG	0.88 ± 0.01
Ala4+5	GGGA AGGG GGGG	0.87 ± 0.02
Ala4+6	GGGA GAGG GGGG	0.7 ± 0.1
Ala4+7	GGGA GGAG GGGG	0.51 ± 0.09
Ala4+8	GGGA GGGA GGGG	0.31 ± 0.04
Ala4+9	GGGA GGGG AGGG	0.24 ± 0.07
Ala4+10	GGGA GGGG GAGG	0.29 ± 0.07
Ala4+11	GGGA GGGG GGAG	0.33 ± 0.06
Ala4+12	GGGA GGGG GGGA	0.31 ± 0.04

Table 3.1 – Degradation of variable-tail substrates in the bacterial cytoplasm
Sequences of all substrate tails tested and the extent of degradation by $ClpX^{\Delta N}P$ in *E. coli* after 35 minutes. For substrates tested in multiple panels, the value presented is from the panel in which they first appear. Values are the average of three biological replicates ± S.D.

Substrate	V_{max} (min^{-1} hex^{-1})	K_M (μM)
Gly_{12}	*No fit*	
GA	2.4 ± 0.1	1.7 ± 0.1
Titin	2.3 ± 0.2	1.1 ± 0.1
Basic	1.1 ± 0.1	1.4 ± 0.1
Acidic	2.7 ± 0.3	1.9 ± 0.2
Tyr1	*No fit*	
Tyr2	0.10 ± 0.03	0.7 ± 0.2
Tyr3	0.7 ± 0.2	1.3 ± 0.2
Tyr4	1.7 ± 0.1	1.9 ± 0.2
Tyr5	1.1 ± 0.1	1.2 ± 0.1
Tyr6	0.08 ± 0.02	0.7 ± 0.3
Tyr7	*No fit*	
Tyr8	*No fit*	
Ala4	0.13 ± 0.06	1.6 ± 0.8
Arg4	0.4 ± 0.1	1.8 ± 0.4
Asn4	*No fit*	
Asp4	*No fit*	
Cys4	0.16 ± 0.04	0.9 ± 0.4
Glu4	0.11 ± 0.06	1.1 ± 0.6
Gln4	0.40 ± 0.07	1.8 ± 0.3
Ile4	1.4 ± 0.3	2.2 ± 0.3
Leu4	1.3 ± 0.1	2.0 ± 0.2
Lys4	0.3 ± 0.1	2.0 ± 0.7
Met4	1.3 ± 0.1	2.1 ± 0.3
Phe4	1.4 ± 0.2	2.5 ± 0.2
Pro4	*No fit*	
Ser4	*No fit*	
Thr4	0.10 ± 0.02	2.0 ± 0.6
Trp4	0.48 ± 0.03	1.4 ± 0.1
Val4	1.7 ± 0.2	2.2 ± 0.2
Ala1	0.19 ± 0.04	1.8 ± 0.8
Ala1+4	2.3 ± 0.2	2.2 ± 0.1
Ala3+4	2.4 ± 0.1	2.1 ± 0.1
Ala4+5	1.7 ± 0.2	1.5 ± 0.1
Ala4+7	0.38 ± 0.09	0.8 ± 0.2
Ala4+9	0.09 ± 0.04	0.5 ± 0.3
Tyr1+4	1.5 ± 0.1	1.2 ± 0.1
Tyr2+4	2.1 ± 0.1	1.1 ± 0.1

Tyr3+4	1.2 ± 0.1	2.4 ± 0.1
Tyr4+5	1.4 ± 0.2	1.2 ± 0.1
Tyr4+6	2.4 ± 0.2	1.2 ± 0.1
Tyr4+7	1.0 ± 0.1	0.67 ± 0.05
Tyr4+8	2.0 ± 0.3	0.92 ± 0.06
Tyr1+3	1.4 ± 0.1	0.9 ± 0.1
Tyr2+3	0.81 ± 0.03	1.1 ± 0.1
Tyr3+5	0.7 ± 0.1	0.43 ± 0.03
Tyr3+6	1.1 ± 0.1	0.62 ± 0.05
Tyr3+7	0.97 ± 0.08	0.61 ± 0.03
Tyr3+8	0.49 ± 0.07	0.39 ± 0.06
Val1+4	2.3 ± 0.1	1.6 ± 0.1
Val2+4	1.8 ± 0.1	1.2 ± 0.1
Val3+4	1.6 ± 0.1	1.1 ± 0.1
Val4+5	2.1 ± 0.1	1.5 ± 0.1
Val4+6	1.4 ± 0.1	1.1 ± 0.1
Val4+7	1.3 ± 0.1	0.93 ± 0.01
Val4+8	1.0 ± 0.1	0.88 ± 0.06

Table 3.2 – Degradation of purified variable-tail substrates *in vitro*

Fitted parameters from Michaelis-Menten analysis of substrate degradation by $ClpX^{\Delta N}P$. *No fit* – substrate degradation too slow to be accurately fit. Values are the average of three biological replicates ± S.D.

Substrate	ATP hydrolysis (min^{-1} hex^{-1})	ATP hydrolyzed per substrate degraded
ClpXP alone	73 ± 4	
Gly_{12}	160 ± 10	
GA	290 ± 20	120 ± 10
Titin	300 ± 20	130 ± 10
Basic	240 ± 20	220 ± 20
Acidic	320 ± 20	120 ± 10
Tyr1	170 ± 10	
Tyr2	180 ± 10	1900 ± 500
Tyr3	190 ± 20	260 ± 70
Tyr4	300 ± 20	180 ± 10
Tyr5	240 ± 20	230 ± 30
Tyr6	280 ± 20	3700 ± 800
Tyr7	270 ± 20	
Tyr8	300 ± 20	
Ala4	280 ± 20	2100 ± 900
Arg4	270 ± 20	700 ± 200
Asn4	180 ± 20	
Asp4	170 ± 10	
Cys4	400 ± 10	2500 ± 600
Glu4	240 ± 20	2000 ± 1000
Gln4	270 ± 20	700 ± 100
Ile4	290 ± 30	200 ± 40
Leu4	300 ± 20	230 ± 20
Lys4	240 ± 20	800 ± 300
Met4	290 ± 20	230 ± 30
Phe4	290 ± 20	200 ± 30
Pro4	260 ± 10	
Ser4	320 ± 10	
Thr4	230 ± 20	2400 ± 600
Trp4	260 ± 20	540 ± 50
Val4	300 ± 20	180 ± 30
Ala1	160 ± 10	800 ± 200
Ala1+4	270 ± 20	120 ± 10
Ala3+4	270 ± 20	110 ± 10
Ala4+5	240 ± 20	140 ± 20
Ala4+7	290 ± 20	800 ± 200
Ala4+9	480 ± 40	5000 ± 2000

Tyr1+4	520 ± 10	350 ± 30
Tyr2+4	480 ± 20	230 ± 10
Tyr3+4	400 ± 10	330 ± 30
Tyr4+5	440 ± 20	320 ± 40
Tyr4+6	520 ± 20	210 ± 20
Tyr4+7	410 ± 20	420 ± 30
Tyr4+8	510 ± 30	250 ± 40
Tyr1+3	380 ± 20	270 ± 20
Tyr2+3	410 ± 20	510 ± 30
Tyr3+5	240 ± 10	210 ± 20
Tyr3+6	310 ± 10	420 ± 70
Tyr3+7	300 ± 10	270 ± 30
Tyr3+8	380 ± 10	390 ± 30
Val1+4	450 ± 10	190 ± 10
Val2+4	390 ± 10	220 ± 10
Val3+4	390 ± 10	250 ± 10
Val4+5	430 ± 10	200 ± 10
Val4+6	390 ± 10	280 ± 20
Val4+7	420 ± 10	320 ± 30
Val4+8	470 ± 10	460 ± 30

Table 3.3 – Stimulation of ClpXP ATP hydrolysis by purified substrates
Values are averages of three biological replicates ± S.D.

METHODS

Plasmid and strain construction

An expression plasmid containing *E. coli* ClpP and *E. coli* $ClpX^{\Delta N}$ was constructed by cloning ClpP into the open reading frame downstream of the pBAD promoter in pBAD18 (Guzman et al., 1995). A second ribosome binding site (5'-CAAGGAGAATAACG-3') and the $ClpX^{\Delta N}$ coding sequence (residues 62-424) was added downstream of the ClpP stop codon to produce a polycistronic expression construct. GFP substrates for cytoplasmic degradation assays were cloned downstream of the constitutive insulated ProD promoter in pSB3C5 (Davis et al., 2011). His_6-GFP substrates for purification were cloned into a pET4b derivative downstream of the pT7 promoter. For all substrates, the 12-residue variable cassette was encoded on an oligonucleotide and introduced upstream of a partial ssrA degron (Gly-Ser-Glu-Asn-Tyr-Ala-Leu-Ala-Ala) using PCR mutagenesis. The seven C-terminal residues of this degron are identical to those of the ssrA tag, but we removed the N-terminal part of the ssrA tag to preclude potential SspB inhibition (Hersch et al., 2004).

T7 Express *ΔclpA ΔclpP ΔclpX* was generated from the *E. coli* strain T7 Express (New England Biolabs). The bicistronic clpP–clpX locus was removed by using lambda *red* recombineering (Yu et al., 2000) to replace the locus with an FRT-Kan^R cassette, which was subsequently removed by FLP recombinase expression. ClpA::FRT-Kan^R was then transduced into this strain with P1 phage from a ClpA::FRT-Kan^R strain in the Keio collection (Baba et al., 2006), and the resistance marker was again removed with FLP recombinase. Modification of the correct loci was verified by PCR at each step in strain construction.

Protein expression and purification

His_6-GFP-cassette-ssrA constructs were expressed as described (Kim et al., 2000) and purified by Ni-NTA affinity, Source 15Q anion exchange, and Superdex 200 size-exclusion chromatography. Purified substrates were assessed to be >99% pure by SDS-PAGE and were stored in 25 mM HEPES-KOH (pH 7.5), 150 mM KCl, 10% glycerol, and 500 μM dithiothreitol.

***Degradation assays* in vivo**

The ProD-GFP plasmid encoding each substrate was transformed into T7 Express *ΔclpA ΔclpP ΔclpX* cells carrying either pBAD18(ClpP/$ClpX^{\Delta N}$) or pBAD18(null). After overnight growth at 30 °C on LB agar plates supplemented with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol, single colonies were picked into 5 mL of the same medium and antibiotics and cultures were grown overnight at 30 °C. At the start of degradation assays, 50 μL of culture of either the ClpP/$ClpX^{\Delta N}$ expression strain or the null control strain for each substrate was inoculated into fresh 5 mL LB plus antibiotics and grown at 37 °C to OD_{600} 0.7-1.0. The cultures were then centrifuged; resuspended at OD_{600} 1.2 in fresh media plus antibiotics; and 500 μL was added to 1 mL of fresh media plus antibiotics supplemented with 120 mM L-arabinose, for a final concentration of 80 mM L-arabinose and OD_{600} of 0.4. After 35 min of growth at 37 °C, 1 mL of culture was removed, centrifuged, and resuspended in 600 μL of phosphate buffered saline (pH 7.4). Three 150 μL technical replicates of resuspended cells were transferred to wells of a clear-bottom black 96-well plate (Greiner). Both the GFP fluorescence of the cell resuspension (excitation 467 nm, emission 511 nm) and the optical density (absorbance 600 nm) were measured on a SpectraMax M5 plate

reader (Molecular Devices). The GFP fluorescence for each $ClpX^{\Delta N}P$ sample and control sample was divided by the measured cell density to give normalized $fl_{protease}$ and $fl_{control}$ values respectively. Fraction degraded was calculated as:

$$1 - (fl_{protease} / fl_{control})$$

Each degradation assay was performed independently in multiple biological replicates, and the calculated value of fraction degraded was averaged across biological replicates. No obvious outliers were observed, and all values were included in the subsequent analysis.

***Biochemical assays* in vitro**

Degradation assays were performed at 37 °C in 25 mM HEPES-KOH (pH 7.5), 5 mM $MgCl_2$, 200 mM KCl, 10% glycerol, with 0.1 μM $ClpX^{\Delta N}$ (hexamer), 0.3 μM ClpP (14-mer), 5 mM ATP, 32 mM creatine phosphate (Roche), and 0.08 mg/mL creatine kinase (Millipore-Sigma). For the Thr-4 substrate, assays were performed with 0.5 μM $ClpX^{\Delta N}$ (hexamer) and 1.5 μM ClpP (14-mer) to measure degradation rates more accurately and facilitate comparison with substrates that were degraded more rapidly. Degradation rates were measured by decrease in fluorescence (excitation 467 nm; emission 511 nm) in 20 μL reactions on a SpectraMax M5 plate reader. To control for signal loss from photobleaching, a parallel set of reactions was measured for each substrate without $ClpX^{\Delta N}$ or ClpP, and changes in GFP fluorescence in this experiment were subtracted from those in the degradation reaction. Each measurement included three technical replicates measured together in parallel, and the average values of these replicates were fit to a hyperbolic equation to determine K_M and V_{max}. Three independently-conducted biological replicates were performed in this manner for each substrate to determine average values (± S.D.) for K_M and V_{max}. No obvious outliers were observed, and all values were included in the subsequent analysis.

ATP hydrolysis rates were measured using a coupled-NADH oxidation assay as described (Martin et al., 2005). Degradation efficiency (ATP hydrolyzed per substrate degraded) was calculated by dividing the rate of ATP hydrolysis at a near saturating substrate concentration (15 μM) by V_{max} for substrate degradation.

ACKNOWLEDGMENTS

This work was supported by US National Institutes of Health (NIH) grant GM-101988 (R.T.S.). T.A. Bell was supported in part by US NIH grant 5T32GM-007287. T.A. Baker is an employee of the Howard Hughes Medical Institute. We thank X. Fei (MIT) for providing information on the ClpXP cryo-EM structure and present and past members of the Sauer and Baker labs for helpful conversations and feedback on the manuscript.

REFERENCES

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, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. *2*, 2006.0008.

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

Barkow, S.R., Levchenko, I., Baker, T.A., and Sauer, R.T. (2009). Polypeptide Translocation by the AAA+ ClpXP Protease Machine. Chem. Biol. *16*, 605–612.

Budenholzer, L., Cheng, C.L., Li, Y., and Hochstrasser, M. (2017). Proteasome Structure and Assembly. J. Mol. Biol. *429*, 3500–3524.

Cordova, J.C., Olivares, A.O., Shin, Y., Stinson, B.M., Calmat, S., Schmitz, K.R., Aubin-Tam, M.-E., Baker, T.A., Lang, M.J., and Sauer, R.T. (2014). Stochastic but Highly Coordinated Protein Unfolding and Translocation by the ClpXP Proteolytic Machine. Cell *158*, 647–658.

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*, 30090–30100.

Davis, J.H., Rubin, A.J., and Sauer, R.T. (2011). Design, construction and characterization of a set of insulated bacterial promoters. Nucleic Acids Res. *39*, 1131–1141.

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.

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

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*, 671–683.

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*, 273–279.

Glynn, S.E. (2017). Multifunctional Mitochondrial AAA Proteases. Front. Mol. Biosci. *4*.

Glynn, S.E., Martin, A., Nager, A.R., Baker, T.A., and Sauer, R.T. (2009). Structures of Asymmetric ClpX Hexamers Reveal Nucleotide-Dependent Motions in a AAA+ Protein-Unfolding Machine. Cell *139*, 744–756.

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*, 1338–1347.

Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. *177*, 4121–4130.

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*, 12136–12141.

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.

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*, 1720–1729.

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

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*, 990–993.

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.-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*, 639–648.

Kraut, D.A. (2013). Slippery Substrates Impair ATP-dependent Protease Function by Slowing Unfolding. J. Biol. Chem. *288*, 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*, 1444–1453.

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. *94*, 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*, 2248–2254.

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*, 459–469.

Majumder, P., Rudack, T., Beck, F., Danev, R., Pfeifer, G., Nagy, I., and Baumeister, W. (2019). Cryo-EM structures of the archaeal PAN-proteasome reveal an around-the-ring ATPase cycle. Proc. Natl. Acad. Sci. *116*, 534–539.

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

Martin, A., Baker, T.A., and Sauer, R.T. (2008a). Diverse Pore Loops of the AAA+ ClpX Machine Mediate Unassisted and Adaptor-Dependent Recognition of ssrA-Tagged Substrates. Mol. Cell *29*, 441–450.

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

Monroe, N., Han, H., Shen, P.S., Sundquist, W.I., and Hill, C.P. (2017). Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. ELife *6*, e24487.

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*, 4–16.

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*, 33–44.

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 *273*, 1392–1395.

Park, E., Rho, Y.M., Koh, O., Ahn, S.W., Seong, I.S., Song, J.-J., Bang, O., Seol, J.H., Wang, J., Eom, S.H., et al. (2005). Role of the GYVG Pore Motif of HslU ATPase in Protein Unfolding and Translocation for Degradation by HslV Peptidase. J. Biol. Chem. *280*, 22892–22898.

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*, eaav0725.

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.

Reid, B.G., and Flynn, G.C. (1997). Chromophore Formation in Green Fluorescent Protein. Biochemistry *36*, 6786–6791.

Ryu, J.-K., Min, D., Rah, S.-H., Kim, S.J., Park, Y., Kim, H., Hyeon, C., Kim, H.M., Jahn, R., and Yoon, T.-Y. (2015). Spring-loaded unraveling of a single SNARE complex by NSF in one round of ATP turnover. Science *347*, 1485–1489.

Sauer, R.T., and Baker, T.A. (2011). AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annu. Rev. Biochem. *80*, 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*, 607–615.

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*, 636–646.

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*, 137–142.

Siddiqui, S.M., Sauer, R.T., and Baker, T.A. (2004). Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. Genes Dev. *18*, 369–374.

Singh, S.K., Rozycki, J., Ortega, J., Ishikawa, T., Lo, J., Steven, A.C., and Maurizi, M.R. (2001). Functional Domains of the ClpA and ClpX Molecular Chaperones Identified by Limited Proteolysis and Deletion Analysis. J. Biol. Chem. *276*, 29420–29429.

Tian, L., Holmgren, R.A., and Matouschek, A. (2005). A conserved processing mechanism regulates the activity of transcription factors Cubitus interruptus and NF-κB. Nat. Struct. Mol. Biol. *12*, 1045–1053.

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*, 13243–13257.

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*, 48–53.

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

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.

Wojtyra, U.A., Thibault, G., Tuite, A., and Houry, W.A. (2003). The N-terminal Zinc Binding Domain of ClpX Is a Dimerization Domain That Modulates the Chaperone Function. J. Biol. Chem. *278*, 48981–48990.

Yamada-Inagawa, T., Okuno, T., Karata, K., Yamanaka, K., and Ogura, T. (2003). Conserved Pore Residues in the AAA Protease FtsH Are Important for Proteolysis and Its Coupling to ATP Hydrolysis. J. Biol. Chem. *278*, 50182–50187.

Yang, F., Moss, L.G., and Phillips, G.N. (1996). The molecular structure of green fluorescent protein. Nat. Biotechnol. *14*, 1246–1251.

Yu, D., Ellis, H.M., Lee, E.-C., Jenkins, N.A., Copeland, N.G., and Court, D.L. (2000). An efficient recombination system for chromosome engineering in Escherichia coli. Proc. Natl. Acad. Sci. *97*, 5978–5983.

Chapter IV

Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate

This chapter has been submitted for publication as:

Fei, X., Bell, Tristan A., Jenni, S., Stinson, B.M., Baker, Tania A., Harrison, Stephen C. and Sauer, Robert T. (2019). Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate.

X.F., S.J., and R.T.S. performed cryo-EM experiments, calculations, and/or model building and refinement. T.A. Bell and B.M.S. performed biochemical experiments. S.P.H., T.A. Baker, and R.T.S. oversaw research. All authors contributed to writing and/or revising the manuscript.

ABSTRACT

ClpXP is an ATP-dependent protease in which the ClpX AAA+ motor binds, unfolds, and translocates specific protein substrates into the degradation chamber of ClpP. We present cryo-EM studies of the *E. coli* enzyme that show how asymmetric hexameric rings of ClpX bind symmetric heptameric rings of ClpP and interact with protein substrates. Subunits in the ClpX hexamer assume a spiral conformation and interact with two-residue segments of substrate in the axial channel, as observed for other AAA+ proteases and protein-remodeling machines. Strictly sequential models of ATP hydrolysis and a power stroke that moves two residues of the substrate per translocation step have been inferred from these structural features for other AAA+ unfoldases, but biochemical and single-molecule biophysical studies indicate that ClpXP operates by a probabilistic mechanism in which five to eight residues are translocated for each ATP hydrolyzed. We propose structure-based models that could account for the functional results.

INTRODUCTION

AAA+ motors harness the energy of ATP hydrolysis to carry out mechanical tasks in cells (Erzberger and Berger, 2006). In the ClpXP protease, for example, AAA+ ClpX ring hexamers bind target proteins, unfold them, and translocate the unfolded polypeptide through an axial channel and into the peptidase chamber of ClpP, which consists of two heptameric rings (Figure 4.1A) (Grimaud et al., 1998; Ortega et al., 2002; Sauer and Baker, 2011; Wang et al., 1997). In the absence of ClpX or another AAA+ partner, small peptides diffuse into the ClpP chamber through narrow axial pores, but larger peptides and native proteins are excluded and escape degradation (Grimaud et al., 1998; Lee et al., 2010b). The substrates of *Escherichia coli* ClpXP include aberrant ssrA-tagged proteins, produced by abortive translation, and normal cellular proteins synthesized with degradation tags that program rapid turnover (Baker and Sauer, 2012; Keiler, 2015).

ClpX subunits consist of a family specific N-terminal domain, which is dispensable for degradation of ssrA-tagged proteins, and large and small AAA+ domains, which contain sequence motifs that mediate ATP binding and hydrolysis, ClpP binding, and substrate recognition (Figure 4.1B) (Baker and Sauer, 2012). Ring hexamers of ClpX bind the ssrA tag within an axial channel (Martin et al., 2008a). Following degron binding, ATP-fueled power strokes pull on and eventually unfold attached native domains (Kenniston et al., 2003). Single-chain ClpX$^{\Delta N}$ pseudohexamers, containing six 'subunits' linked by genetically encoded tethers, support ClpP degradation of ssrA-tagged substrates at rates similar to wild-type ClpX (Martin et al., 2005). Eliminating ATP hydrolysis in four or five subunits of single-chain pseudohexamers slows but does not prevent ClpP-mediated degradation, suggesting that ATP hydrolysis in any one of multiple subunits in the ClpX ring is sufficient to power unfolding and translocation (Martin et al., 2005). In optical-

trapping experiments using single-chain ClpX and ClpP, the smallest observed translocation steps correspond to movement of five to eight amino acids of the substrate, and kinetic bursts of power strokes produce fast translocation steps two, three, or four-fold larger in terms of the number of residues translocated (Aubin-Tam et al., 2011; Cordova et al., 2014; Maillard et al., 2011; Olivares et al., 2017; Sen et al., 2013). Such bursts do not occur in repeating patterns, supporting probabilistic but coordinated ATP hydrolysis within the ClpX ring.

We describe here near atomic-resolution single-particle cryo-EM structures of single-chain ClpX pseudohexamers bound to ClpP and protein substrates. These structures show how asymmetric hexameric rings of ClpX dock with symmetric heptameric rings of ClpP and reveal how the pore-1, pore-2 and RKH loops of ClpX function in substrate binding. ClpX adopts a spiral conformation, with neighboring pore-1 loops interacting with every two residues of substrate in the axial channel, as observed in other AAA+ unfolding and remodeling machines (Cooney et al., 2019; Dong et al., 2019; Gates et al., 2017; Han et al., 2017; Majumder et al., 2019; de la Peña et al., 2018; Puchades et al., 2017; Ripstein et al., 2017; Rizo et al., 2019; Shin et al., 2019; Su et al., 2017; Sun et al., 2017; Twomey et al., 2019; White et al., 2018; Yu et al., 2018; Zehr et al., 2017). Based on these structural features, strictly sequential models of ATP hydrolysis and a power stroke that moves two residues of the substrate per translocation step have been proposed. As noted, however, ClpX does not need to operate in a strictly sequential manner and takes translocation steps substantially longer than two residues. Thus, an apparent incongruity exists between the structural and functional studies. We discuss this conflict and propose structure-based translocation models that reconcile how ClpX might use probabilistic ATP hydrolysis to take larger translocation steps of varying length.

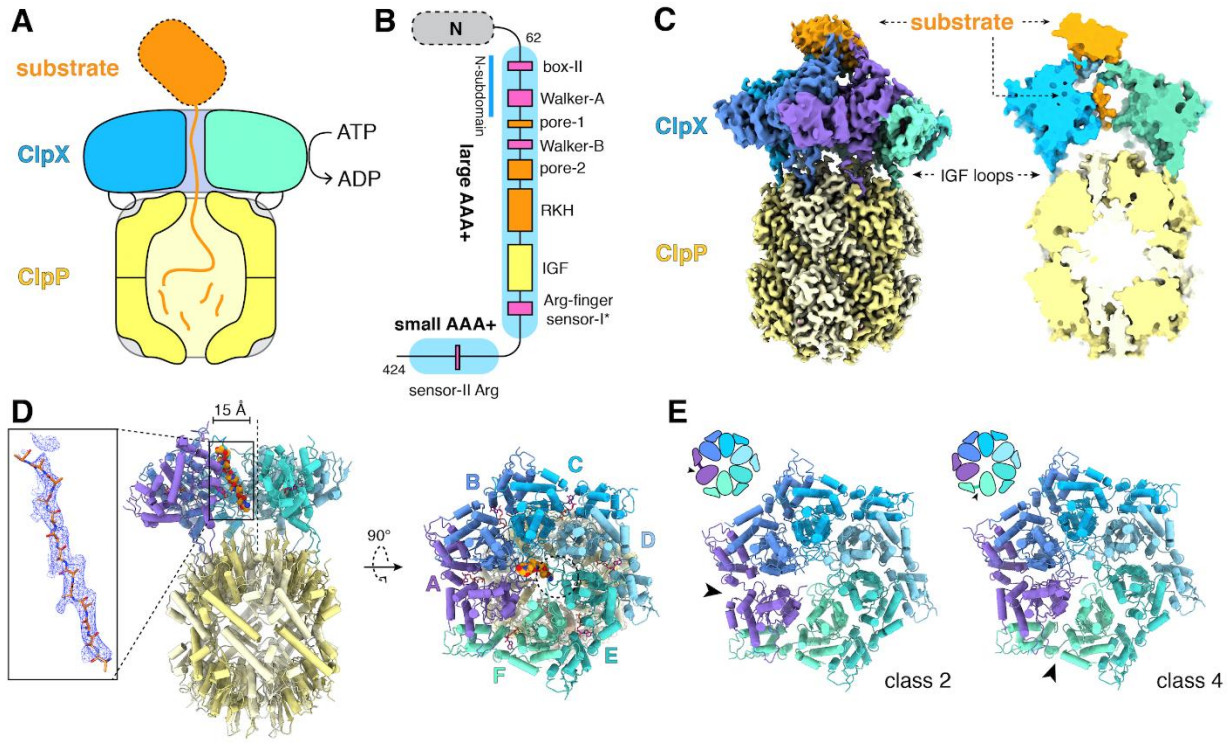


Figure 4.1 – ClpXP protease

(A) Schematic representation of ClpXP. **(B)** ClpX domain structure and positions of sequence motifs important for ATP binding and hydrolysis (magenta), substrate binding (orange), and ClpP binding (yellow). **(C)** Left. Composite cryo-EM density of $ClpX^{\Delta N}$/ClpP complex. ClpP is yellow; $ClpX^{\Delta N}$ (class 3) is blue, green, or purple; and substrate is orange. Right. Slice through surface representation of the model showing substrate in the axial channel and the ClpP degradation channel. **(D)** Left. Cartoon representation of the class-4 $ClpX^{\Delta N}$ hexamer (subunit F removed) and its docking with a heptameric ClpP ring. Substrate in the ClpX channel is shown in space-filling representation as a poly-alanine chain; the inset shows substrate as a ball-and-stick model with associated density. The dashed line shows the 7-fold symmetry axis of ClpP. Right. View rotated by 90°. The dashed circle shows the position of the ClpP pore. **(E)** Cartoon representation of class-2 and class-4 $ClpX^{\Delta N}$ hexamers with substrate removed for clarity. Arrows point to the different positions of the seam interface.

RESULTS

Cryo-EM Structures

For cryo-EM studies, we used epitope-tagged variants of *Escherichia coli* ClpP and a single-chain variant of *E. coli* ClpX$^{\Delta N}$ with an E185Q mutation to eliminate ATP hydrolysis without compromising nucleotide, ClpP, or substrate binding (Hersch et al., 2005; Martin et al., 2005). Single-chain ClpX$^{\Delta N}$ was used to ensure subunits of the pseudohexamer do not dissociate during sample preparation. This enzyme has been used previously for many biochemical and single-molecule studies (Amor et al., 2019; Aubin-Tam et al., 2011; Bell et al., 2018; Cordova et al., 2014; Glynn et al., 2012; Iosefson et al., 2015a, 2015b; Maillard et al., 2011; Martin et al., 2005, 2007, 2008a, 2008b; Olivares et al., 2017; Rodriguez-Aliaga et al., 2016; Sen et al., 2013; Stinson et al., 2013). We purified enzymes separately and incubated ClpX$^{\Delta N}$ (4 μM pseudohexamer), ClpP (2 μM 14-mer), and ATPγS (5 mM) for five min before vitrification. Experiments using ATP instead of ATPγS resulted in fewer ClpXP complexes. Imaging revealed complexes with ClpP bound to one or two ClpX$^{\Delta N}$ hexamers (Figures 4.2A-C). We analyzed the more abundant doubly capped complexes. As single-particle data processing with C_2 symmetry produced maps with conformational heterogeneity, we used signal-subtraction methods to allow independent classification and refinement of ClpX$^{\Delta N}$ or ClpP density (Figures 4.2D-F). We calculated a D_7 symmetric map for ClpP and parts of ClpX$^{\Delta N}$ making symmetric contacts, four different classes of symmetry-free ClpX$^{\Delta N}$ maps, and then extended each asymmetric ClpX$^{\Delta N}$ map to include one heptameric ClpP ring. Final structures have good stereochemistry with resolutions from 3.2 to 4.3 Å (Figure 4.3, Table 4.1). Substrates were observed in all ClpX$^{\Delta N}$ structures and probably represent bound endogenous peptides/proteins or partially denatured portions of ClpP or ClpX$^{\Delta N}$.

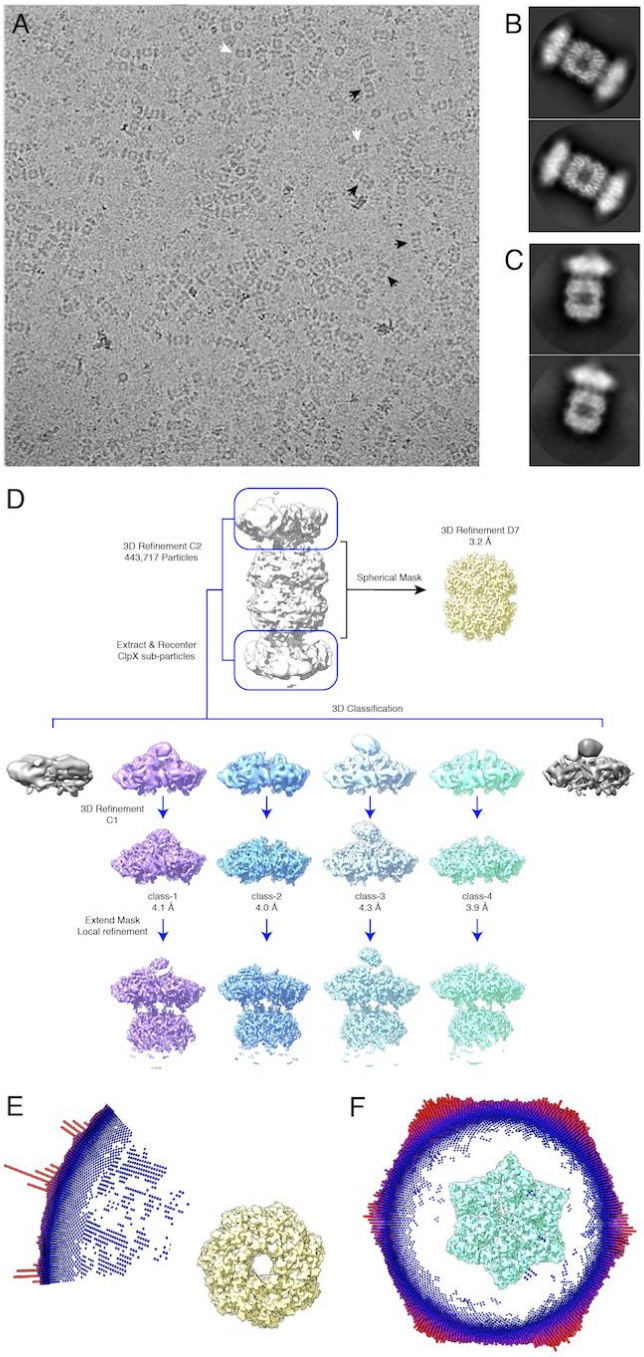


Figure 4.2 – Cryo-EM data and strategy
(A) Cryo-EM micrograph at 36000X magnification of doubly capped particles (black arrows) and singly capped particles (white arrows). **(B)** 2D-class averages of doubly capped complexes. **(C)** 2D-class averages of singly capped complexes. **(D)** Relion data processing scheme used to obtain 3D reconstructions. **(E)** Euler-angle distribution of particles used to reconstruct the D7 symmetric ClpP map. **(F)** Euler-angle distribution of particles used to reconstruct the non-symmetric class-4 ClpXΔN map.

Figure 4.1C shows density for a composite $ClpX^{\Delta N}$/ClpP/substrate complex. The six subunits of $ClpX^{\Delta N}$ formed a shallow spiral (labeled ABCDEF from top to bottom) in all structural classes, which differed largely in substrate density or nucleotide state, and the hexameric $ClpX^{\Delta N}$ and heptameric ClpP rings were slightly offset (Figure 4.1D). The structure of the $ClpX^{\Delta N}$ hexamers in the class-1, class-3, and class-4 EM structures were very similar to each other (pair-wise Cα RMSDs 1.2-1.9 Å). The hexamer in the class-2 structure was generally similar (pair-wise Cα RMSDs 2.6 Å) but differed in the position of the 'seam', a dilated inter-subunit interface that occurs as a result of ring closure. This seam was located between subunits A and B in the class-2 structure and between subunits F and A in the class-1, class-3, and class-4 structures, (Figure 4.1E). As discussed below, this difference appears to be related to the identity of the nucleotide bound in subunits A or F.

ClpX docking with ClpP

Our D_7-symmetric map included $ClpP_{14}$ and symmetric interface contacts with $ClpX^{\Delta N}$ (Figure 4.4A). In both heptameric ClpP rings, the N-terminal residues of each subunit formed a collar of β-hairpins creating a pore into the degradation chamber (Figures 4.4A-B, 4.5A). ClpP had essentially the same structure in the D_7 and symmetry-free ClpXP maps. ClpX IGF loops, named for an Ile^{268}-Gly^{269}-Phe^{270} sequence, are critical for ClpP binding (Amor et al., 2019; Joshi et al., 2004; Kim et al., 2001; Martin et al., 2007) and were responsible for most contacts in our structures (Figure 4.4C). These interactions included packing of the Ile^{268}, Phe^{270}, and Val^{274} side chains into pockets at interfaces between ClpP subunits (Figure 4.4D). I268L, F270L, and V274A variants have severe ClpP binding defects (Amor et al., 2019). Asymmetric ClpX-ClpP docking relies on conformational adjustments in the N- and C-terminal residues of individual IGF loops, allowing

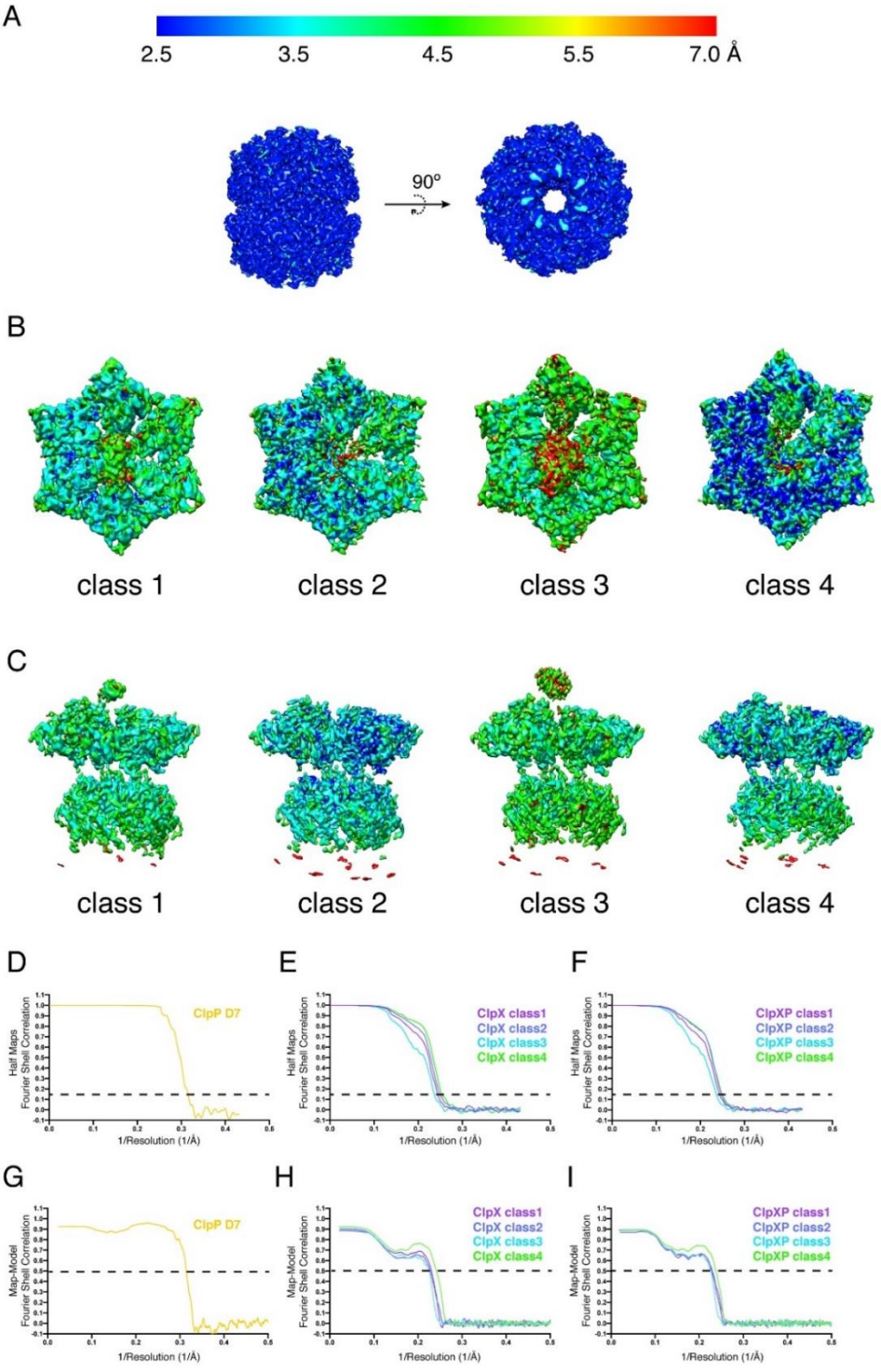


Figure 4.3 – Local map resolution and FSC plots

(A) Side views and top views of ClpP (D_7). **(B)** Top views of the four $ClpX^{\Delta N}$ classes. **(C)** Side views of the four ClpXP classes. ResMap was used to estimate local resolutions (Kucukelbir et al., 2014). **(D-F)** FSC plots of all masked final 3D reconstructions. **(G-I)** FSC plots calculated using masked 3D reconstructions and models.

the central portion of each IGF loop to contact the flat ClpP ring despite projecting from a spiral (Figure 4.4E). Changes in IGF-loop length decrease ClpP affinity and degradation activity (Amor et al., 2019), suggesting that these loops act as shock absorbers to maintain ClpP contacts during ClpXP machine function. In the symmetry-free maps, the unoccupied pocket in each ClpP heptamer was always located between the pockets bound by IGF loops from ClpX subunits E and F (Figure 4.5B).

In our structures, the side chain of Arg^{192} in ClpP appeared to hydrogen bond to the backbone of the IGF loop (Figure 4.4F**)**. Unlike wild-type ClpP, ^{R192K}ClpP neither bound $ClpX^{\Delta N}$ in pull-down assays (Figure 4.4G) nor degraded protein substrate in the presence of $ClpX^{\Delta N}$ (Figure 4.4H). Small-molecule acyldepsipeptides (ADEPs) bind in the same ClpP pockets as the ClpX IGF loops and kill bacteria by opening the ClpP pore to facilitate rogue degradation of unstructured proteins (Brötz-Oesterhelt et al., 2005; Kirstein et al., 2009; Lee et al., 2010a; Li et al., 2010). ADEPs stimulated R192K and wild-type ClpP decapeptide cleavage to similar extents (Figure 4.4I), establishing that ^{R192K}ClpP has normal peptidase activity. Thus, Arg^{192} is critical for ClpX binding but not for ADEP binding or ClpP pore opening. Modifying ADEPs to interact with Arg^{192} may increase affinity for ClpP and improve their efficacy as antibiotics.

Crystal structures show that ADEP binding to ClpP results both in opening of the axial pore and assembly of its N-terminal sequences into a collar of β-hairpins (Lee et al., 2010a; Li et al., 2010). Based on a recent cryo-EM structure of *Listeria monocytogenes* ClpXP, it was concluded that ClpX binding does not induce the same widening of the ClpP pore as does ADEP binding (Gatsogiannis et al., 2019). By contrast, our results support the opposite conclusion, as the diameter

Name	ClpP	ClpX-class 1	ClpX-class 2	ClpX-class 3	ClpX-class 4	ClpXP-class 1	ClpXP-class 2	ClpXP-class 3	ClpXP-class 4
PDB ID	6PPE	6PP8	6PP7	6PP6	6PP5	6POS	6POD	6PO3	6PO1
EMDB ID	20434	20422	20421	20420	20419	20418	20412	20408	20406
Data collection and processing									
Microscope	Talos Arctica								
Camera	K2 summit								
Magnification	36000X								
Voltage (kV)	200								
Total electron dose (e-/A2)	58								
Defocuse range (um)	-1.2 to -2.5								
Pixel size (Å)	0.58								
Micrographs collected	3657								
Final particles	443717	151652	241899	115751	227257	151652	241899	115751	227257
Symmetry	D_7	C_1	C_1	C_1	C_1	C_1	C_1	C_1	C_1
Resolution (FSC 0.143)	3.19	4.12	4.05	4.28	3.98	4.12	4.05	4.28	4.05
Model composition									
Non-hydrogen atoms	21840	15498	15593	15599	15530	26656	26626	26597	26336
Protein residues	2814	1994	2012	2006	2000	3432	3431	3425	3396
Ligands	0	6	6	6	6	6	6	6	6
Refinement									
Map-model CC	0.90	0.76	0.75	0.76	0.8	0.76	0.73	0.72	0.77
Map sharpening B factors ($Å^2$)	-180	-186	-235	-216	-170	-160	-176	-184	-184
R.m.s deviations - bond length (Å)	0.004	0.002	0.003	0.003	0.003	0.002	0.002	0.002	0.002
R.m.s deviations - bond angles (degrees)	0.648	0.637	0.667	0.664	1.164	0.650	0.666	0.658	0.660
Validation									
MolProbity score	0.60	0.94	0.94	0.94	0.91	0.91	0.83	0.91	0.90
Clashscore	0.27	1.83	1.78	1.78	1.63	1.60	1.16	1.63	1.53

C-beta deviations	0	0	0	0	0	0	0	0	0
Rotamers outlier (%)	0	0	0.18	0	0	0	0.04	0	0
Ramachandran favored (%)	98.51	99.95	99.85	99.90	99.85	99.91	99.76	99.73	99.97
Ramachandran allowed (%)	1.49	0.05	0.15	0.10	0.15	0.09	0.24	0.27	0.03
Ramachandran disallowed (%)	0	0	0	0	0	0	0	0	0

Table 4.1 – Cryo-EM data collection, processing, model building and validation statistics

and overall structure of the ClpP pore in our *E. coli* ClpXP structures were extremely similar to the crystal structure of ADEP-activated *E. coli* ClpP (Li et al., 2010), with an overall RMSDs of 0.8 Å for all Cα's in a single ClpP ring and 0.6 Å for all Cα's of the seven N-terminal β-hairpins and adjacent a-helices that define the pore diameter.

Substrate binding

All $ClpX^{\Delta N}$ maps contained substrate density within the axial channel, which we generally modeled as an extended poly-alanine chain (Figure 4.1D), although additional side-chain density at residue 3 of the substrate was modeled as arginine in the class-1 structure and histidine in the class-3 structure (Figure 4.6A). Substrate was built with the N terminus facing ClpP in the class-2 and class-4 structures and in the opposite orientation in the class-1 and class-4 structures. ClpXP can translocate substrates in either the N-to-C or C-to-N direction (Olivares et al., 2017). In all structures, the top of the ClpX channel was most constricted by tight packing between ClpX side chains and substrate (Figures 4.7A-B), providing a structural basis for experiments showing that interactions with substrate near the top of the ClpX channel are most important for unfolding grip (Bell et al., 2019). The class-1 and class-3 structures also had density corresponding to a roughly globular native domain above the channel (Figures 4.1C, 4.6B).

In our structures, the pore-1, pore-2, and RKH loops of ClpX contacted substrate in accord with genetic and biochemical studies (Farrell et al., 2007; Iosefson et al., 2015a, 2015b; Martin et al., 2008a, 2008b; Rodriguez-Aliaga et al., 2016; Siddiqui et al., 2004). For example, the pore-1 loops of subunits B, C, D, and E interacted with substrate in the channels of all classes, whereas these loops from subunits A and F did so only in some classes (Figures 4.6C-D, 4.8A-C). For each

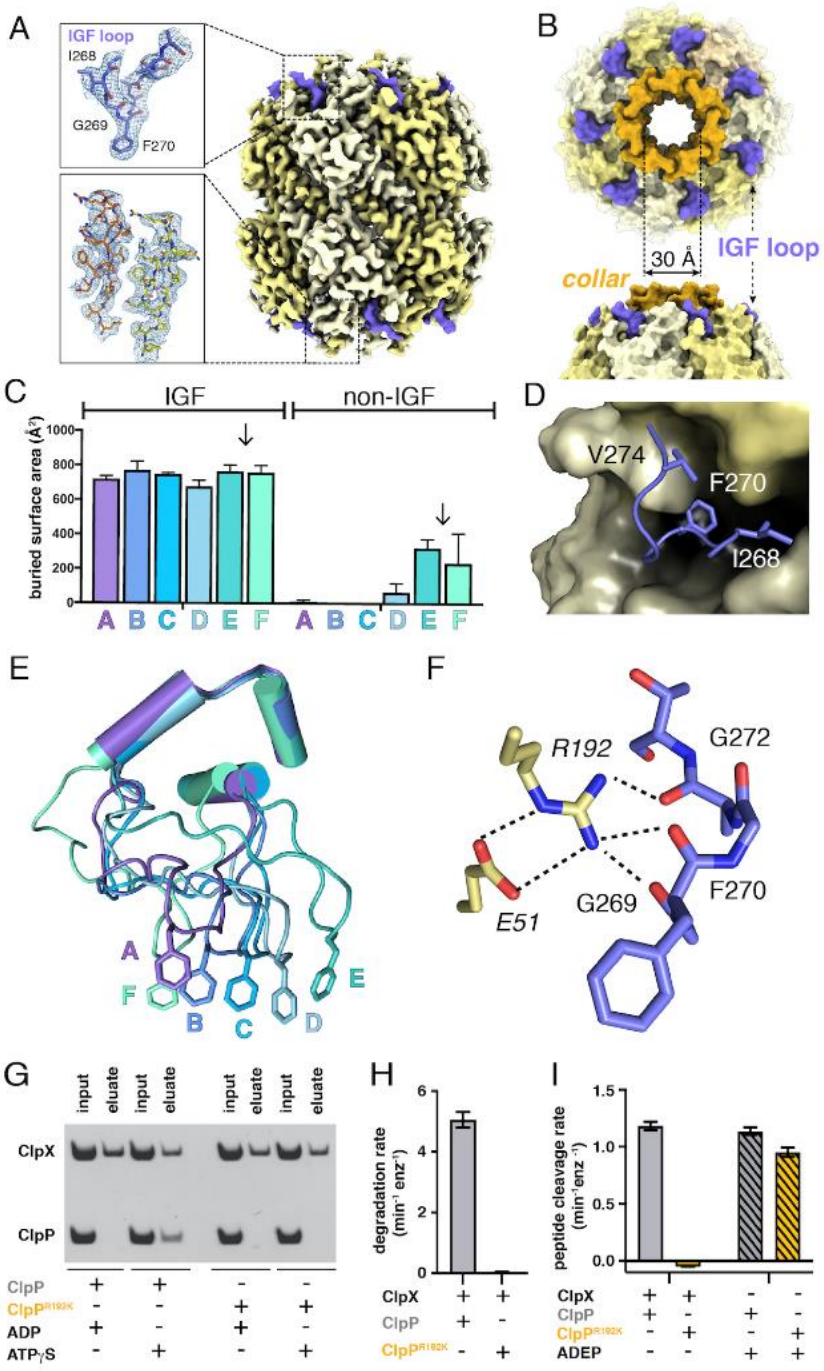


Figure 4.4 – ClpP and symmetric IGF loop contacts

(A) Density for $ClpP_{14}$ in the D_7 map is shown in yellow. Density for part of the IGF loops of ClpX is shown in blue. The upper inset shows the density as mesh and the fitted model for residues 267-275 of one IGF loop; the lower inset shows density and the fitted model for resides 2-18 of two

β-hairpins in the ClpP collar. **(B)** Top views of ClpP showing the axial collar (darker yellow) and pore. **(C)** IGF loops of ClpX make the majority of contacts with ClpP as assessed by buried surface area. Each bar represents a different $ClpX^{\Delta N}$ subunit in the spiral and values are means ± 1 SD for the four classes. The arrows between subunits E and F mark the position of the unoccupied ClpP cleft. **(D)** The side chains of IGF residues Ile^{268} (I268) Phe^{270} (F270), and Val^{274} (V274) pack into hydrophobic ClpP clefts. **(E)** After alignment of the large AAA+ cores of the six subunits in the class-4 $ClpX^{\Delta N}$ hexamer, the IGF loops adopt a variety of conformations, helping mediate ClpP binding despite the symmetry mismatch. The side chains of Phe^{270} are shown in stick representation with the IGF loop (residues 263-283) and three flanking helices (residues 254-262 and 284-298) shown in cartoon representation. **(F)** The side chain of Arg^{192} (R192) in ClpP makes hydrogen bonds with carbonyl oxygens in the IGF loop of ClpX and also forms a salt bridge with Glu^{51} (E51) in a neighboring ClpP subunit. **(G)** ClpP but not ^{R192K}ClpP binds $ClpX^{\Delta N}$ in pull-down assays performed in the presence of ATPγS. As a negative control, neither ClpP variant binds $ClpX^{\Delta N}$ in the presence of ADP (Joshi et al., 2004). **(H)** $ClpX^{\Delta N}$ supports degradation of ^{cp7}GFP-ssrA by ClpP but not by ^{R192K}ClpP. **(I)** $ClpX^{\Delta N}$ supports degradation of a decapeptide by ClpP but not by ^{R192K}ClpP. ADEP-2B activates decapeptide cleavage by both ClpP and ^{R192K}ClpP. Error bars represent means (n=3) ± 1 SD.

engaged pore-1 loop, the side chains of Tyr^{153} and Val^{154} packed between α-carbons spaced two-residues apart on opposite sides of the extended substrate (Figures 4.6C-D, 4.8A-C). Depending on the class, from three to five pore-2 loops also contacted substrate, with Arg^{200} making most of these interactions (Figures 4.6C-D, 4.8A-B, 4.8D).

The RKH loop, which has not been visualized in previous structures and is unique in ClpX-family enzymes (Baker and Sauer, 2012), consisted of an antiparallel β-ribbon stem and a short helix that includes part of the conserved Arg^{228}-Lys^{229}-His^{230} motif (Figure 4.8E). When contacting substrate above the pore, the RKH loops mimicked adjustable structural jacks supporting a house during foundation repair (Figure 4.8E). The RKH loops alter substrate specificity (Farrell et al., 2007; Martin et al., 2008a), but whether they are critical determinants of substrate binding is unknown.

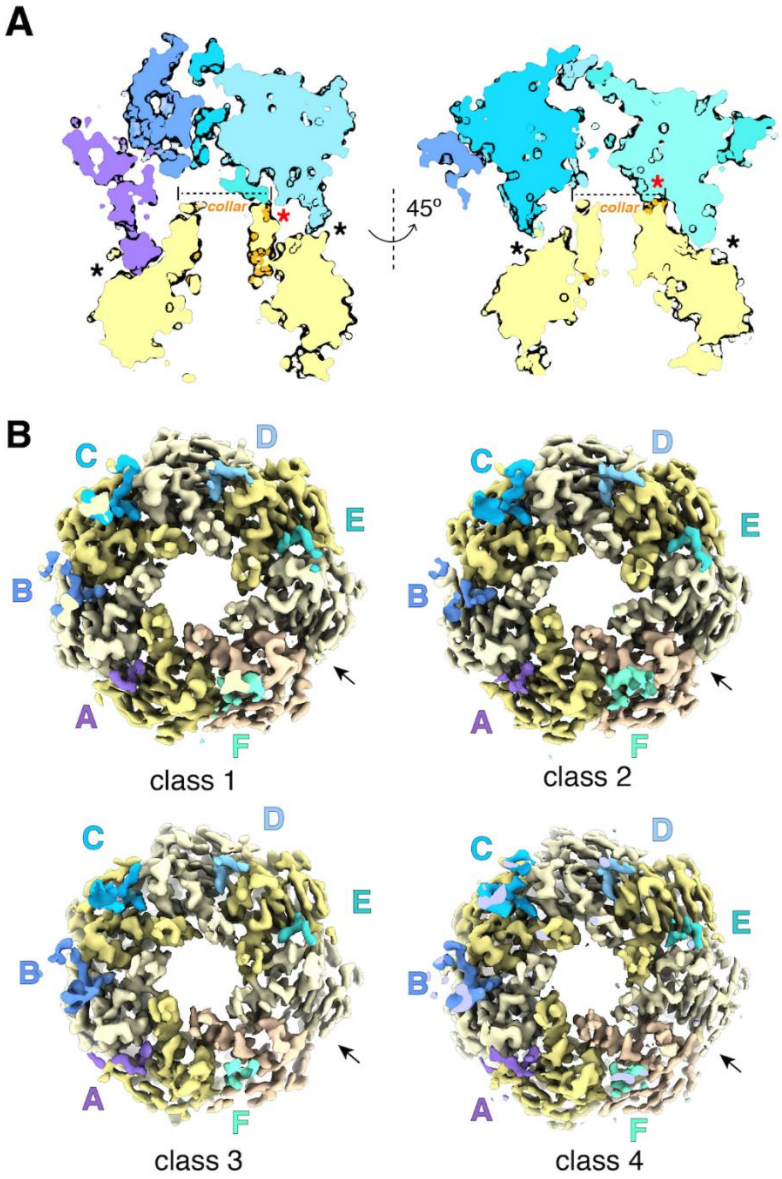


Figure 4.5 – Structural features of ClpX$^{\Delta N}$/ClpP interface
(A) Cutaway showing axial collar of ClpP fitting into a bowl-shaped cavity at the bottom of the ClpX$^{\Delta N}$ hexamer. Black asterisks show contacts between the ClpX IGF loops and ClpP clefts. Red asterisks show contacts between ClpX and the ClpP collar. **(B)** Top views of ClpXP classes with substrate and most ClpX$^{\Delta N}$ except IGF loops (blue or purple) removed. Arrows show the position of the empty ClpP cleft without an IGF loop.

To test this possibility, we constructed and assayed RKH-loop mutants. The most severe mutant (RKH→AAA) did not support ClpP degradation of an ssrA-tagged substrate and hydrolyzed ATP ~5-fold faster than the parent (Figures 4.8F-G). Changing RKH to AKH, KHR, or KRH slowed degradation (Figure 4.8F). A variant $[RKH_2\text{-}AAA]_2$ pseudohexamer with four wild-type RKH loops and two AAA mutations had a ~3-fold higher K_M for substrate degradation than the parent and hydrolyzed ATP slightly faster (Figures 4.8G-H). Thus, the RKH loops play important roles in substrate recognition and in regulating rates of ATP hydrolysis.

Nucleotide binding and motor conformations

We observed density for five ATPγS molecules in each $ClpX^{\Delta N}$ hexamer (Figure 4.9A, Table 4.2). In the sixth site, bound nucleotide fit best as ADP in subunit A (class 2) or subunit F (classes 1, 3, and 4), although with slightly less convincing density than for ATPγS in other subunits (Fig. 4.9B, Table 4.2), raising the possibility of averaging of multiple nucleotide-binding states. These 'ADP-bound' subunits had poorly structured channel loops, made fewer contacts with substrate and with neighboring subunits, and were offset from the axis of the hexamer by ~2.5 Å (Figures 4.9C-E, Table 4.3).

Side-chain contacts with nucleotide included ClpX residues Val^{78}-Ile^{79} (box-II), Lys^{125}-Thr^{126}-Leu^{127} (Walker A), Asp^{184}-Gln^{185} (Walker B with E185Q mutation), Arg^{307} (arginine finger), and Arg^{370} (sensor-II) (Figures 4.9A-B, 4.10A). V78A/I79A, E185Q, and R370K mutations eliminate ATP hydrolysis or weaken nucleotide binding (Hersch et al., 2005; Joshi et al., 2004; Martin et al., 2005; Stinson et al., 2013). We found that K125M, T126A, L127A, D184A, and R307A mutations also severely inhibited ATP hydrolysis (Figure 4.10B). ClpX does not have a traditional sensor-I

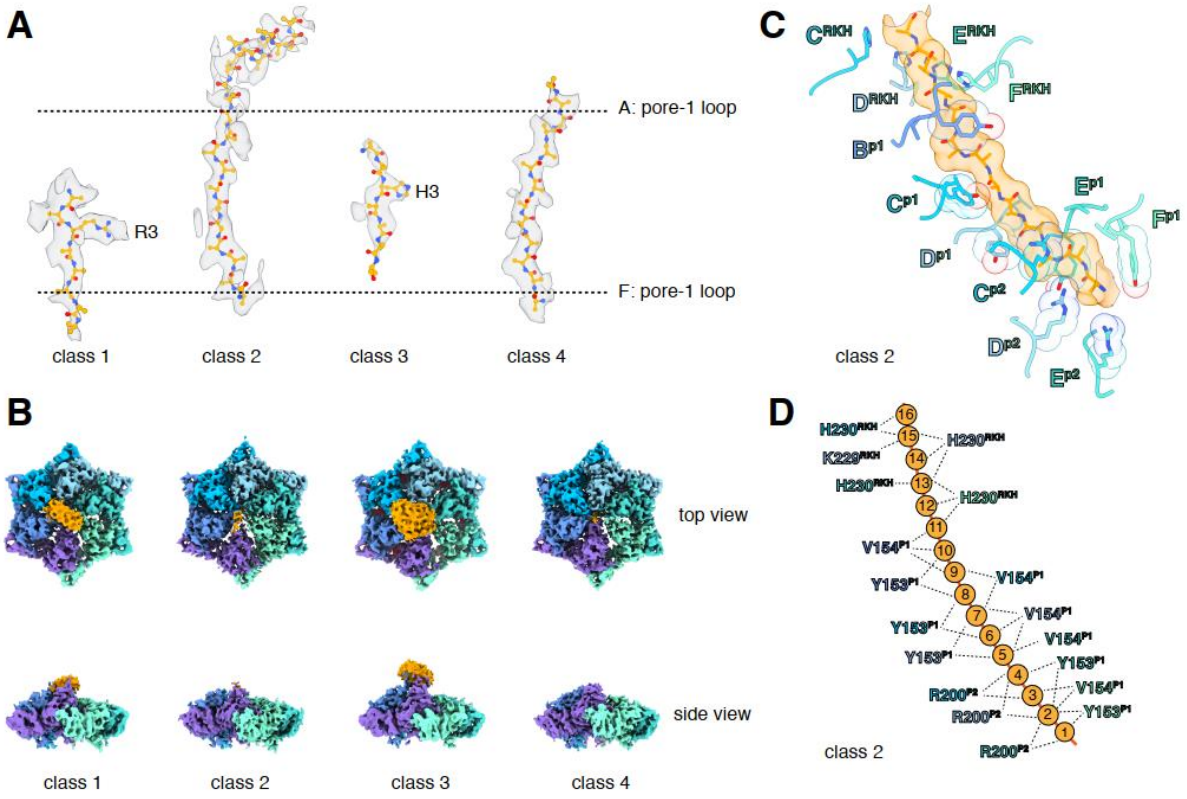


Figure 4.6 – Substrate interactions with ClpX

(A) Substrate density within the ClpX channel is rendered as a transparent surface (grey) with atoms shown in ball and stick representation. Dashed lines shown approximate positions of the pore-1 loop from the ClpX subunit A (top) and subunit F (bottom). **(B)** Substrate density above the ClpX hexamer. Substrate density is colored in orange and ClpX subunits are colored in different shades from purple to aquamarine. **(C)** Interactions between substrate (orange, stick and surface representation) and ClpX pore loops (blue and purple, stick and transparent space-filling representation) in the class-2 structure. Capital letters indicate ClpX subunits; superscripts indicate RKH loops, pore-1 (p1) loops, or pore-2 (p2) loops. **(D)** Scheme of interactions shown in panel C. Dashed lines represent distances of 6.5 Å or less between the Cb atoms of substrate alanines and the Cβ atoms of Y153/V154 (p1), or the Cε atom of R200 (p2), or the Cγ atom of H230 (RKH), or the Nζ atom of R228 (RKH).

residue, which in many AAA+ enzymes positions a water molecule for hydrolysis of ATP bound to the same subunit (Erzberger and Berger, 2006). For ATPγS bound in subunits A-E of our structures, the side chain of Glu^{303} was close to the γ-thiophosphate, at a sensor-I-like position (Figure 4.9A). We found that an E303A mutation severely inhibited ATP hydrolysis, whereas E303Q was partially defective (Figures 4.10B-C). Based on these results, we refer to Glu^{303} as a sensor-I* element. As Glu^{303} is in the same helix as Arg^{307}, the arginine-finger residue, these residues could coordinate structurally to activate ATP hydrolysis in a neighboring subunit.

A short hinge connects the large and small AAA+ domain of each $ClpX^{\Delta N}$ subunit. In all of our structures, the conformations of the small domains were very similar to each other, as were those of the large domains after removing an N-terminal subdomain (residues 65-114) and the pore-1, pore-2, RKH, and IGF loops (Figure 4.11). The relative orientations of the large and small AAA+ domains were similar across all four structural classes for subunits at equivalent spiral positions, but conformational changes in the hinge resulted in different orientations or the large and small domains at many positions in the spiral for each hexamer (Figure 4.9F). Changes in hinge length or deletion of one hinge largely eliminate ClpX function (Bell et al., 2018; Glynn et al., 2012). Thus, the conformational changes associated with a power stroke likely arise from changes in hinge conformations, whereas movements of different large-domain loops or the N-subdomain mediate ring closure and asymmetric contacts with ClpP and substrate.

ATP hydrolysis requires proper positioning of the box-II, Walker-A, and Walker-B elements in the large AAA+ domain of one ClpX subunit, the sensor-II arginine from the small AAA+ domain of the same subunit, and the arginine-finger/sensor-I* element from the large domain of the

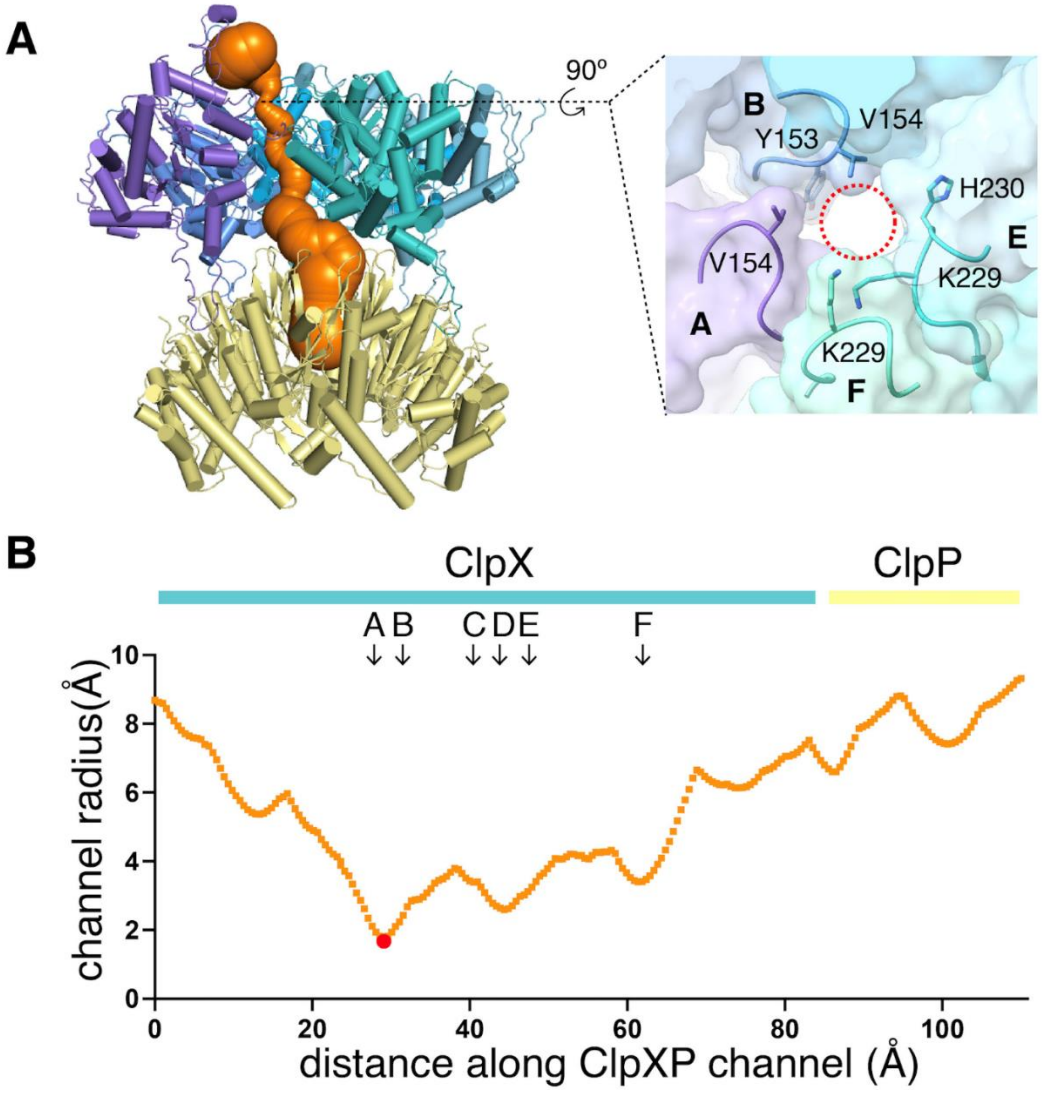


Figure 4.7 – Dimensions of axial channel

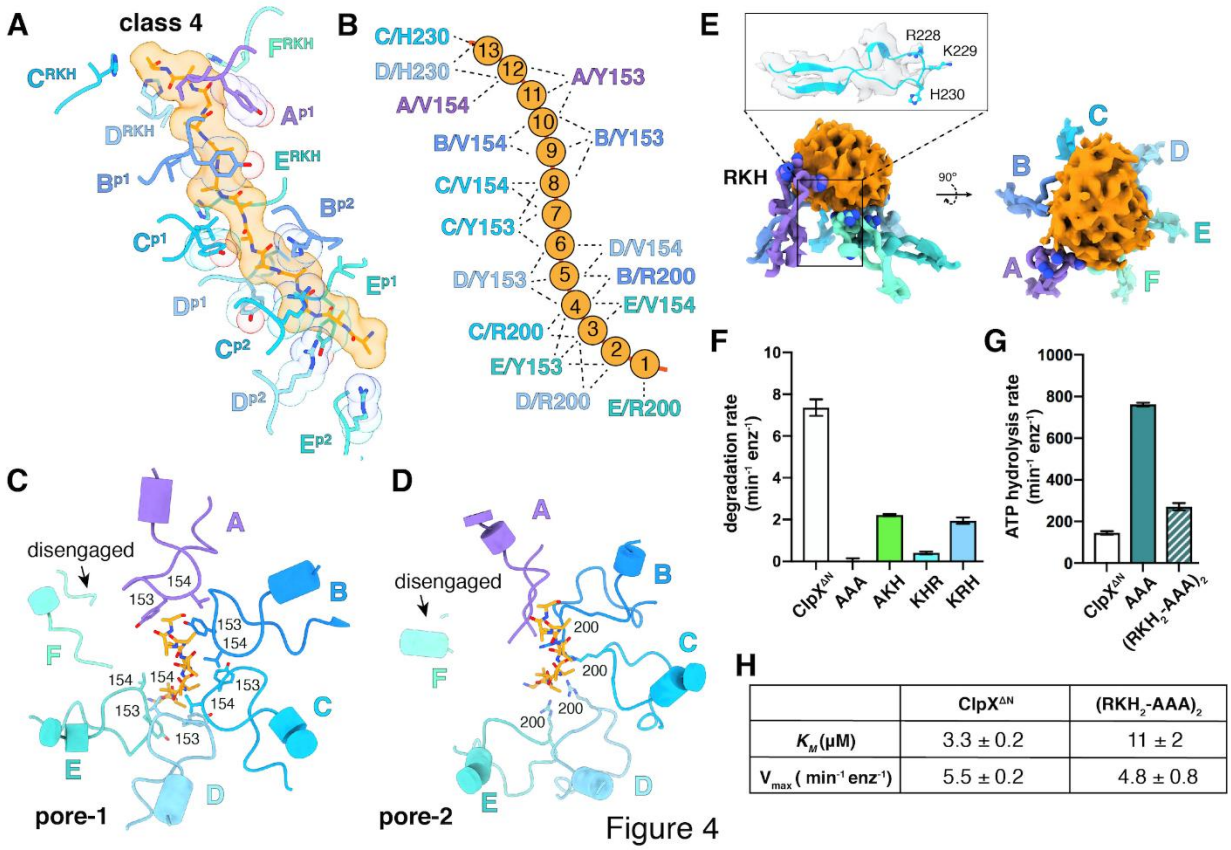
(A) The channel though ClpX and into ClpP (class-1) is most constricted near the top of the ClpX$^{\Delta N}$ hexamer. This panel was made using *Caver* (Pavelka et al., 2016). The inset shows that the constriction involves the pore-1 loops of subunits A and B and the RKH loops of subunits E and F. **(B)** Plot of channel radius as a function of channel length. Arrows mark positions of pore-1 loops. Red dot marks the bottleneck of this channel.

clockwise subunit (viewed from the top of the ring). These structural features depend on how the small AAA+ domain of each subunit packs against the large AAA+ domain of its clockwise neighbor, which was similar for units A/B, B/C, C/D, D/E, and E/F, suggesting that the associated ATP-binding sites are hydrolytically active. By contrast, the structure of the FA interface was different as a consequence of changes in rotation of the N-subdomain in subunit A relative to the rest of the large AAA+ domain and changes in a loop that contains the sensor-II arginine in subunit F (Figure 4.9G). These changes resulted in disengagement of the Arg-finger and sensor-I* side chains in subunit A from the nucleotide bound to subunit F (Figure 4.9B, Table 4.4). Thus, in each of our structures, the nucleotide-binding site in subunit F appears to be catalytically inactive.

DISCUSSION

ClpX interactions with ClpP and substrates

Our cryo-EM structures provide snapshots of ClpX binding to ClpP, protein substrates, and nucleotides. In each of our structures, the six subunits of the ClpX ring hexamer are arranged in a shallow spiral. Slightly altered orientations of the large and small AAA+ domains in each ClpX subunit allow the hexameric ring to remain topologically closed with each large domain contacting the small domain of one neighbor. These structural results are consistent with biochemical experiments that show that ClpX is fully functional when all of its subunit-subunit interfaces are covalently crosslinked (Glynn et al., 2012) and support a model in which the architectural changes in the spiral that drive a power stroke result from changes in the conformations of the hinges connecting the large and small AAA+ domains of each subunit. Our structures show that relatively flexible interactions between IGF loops of ClpX and binding pockets on ClpP heptamers allow docking of these symmetry-mismatched partners. Although IGF-ClpP contacts are highly dynamic



	ClpX$^{\Delta N}$	(RKH$_2$-AAA)$_2$
K_M (μM)	3.3 ± 0.2	11 ± 2
V_{max} (min^{-1} enz^{-1})	5.5 ± 0.2	4.8 ± 0.8

Figure 4

Figure 4.8 – Substrate contacts

(A) Interactions between substrate (orange, stick and surface representation) and ClpX$^{\Delta N}$ loops (blue and purple, stick and transparent space-filling representation) in the class-4 structure. Capital letters indicate subunits; superscripts indicate RKH loops, pore-1 (p1) loops, or pore-2 (p2) loops. **(B)** Scheme of interactions shown in panel (A). Dashed lines represent distances of 6.5 Å or less between the Cβ atoms of substrate alanines and the Cβ atoms of Y153/V154 (p1), or the Cε atom of R200 (p2), or the Cγ atom of H230 (RKH). **(C)** Top view of interactions of pore-1 loops with substrate (class 4). **(D)** Top view of interactions of pore-2 loops with substrate (class 4). **(E)** Interaction of RKH loops in the class-3 structure with the globular portion of the substrate above the channel. Inset – representative RKH-loop density (class 4, subunit C) and positions of R228, K229, and H230. **(F)** Mutation of RKH motifs in each subunit of a ClpX$^{\Delta N}$ hexamer inhibits degradation of Arc-st11-ssrA. **(G)** Effects of RKH mutations on ATP hydrolysis. **(H)** Mutating two RKH motifs in a single chain pseudohexamer to AAA increases K_M for steady-state degradation of ^{CP7}GFP-ssrA.

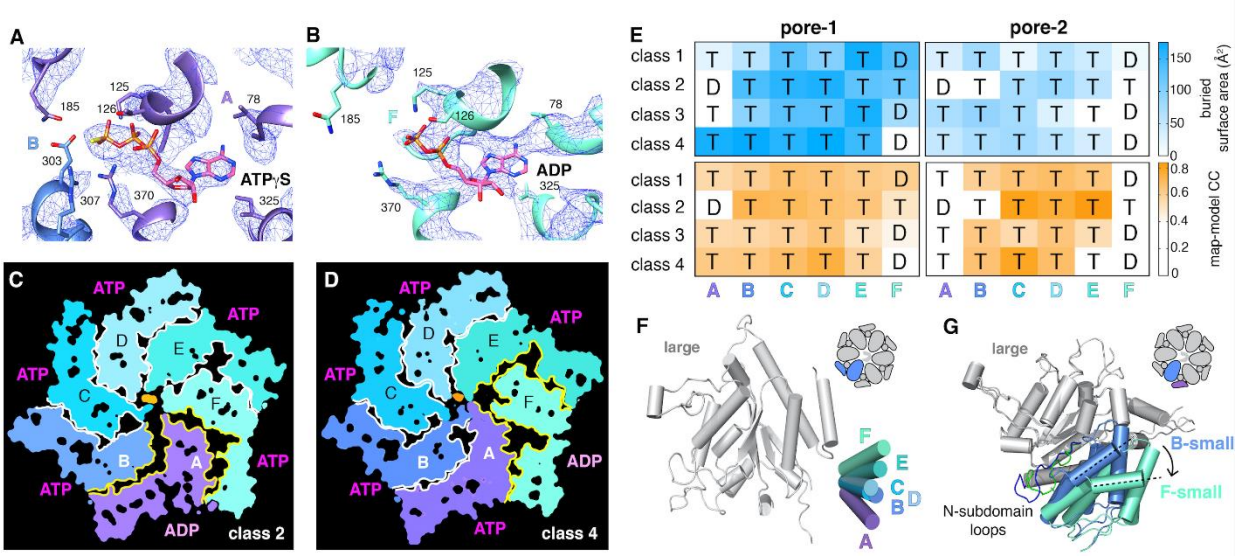


Figure 4.9 – Nucleotide binding and subunit interactions in ClpX$^{\Delta N}$ hexamers

(A) Density for ATPγS in subunit A of the class-4 hexamer. The positions of ATP-binding/hydrolysis motifs are shown in cartoon and stick representation. **(B)** Weaker nucleotide density built as ADP in subunit F of the class-4 hexamer. **(C, D)** Slices through the density maps of class-2 and class-4 hexamers show that the ADP-bound subunit (A or F, respectively) makes fewer neighbor contacts than ATP-bound subunits (yellow). **(E)** Contacts between substrate and pore-1 loops (left) or pore-2 loops (right) were characterized by buried surface area (top) or model-to-map correlation (bottom). T represents an ATP-bound subunit; D represents an ADP-bound subunit. **(F)** Superimposition of the large AAA+ domains of six subunits from a hexamer (class 4) showing variation in the angle between small and large domains of the same subunit. One of the large AAA+ domains is light grey, and the small AAA+ domains are in different shades of purple, blue and green. **(G)** Superimposition of two large AAA+ domains from a class-4 hexamer, illustrating different types of packing against the neighboring small AAA+ domain. The large domains of subunits A and C are light grey; the small domains of subunit F and B are cornflower blue and mint green. N sub-domain loops are colored green (subunit F) and blue (subunit B). The dashed lines and arrow show an ~30° rotation at the interface. The F/A interface is the seam.

in solution (Amor et al., 2016), they were well defined in our structures and revealed essential interactions.

Protein substrates were observed in the axial channel of ClpX in all structures and above the channel in two structures. The ClpX pore-1 and pore-2 loops were responsible for most substrate contacts in the channel, with a periodicity of two substrate residues per ClpX subunit. The RKH loops of ClpX, which we found play critical roles in substrate recognition and control of ATP-hydrolysis rates, also contacted substrate near the top and above the channel. Substrate contacts near the top of the channel are most important in determining substrate grip during unfolding (Bell et al., 2019), and substrate-ClpX contacts were tightest in this region in our structures. Finally, we observed ATPγS bound to five of the six ClpX subunits with ADP apparently bound to the sixth subunit, and determined how ClpX functions without a traditional sensor-I residue.

Our cryo-EM structures contrast markedly with crystal structures of $ClpX^{\Delta N}$ pseudohexamers, which did not form a spiral, bound only four nucleotides, and had conformations that now seem incompatible with ClpP and/or substrate binding (Glynn et al., 2009; Stinson et al., 2013).

Proteolytic motors from different AAA+ clades have similar structures

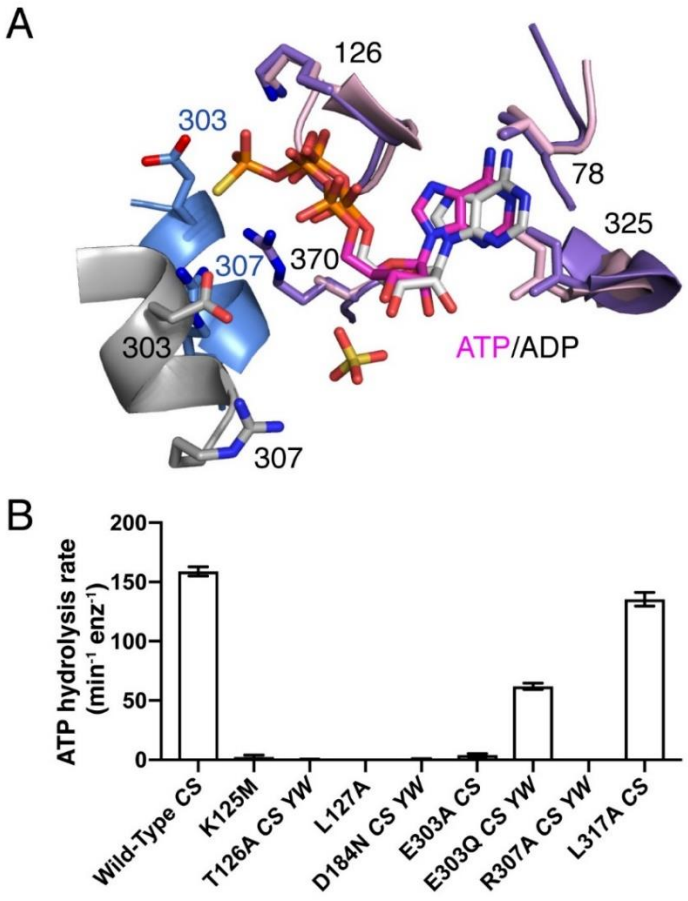
AAA+ protease motors belong to either the classic or HCLR clades (Erzberger and Berger, 2006). For HCLR-clade members ClpX and Lon, the spiral hexamer architectures and pore-1-loop interactions with substrate in the axial channel are similar to those for the classic-clade YME1/FtsH and the proteasomal Rpt_{1-6}/PAN motors (Dong et al., 2019; Majumder et al., 2019; de la Peña et al., 2018; Puchades et al., 2017; Shin et al., 2019). Thus, from a structural perspective, it is reasonable to suggest that motors from different clades may operate by a common fundamental mechanism.

		A	B	C	D	E	F
class 1	protein	0.78	0.79	0.81	0.81	0.81	0.80
	nucleotide	0.77	0.80	0.82	0.81	0.84	0.74
class 2	protein	0.81	0.83	0.82	0.82	0.80	0.80
	nucleotide	**0.59**	0.87	0.80	0.80	0.81	**0.67**
class 3	protein	0.82	0.83	0.80	0.80	0.80	0.78
	nucleotide	0.79	0.84	0.83	0.80	0.77	0.76
class 4	protein	0.81	0.76	0.74	0.73	0.76	0.80
	nucleotide	0.83	0.77	0.78	0.74	0.78	**0.65**

Table 4.2 – Model to map correlation coefficient (CC) of amino acids and nucleotides
Calculated using phenix_realspace.correlation (Adams et al., 2010). Nucleotide sites with suspected partial occupancy are marked in bold.

	A-B	B-C	C-D	D-E	E-F	F-A
class 1	1.76	2.02	2.10	2.36	2.02	1.51
class 2	1.29	1.97	2.21	2.23	1.99	1.38
class 3	1.89	2.18	2.06	2.30	2.03	1.35
class 4	1.99	2.22	2.25	2.24	1.70	1.26

Table 4.3 – Buried surface area (in units of 1000 Å^2) between ClpX$^{\Delta N}$ subunits
The interface in each hexamer with the smallest amount of buried surface is highlighted in gray and corresponds to the ADP-bound subunit.



C

Construct	Vmax ($min^{-1}enz^{-1}$)	K_M (μM)	n
Wild-Type *CS*	120 ± 10	57 ± 3	1.9 ± 0.1
E303Q *CS YW*	86 ± 3	43 ± 3	2.3 ± 0.1
L317A *CS*	130 ± 10	270 ± 10	1.6 ± 0.1

Figure 4.10 – Nucleotide-binding pocket and mutations

(A) Alignment of residues contacting bound nucleotide in the class-4 cryo-EM structure (A/B subunit interface; purple/blue) and crystal structure of $ClpX^{\Delta N}$ (pink/grey; PDB 3HWS) (Glynn et al., 2009). **(B)** ATP-hydrolysis activity for $ClpX^{\Delta N}$ variants with mutations in nucleotide-binding pocket. Values are averages of three independent replicates ± SD. **(C)** Fitted Michealis-Menten-Hill parameters for ATP hydrolysis for variants with significant activity. Values are averages of three independent replicates ± SD. In panels B and C, *CS* (C169S) and *YW* (Y77W) indicate additional mutations in variants that do not, by themselves, alter ATP hydrolysis activity.

A sequential translocation model has been proposed for Yme1 and the Rpt_{1-6} ring of the 26S proteasome, based on placing distinct cryo-EM structures with different nucleotide and substrate-engagement states in a defined kinetic pathway, (Dong et al., 2019; de la Peña et al., 2018; Puchades et al., 2017). In this model, ATP hydrolysis in the fifth spiral subunit (subunit E in ClpX) drives a power stroke that moves each subunit from its previous location to a position offset by one subunit in the clockwise direction, generating a two-residue translocation step (Figure 4.12A). We refer to this translocation model as SC/2R (Sequential Clockwise/2-Residue Step). Based on similar hexamer architectures and/or substrate interactions, a SC/2R translocation model has also been proposed for Lon, for PAN/20S, and for the AAA+ protein-remodeling machines ClpB/Hsp104, CDC48/p97, NSF, and Vps4 (Cooney et al., 2019; Gates et al., 2017; Han et al., 2017; Majumder et al., 2019; Ripstein et al., 2017; Rizo et al., 2019; Shin et al., 2019; Su et al., 2017; Sun et al., 2017; Twomey et al., 2019; White et al., 2018; Yu et al., 2018; Zehr et al., 2017).

Discrepancies between ClpX function and SC/2R-model predictions

The ClpX structures reported here resemble those of Yme1 and Rpt_{1-6} in the spiral architecture of the hexamer, the positions of subunits that contain a nucleoside triphosphate, the interaction of successive pore-1 loops with two-residue segments of the substrate, and the patterns of substrate engaged and disengaged pore-1 and pore-2 loops in the ring (Dong et al., 2019; de la Peña et al., 2018; Puchades et al., 2017). Thus, ClpXP might also be expected to operate by an SC/2R mechanism, but multiple experimental observations suggest otherwise.

The first issue involves the length of the smallest, fundamental translocation steps. Structures show that axial-channel binding in the AAA+ ring enforces an extended polypeptide conformation in

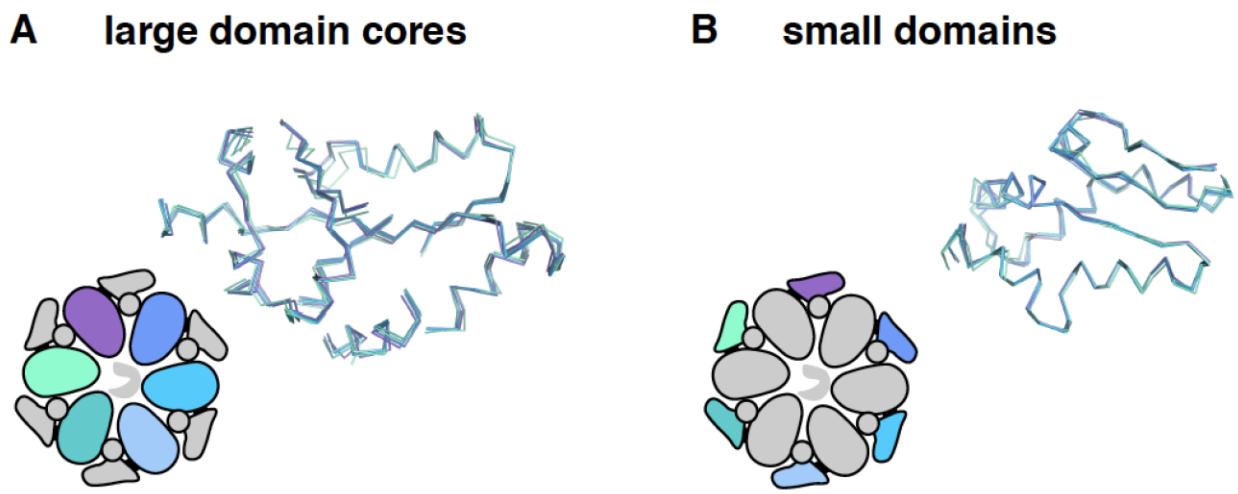


Figure 4.11 – Invariant and flexible portions of ClpX$^{\Delta N}$
(A) Superimposition of the cores of the large AAA+ domains from each subunit of the class-4 ClpX$^{\Delta N}$ structure. **(B)** Superimposition of the small AAA+ domains from the class-4 ClpX$^{\Delta N}$ hexamer.

class 1	R307 (Arg finger)	R370 (sensor II)	E185 (Walker B)
A	6.2	4.2	7.1
B	4.5	4.6	7.0
C	4.7	4.7	6.2
D	4.8	4.4	6.2
E	5.7	4.6	6.1
F	15.1	4.7	8.8
class 2	**R307**	**R370**	**E185**
A	8.0	4.0	7.9
B	4.9	4.0	6.8
C	5.2	4.2	6.0
D	4.4	4.3	6.6
E	5.9	4.5	6.6
F	16.2	4.6	6.3
class 3	**R307**	**R370**	**E185**
A	5.6	4.2	6.7
B	4.6	4.3	6.7
C	5.4	4.2	6.5
D	4.7	4.6	5.9
E	5.6	6.3	6.4
F	15.9	5.8	5.8
class 4	**R307**	**R370**	**E185**
A	6.2	4.0	7.0
B	4.6	4.3	6.6
C	4.7	4.6	6.9
D	4.6	4.3	6.4
E	8.8	5.1	5.9
F	13.8	4.6	7.7

Table 4.4 – Distances between nucleotides and key residues
ADP-occupied subunits are highlighted in gray. Measurements in Å are between the γ thiophosphate of ATPγS or β phosphate of ADP and the Cε of arginine or Cγ of glutamine.

which two residues of the substrate span ~0.6 nm, and the SC/2R model predicts corresponding step sizes in terms of two residues translocated or ~0.6 nm traversed. By contrast, the basic translocation step of ClpXP measured in optical-trapping experiments is ~1 nm (Aubin-Tam et al., 2011; Cordova et al., 2014; Iosefson et al., 2015b; Maillard et al., 2011; Olivares et al., 2017; Rodriguez-Aliaga et al., 2016; Sen et al., 2013). Although ~1 nm is larger than ~0.6 nm, this difference, by itself, is not a compelling argument against the SC/2R model. In the optical trap, however, step size represents the average distance that unfolded polypeptide outside of the axial channel moves between successive translocation steps, which can be converted into amino-acid residues using the wormlike-chain model (Bustamante et al., 1994). Because the unfolded substrate outside the channel is in a partially compact conformation at the forces used in these experiments, ~1 nm corresponds to a translocation step of 5-8 residues, which appears inconsistent with the two-residue step predicted by the SC/2R model.

Because the sensitivity of optical trapping precludes direct identification and quantification of translocation steps as short as two residues, it could be argued that the single 5-8 residue steps observed for ClpXP actually consist of three or four unresolved SC/2R sub-steps. Consideration of kinetics makes this possibility unlikely, however. For example, the time from the beginning to the end of each translocation step is less than 0.1 s in the optical trap (Aubin-Tam et al., 2011; Cordova et al., 2014; Maillard et al., 2011; Olivares et al., 2017; Sen et al., 2013). Under similar conditions, the steady-state rate of $ClpX^{\Delta N}$/ClpP ATP hydrolysis is 3.6 ± 0.1 s^{-1} (Figure 4.13), corresponding to a time constant of 0.28 s. In the SC/2R model, each sub-step would require an independent ATP hydrolysis event, and ~1 s would be required to take three or four sub-steps. This kinetic problem becomes worse for ClpXP translocation bursts that move 20 to 32 residues in less

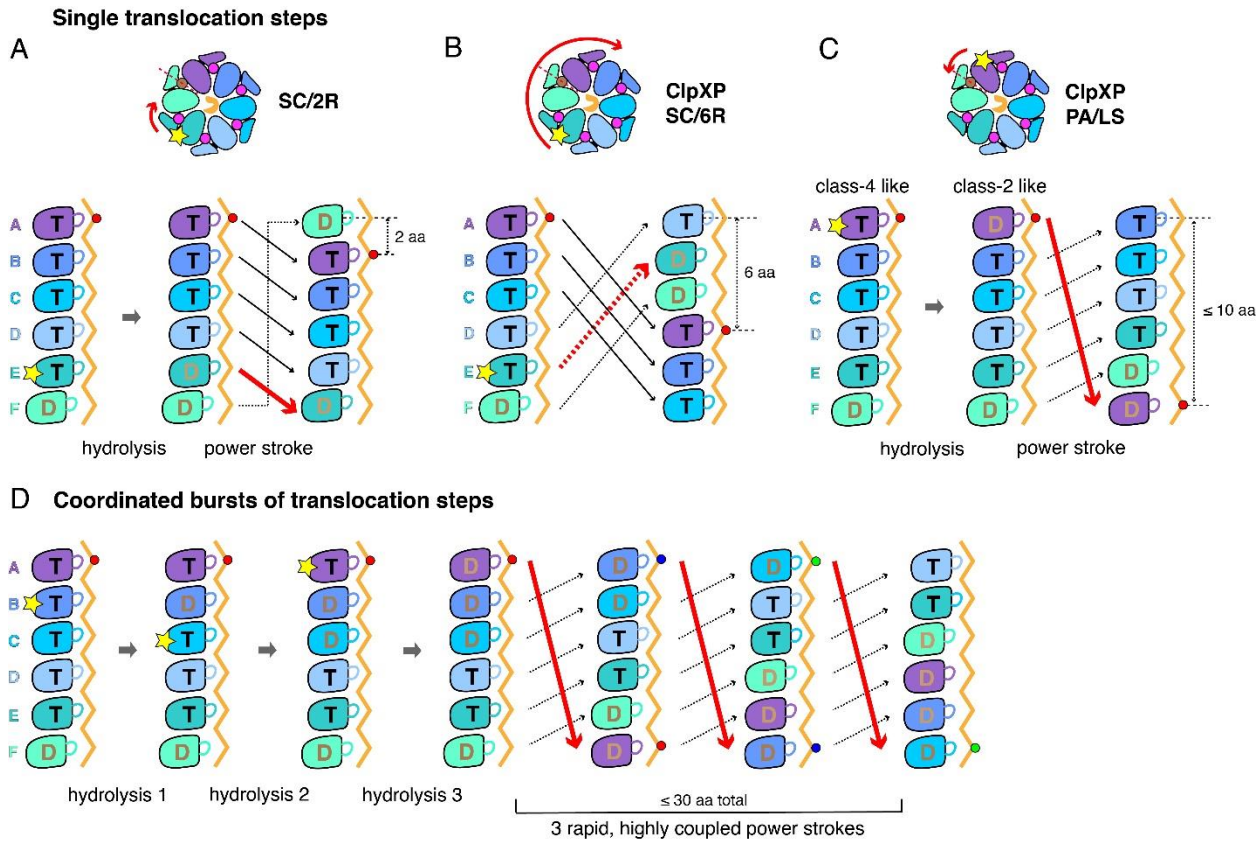


Figure 4.12 – Single-step and burst translocation models

In all panels, T represents ATP-bound subunits, and D represents subunits containing ADP and possibly inorganic phosphate. **(A)** SC/2R translocation model proposed for the Yme1 protease and 26S proteasome (Dong et al., 2019; de la Peña et al., 2018; Puchades et al., 2017). Only subunit E hydrolyzes ATP (depicted by a star). Hydrolysis and/or product release results in a two-residue translocation step, in which the top five subunits in the spiral move down one position in the clockwise direction and the bottom subunit moves up to the top position. **(B)** SC/6R translocation model. During a six-residue translocation step, subunits A, B, and C each move down in the spiral to positions D, E, and F, respectively, dragging substrate with them; at the same time, subunits D, E and F each move up to positions A, B, and C. Subunit displacement is in the clockwise direction. **(C)** PA/LS model in which ATP hydrolysis in subunit A results in a single translocation step of up to 2.5 nm in length as a consequence of anti-clockwise movement of this subunit to position F at the bottom of the spiral. At the same time, subunits BCDEF move up to positions ABCDE. **(D)** One variation of a PA/LS model resulting in a burst of three long translocation steps. Initial probabilistic ATP hydrolysis in subunits D and then B creates strain in the spiral, which is released in burst of fast steps upon ATP hydrolysis in subunit A. In the general PA/LS model, the initial ATP-hydrolysis event can occur with different probabilities in subunits A-E.

than 0.1 s. Thus, ClpXP translocation steps occur ~10 to ~40-fold faster than predicted by the SC/2R mechanism and the experimentally determined rate of ATP hydrolysis.

Another caveat might be that ClpX step sizes could be longer under tension in optical-trapping experiments, or that ClpXP uses a different translocation mechanism under these conditions. We consider both possibilities unlikely as the distribution of translocation step lengths shows little dependence on trap force (Aubin-Tam et al., 2011; Cordova et al., 2014; Maillard et al., 2011; Olivares et al., 2017; Sen et al., 2013) (Aubin-Tam et al., 2011; Maillard et al., 2011; Sen et al., 2013; Cordova et al., 2014; Olivares et al., 2017).

A separate concern involves the SC/2R prediction that only subunit E in the ClpX spiral hydrolyzes ATP during normal translocation, and thus that single-chain variants containing subunits that cannot hydrolyze ATP should stall when an ATPase inactive subunit moves into the hydrolysis position. However, $ClpX^{\Delta N}$ hexamers with just two ATPase-active subunits support ClpP degradation at ~30% of the wild-type rate and with the same thermodynamic efficiency as wild-type ClpXP (Martin et al., 2005). Thermal motions might move an ATPase-inactive subunit out of the hydrolysis position in the ClpX spiral and an active subunit into it, allowing continued function. This possibility seems unlikely, as the variant with just two ATPase active subunits translocates substrates directionally against force in optical-trap experiments (Cordova et al., 2014). Thus, at some level, ClpX translocation must operate probabilistically to avoid stalling in cases where only some of its subunits are hydrolytically active.

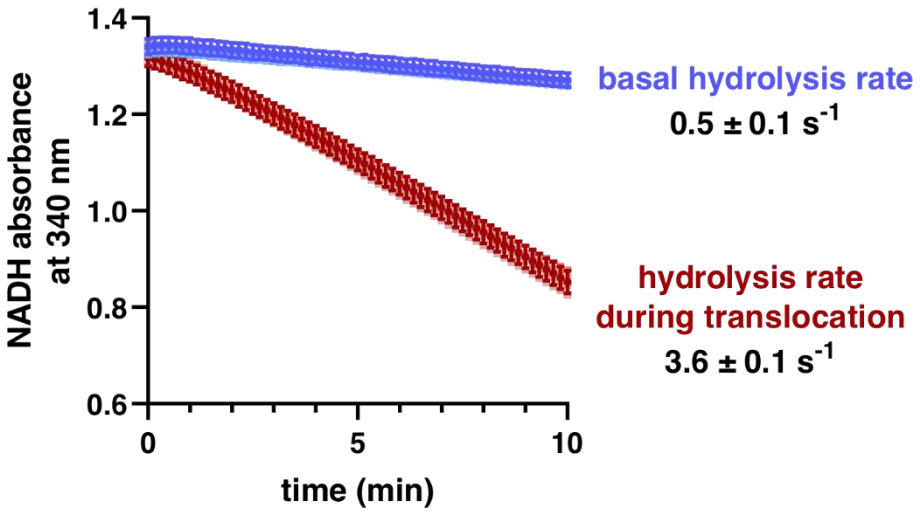


Figure 4.13 – ATP-hydrolysis rates

ATP-hydrolysis rates by $ClpX^{\Delta N}$/ClpP were measured using a coupled assay (Martin et al., 2005) that results in loss of NADH absorbance at 23 °C in PD buffer, conditions that approximate the conditions of optical-trapping experiments (Cordova et al., 2014). Shown are experiments in the absence (basal) or presence (translocation) of a near-saturating concentration (20 μM) of an unfolded $titin^{I27}$-ssrA substrate. The figure shows an overlay of three independent replicates. The rates shown are averages of the three independent replicates ± SD in units corresponding ATP molecules hydrolyzed per second per enzyme.

Structure-based translocation models for ClpXP

To account for the experimentally determined translocation properties of ClpXP, any model needs to explain: (1) how structural changes in the spiral result in a fundamental translocation step of 5-8 residues despite the observed two-residue periodicity of substrate contacts; (2) how kinetic bursts could generate very fast translocation of 20-32 residues without requiring multiple ADP-dissociation and ATP-rebinding events; and (3) how the motor functions efficiently without strict requirements for ATP hydrolysis in any particular subunit in the spiral.

Modifications of the SC/2R model could, in principle, address some of these issues. For example, upon ATP hydrolysis in subunit E, a clockwise movement a three-subunit shift could generate a six-residue power stroke in a SC/6R model (Figure 4.12B). Potential issues with this mechanism include that substrate contacts with subunits D, E, and possibly F would need to break during each translocation step; the structural features that would drive a three-subunit rearrangement are unclear; and this SC/6R model doesn't account for kinetic bursts resulting in translocation steps longer than six residues. By altering this clockwise three-subunit shift mechanism to allow probabilistic hydrolysis, long translocation bursts can be explained. Specifically, assume that multiple subunits in the spiral have some probability of hydrolyzing ATP, with the stipulation that hydrolysis at subunits other than E causes strain that is only released in a burst of six-residue power strokes when ATP in subunit E is eventually hydrolyzed or sufficient strain in the spiral accumulates (Probabilistic Clockwise/6-Reside Step or PC/6R model). Five of the six nucleotide-binding sites in our ClpX structures have similar geometries and thus each of these sites could plausibly hydrolyze ATP. Probabilistic models do not preclude neighboring subunits in the ring from firing sequentially, but rather exclude any mechanistic requirement for sequential action, and

are fully compatible with communication between ClpX subunits and cooperative ATP hydrolysis (Hersch et al., 2005; Martin et al., 2005).

Examination of our ClpXP structures reveals an alternative mechanism by which a fundamental translocation step of 5-8 residues might occur. Specifically, a power stroke that moves the top subunit in the ClpX spiral to the bottom position in the anti-clockwise direction could generate a movement of up to 2.5 nm through the axial channel. A conformational change of this type is captured in a morph in which subunit A in class 2 is aligned with subunit F in class 4. Figure 4.12C shows the corresponding translocation model in which probabilistic ATP hydrolysis in subunit A results in a single long step (Probabilistic Anti-clockwise Long-Step or PA/LS). The long-step nomenclature implies that step size need not be fixed in terms of residues, and could depend on conformational variability in the substrate sequence. A similar anticlockwise A-to-F subunit transition has been proposed for the Vps4 unfoldase (Su et al., 2017). Probabilistic ATP hydrolysis at multiple subunits in the spiral could then account for a kinetic burst of translocation steps, obviating the need to release ADP and rebind ATP after each step. Figure 4.12D shows one variation of this model. In any PA/LS model, subunit A would need to bind substrate and drag it to the bottom of the spiral, at least transiently breaking contacts with other subunits. Whether these contacts would be physically broken or simply released during the power stroke is currently unclear, as not all conformational states in the ClpX reaction cycle are likely to be known at this point. Crystal structures of ClpX reveal rotations between the large and small AAA+ domains so large that nucleotide cannot bind some subunits (Glynn et al., 2009; Stinson et al., 2013). Although subunits of this type are not present in the current cryo-EM structures, they could be representative of transient functional conformations, as crosslinks engineered to trap these hexamer

conformations form in solution and prevent ClpXP degradation but not ATP hydrolysis (Stinson et al., 2013, 2015).

Additional observations indicate that functionally relevant conformations remain to be discovered. For example, ClpXP degrades disulfide-bonded and knotted proteins in reactions that require simultaneous translocation of two or more polypeptides (Bolon et al., 2004; Burton et al., 2001; Martín et al., 2017; Sivertsson et al., 2019). In our cryo-EM structures, a single polypeptide strand fills the axial channel, and thus structures must exist in which the channel expands to accommodate multiple strands during translocation.

Other AAA+ unfolding/remodeling machines and SC/2R mechanisms

ClpAP takes ~1 nm and ~2 nm residue translocation steps (Olivares et al., 2014, 2017), suggesting that it can also operate by a non-SC/2R mechanism, but we are unaware of experiments that establish the translocation step size for other AAA+ protein unfolding or remodeling enzymes. Like ClpX, some of these enzymes do not stall if an ATPase-inactive subunit occupies the hydrolysis position. In the Rpt_{1-6} hexamer of the yeast 26S proteasome, for example, ATPase-defective Rpt_3, Rpt_4, or Rpt_6 subunits cause virtually complete loss of degradation activity, whereas ATPase-defective Rpt_1, Rpt_2, and Rpt_5 subunits do not (Beckwith et al., 2013). In the alternating Yta10/Yta12 hetero-hexamer (a Yme1 homolog), a Walker-B mutation in Yta12 prevents ATP hydrolysis in Yta10, as expected for tight subunit-subunit coupling in a strictly sequential mechanism, but a Walker-B mutation in Yta10 does not prevent ATP hydrolysis in Yta12 (Augustin et al., 2009). Given the strong similarities between the structures and substrate interactions of a large number of AAA+ unfolding and remodeling machines and the fact that the

SC/2R mechanism does not account for ClpX and ClpA experimental results, it will be important to continue to use structural and functional studies to investigate the molecular mechanisms by which these AAA+ machines function.

METHODS

Protein expression and purification

The single-chain ClpX variant used for microscopy contained six copies of *E. coli* E185Q ClpX$^{\Delta N}$ (residues 62-424), with neighboring units connected by six-residue linkers of variable composition, and a C-terminal TEV cleavage site and His_6 tag. The ClpP variant consisted of full-length *E. coli* ClpP followed by a TEV cleavage site and His_6 tag. Both proteins were expressed separately in *E. coli* strain ER2566 and purified as described (Stinson et al., 2013). After purification, TEV protease was used to remove the His_6 tags. TEV protease and uncleaved proteins were removed by Ni^{2+}-NTA affinity chromatography, and purified proteins were flash frozen in storage buffer (20 mM HEPES, pH 7.5, 300 mM KCl, 0.5 mM EDTA, 10% glycerol) and stored at –80 °C.

Sample preparation, data acquisition, and image processing

To assemble complexes, the ClpX$^{\Delta N}$ pseudohexamer and $ClpP_{14}$ were diluted into EM buffer (20 mM HEPES, pH 7.5, 100 mM KCl, 25 mM $MgCl_2$, 5 mM ATPγS) to final concentrations of 4 μM and 1.8 μM, respectively. After 5 min at 25 °C, 3 μL of the mixture was applied to glow discharged R1.2/1.3 400 mesh grids (Quantifoil). Grids were blotted with filter paper 494 (VWR) and plunged into liquid ethane using a Cryoplunge-3 system (Gatan). Electron micrographs were collected using a Talos Arctica with a Gatan K2-Summit direct electron detector in super-resolution mode.

High-resolution movies were recorded at a magnification of 36000X (0.58 Å pixel size). Each movie was composed of fifty frames (200 ms per frame) and a total dose of ~58 $e^{-}/Å^2$ per movie. The final dataset consisted of 3657 movies recorded in two separate sessions. Frames in each movie were 2X binned, aligned, gain-corrected, and dose-weighted using Motioncor2 (Zheng et al., 2017), to generate a single micrograph. The contrast transfer function (CTF) was estimated using CTFFIND4 (Rohou and Grigorieff, 2015). Unless noted, Relion 2.0 (Kimanius et al., 2016) was used for 2D/3D classification and refinement.

We first attempted to construct a density map of doubly capped ClpX-ClpP-ClpX using standard protocols. 1.4 million doubly capped particles were automatically picked and filtered by 2D classification. 443,717 "good" particles were selected for 3D map reconstruction. To generate an initial model for 3D refinement, the crystal structures of a $ClpX^{\Delta N}$ hexamer (PDB 3HWS) (Glynn et al., 2009) and a ClpP tetradecamer (PDB code 3MT6) (Li et al., 2010) were merged in PyMOL and low-pass filtered to 40 Å. 3D refinement with C_2 symmetry yielded a map with a resolution of ~4 Å, but the quality of this map was poor and interpretation of secondary structure elements was impossible. Using C_1 or C_7 symmetry did not improve the map.

To minimize problems caused by the symmetry mismatch between ClpX and ClpP, we treated ClpX and ClpP separately before the last step of refinement. For the ClpP reconstruction, we applied a soft circular mask, including ClpP and the tips of the ClpX IGF loops, to the 3D reference and 2D images, respectively. Because we observed predominantly side views of ClpXP (perpendicular to the axial channel; Figure 4.2A-C), this simple masking of the particle images allowed us to remove most of the ClpX density and to focus particle alignment on ClpP. Starting

from a low-pass filtered ClpP structure as an initial reference, 3D refinement with D_7 symmetry yielded a 3.2-Å resolution map with clear secondary structure and side-chain features. For the ClpX reconstruction, we extracted ClpX sub-particles from both ends of ClpP based on the previously determined ClpP alignment using a python script and IMOD (Mastronarde and Held, 2017). We prepared an original and a ClpP-signal-subtracted ClpX particle stack. To remove misaligned particles, the ClpX stacks were 2D classified without alignment. Particles from 2D classes that showed clear secondary structure were used in subsequent 3D-classification and reconstruction steps.

Clean ClpX sub-particles were 3D classified (K=6) and refined, resulting in four distinct ClpX classes with resolutions ranging from 3.9 to 4.2 Å. To recover the $ClpX^{\Delta N}$-ClpP interface, we re-extracted ClpXP sub-particles using a larger box that included the cis ClpP ring. Alignment and classification of each $ClpX^{\Delta N}$ sub-particle was transferred to corresponding $ClpX^{\Delta N}$/ClpP sub-particle and four classes were refined with local alignment optimization, resulting in four $ClpX^{\Delta N}$/ClpP maps (resolutions 4.0 to 4.3 Å).

To test the robustness of this workflow, we performed 3D classification of ClpX sub-particles multiple times using K=4 or K=8. The quality of maps suffered slightly, but the overall structures of four predominant classes of ClpX hexamers remained unchanged.

Model building and refinement

A ClpP tetradecamer (PDB 3MT6) (Li et al., 2010) was docked into the D_7 map using Chimera's "fit to map" function (Pettersen et al., 2004). For $ClpX^{\Delta N}$, six copies of the large and small AAA+

domains (PDB 3HWS) (Glynn et al., 2009) were docked into the map sequentially and refined for three iterations in Chimera. Real-space refinement of docked ClpP and ClpX was performed using PHENIX (Adams et al., 2010), and model building was performed using COOT (Emsley and Cowtan, 2004). ChimeraX (Goddard et al., 2018), Chimera (Pettersen et al., 2004), and PyMOL were used to create figures.

Biochemical assays

Assays were conducted at 37 °C in PD buffer (25 mM HEPES-KOH, pH 7.5, 5 mM $MgCl_2$, 200 mM KCl, 10% glycerol). Experiments were performed in triplicate and reported values are averages ± SD. Degradation of ^{CP7}GFP-ssrA (15 μM) by $ClpX^{\Delta N}$ (0.3 μM hexamer) and $ClpP_{14}$ (0.9 μM) was assayed in the presence of 5 mM ATP and was monitored by loss of fluorescent signal (excitation 467 nm; emission 511 nm). For RKH loop mutants with reduced substrate affinity, degradation was measured at high concentrations of fluorescent substrate, with excitation at an off-peak wavelength (excitation 420 nm; emission 511 nm). Degradation of fluorescein-labeled Arc-st11-ssrA by $ClpX^{\Delta N}$ (0.3 μM hexamer) and $ClpP_{14}$ (0.9 μM) was assayed as described (Bell et al., 2018). K_M and V_{max} were determined by fitting the average values of replicates to a hyperbolic equation. ATP hydrolysis rates were measured using a coupled-NADH oxidation assay as described (Martin et al., 2005), using $ClpX^{\Delta N}$ (0.3 μM hexamer), with or without $ClpP_{14}$ (0.9 μM), and 5 mM ATP. Activation of decapeptide cleavage by ClpP or variants was performed as described (Lee et al., 2010b), using $ClpP_{14}$ (50 nM), RseA decapeptide (15 μM), ATP (5 mM) and a regeneration system, and either $ClpX^{\Delta N}$ (0.5 μM hexamer) or ADEB-2B (100 μM), which was a generous gift from J. Sello (Brown). Pull-down experiments were performed in 25 mM HEPES-KOH (pH 7.5), 5 mM $MgCl_2$, 150 mM KCl, 20 mM imidazole, 10% glycerol, 500 μM

dithiothreitol and 2 mM ADP or ATPγS. 40 μL of a mixture of $ClpX^{\Delta N}$ C169S (1 μM hexamer concentration) and ClpP (1 μM 14-mer concentration) were mixed, incubated for 10 min at 30 °C, and then added to 20 μL of Ni^{2+}-NTA resin (Thermo Scientific) equilibrated in the same buffer. Binding reactions were incubated for 15 min at room temperature with rotation, centrifuged for 1 min at 9400 x g, and the supernatant was discarded. Reactions were washed, centrifuged, and the supernatant discarded three times. Bound protein was eluted with 40 μL of buffer supplemented with 300 mM imidazole for 15 min with rotation. Reactions were then centrifuged again and the eluant collected. Input and elution samples for each reaction were resolved by SDS-PAGE on a 10% Bis-Tris/MES gel run at 150V and visualized by staining with Coomassie Brilliant Blue R250. Results were validated in three independent replicates.

ACKNOWLEDGMENTS

Supported by NIH grant GM-101988 (R.T.S) and the Howard Hughes Medical Institute (T.A. Baker; S.C. Harrison). We thank S. Glynn, M. Lang, H. Manning, A. Olivares, and K. Schmitz for helpful discussions, J. Sello for providing ADEP-2B, and C. Xu and K. Song at the Electron Microscopy Facility at the University of Massachusetts Medical School for advice and data collection. The authors declare no competing financial interests.

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.

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.

Augustin, S., Gerdes, F., Lee, S., Tsai, F.T.F., Langer, T., and Tatsuta, T. (2009). An Intersubunit Signaling Network Coordinates ATP Hydrolysis by m-AAA Proteases. Mol. Cell *35*, 574–585.

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

Beckwith, R., Estrin, E., Worden, E.J., and Martin, A. (2013). Reconstitution of the 26S proteasome reveals functional asymmetries in its AAA+ unfoldase. Nat. Struct. Mol. Biol. *20*, 1164–1172.

Bell, T.A., Baker, T.A., and Sauer, R.T. (2018). Hinge–Linker Elements in the AAA+ Protein Unfoldase ClpX Mediate Intersubunit Communication, Assembly, and Mechanical Activity. Biochemistry *57*, 6787–6796.

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.

Bolon, D.N., Grant, R.A., Baker, T.A., and Sauer, R.T. (2004). Nucleotide-Dependent Substrate Handoff from the SspB Adaptor to the AAA+ ClpXP Protease. Mol. Cell *16*, 343–350.

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*, 1082–1087.

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*, 3092–3100.

Bustamante, C., Marko, J.F., Siggia, E.D., and Smith, S. (1994). Entropic elasticity of lambda-phage DNA. Science *265*, 1599–1600.

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*, 502–505.

Cordova, J.C., Olivares, A.O., Shin, Y., Stinson, B.M., Calmat, S., Schmitz, K.R., Aubin-Tam, M.-E., Baker, T.A., Lang, M.J., and Sauer, R.T. (2014). Stochastic but Highly Coordinated Protein Unfolding and Translocation by the ClpXP Proteolytic Machine. Cell *158*, 647–658.

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.

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

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*, 161–166.

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*, 273–279.

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.

Glynn, S.E., Martin, A., Nager, A.R., Baker, T.A., and Sauer, R.T. (2009). Structures of Asymmetric ClpX Hexamers Reveal Nucleotide-Dependent Motions in a AAA+ Protein-Unfolding Machine. Cell *139*, 744–756.

Glynn, S.E., Nager, A.R., Baker, T.A., and Sauer, R.T. (2012). Dynamic and static components power unfolding in topologically closed rings of a AAA+ proteolytic machine. Nat. Struct. Mol. Biol. *19*, 616–622.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., and Ferrin, T.E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. *27*, 14–25.

Grimaud, R., Kessel, M., Beuron, F., Steven, A.C., and Maurizi, M.R. (1998). Enzymatic and Structural Similarities between theEscherichia coli ATP-dependent Proteases, ClpXP and ClpAP. J. Biol. Chem. *273*, 12476–12481.

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.

Hersch, G.L., Burton, R.E., Bolon, D.N., Baker, T.A., and Sauer, R.T. (2005). Asymmetric Interactions of ATP with the AAA+ ClpX6 Unfoldase: Allosteric Control of a Protein Machine. Cell *121*, 1017–1027.

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

Iosefson, O., Olivares, A.O., Baker, T.A., and Sauer, R.T. (2015b). Dissection of Axial-Pore Loop Function during Unfolding and Translocation by a AAA+ Proteolytic Machine. Cell Rep. *12*, 1032–1041.

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.

Keiler, K.C. (2015). Mechanisms of ribosome rescue in bacteria. Nat. Rev. Microbiol. *13*, 285–297.

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.-I., Levchenko, I., Fraczkowska, K., Woodruff, R.V., Sauer, R.T., and Baker, T.A. (2001). Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. Nat. Struct. Biol. *8*, 230–233.

Kimanius, D., Forsberg, B.O., Scheres, S.H., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. ELife *5*, e18722.

Kirstein, J., Hoffmann, A., Lilie, H., Schmidt, R., Rübsamen-Waigmann, H., Brötz-Oesterhelt, H., Mogk, A., and Turgay, K. (2009). The antibiotic ADEP reprogrammes ClpP, switching it from a regulated to an uncontrolled protease. EMBO Mol. Med. *1*, 37–49.

Kucukelbir, A., Sigworth, F.J., and Tagare, H.D. (2014). Quantifying the local resolution of cryo-EM density maps. Nat. Methods *11*, 63–65.

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*, 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*, 707–718.

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*, 959–969.

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*, 459–469.

Majumder, P., Rudack, T., Beck, F., Danev, R., Pfeifer, G., Nagy, I., and Baumeister, W. (2019). Cryo-EM structures of the archaeal PAN-proteasome reveal an around-the-ring ATPase cycle. Proc. Natl. Acad. Sci. *116*, 534–539.

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

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.

Martin, A., Baker, T.A., and Sauer, R.T. (2008a). Diverse Pore Loops of the AAA+ ClpX Machine Mediate Unassisted and Adaptor-Dependent Recognition of ssrA-Tagged Substrates. Mol. Cell *29*, 441–450.

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

Martín, Á.S., Rodriguez-Aliaga, P., Molina, J.A., Martin, A., Bustamante, C., and Baez, M. (2017). Knots can impair protein degradation by ATP-dependent proteases. Proc. Natl. Acad. Sci. *114*, 9864–9869.

Mastronarde, D.N., and Held, S.R. (2017). Automated tilt series alignment and tomographic reconstruction in IMOD. J. Struct. Biol. *197*, 102–113.

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.

Ortega, J., Lee, H.S., Maurizi, M.R., and Steven, A.C. (2002). Alternating translocation of protein substrates from both ends of ClpXP protease. EMBO J. *21*, 4938–4949.

Pavelka, A., Sebestova, E., Kozlikova, B., Brezovsky, J., Sochor, J., and Damborsky, J. (2016). CAVER: Algorithms for Analyzing Dynamics of Tunnels in Macromolecules. IEEE/ACM Trans. Comput. Biol. Bioinform. *13*, 505–517.

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*, eaav0725.

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*, 1605–1612.

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.

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.

Rizo, A.N., Lin, J., 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*, 1–12.

Rodriguez-Aliaga, P., Ramirez, L., Kim, F., Bustamante, C., and Martin, A. (2016). Substrate-translocating loops regulate mechanochemical coupling and power production in AAA+ protease ClpXP. Nat. Struct. Mol. Biol. *23*, 974–981.

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

Sauer, R.T., and Baker, T.A. (2011). AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annu. Rev. Biochem. *80*, 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*, 636–646.

Shin, M., Asmita, A., Puchades, C., Adjei, E., Wiseman, R.L., Karzai, A.W., and Lander, G.C. (2019). Distinct Structural Features of the Lon Protease Drive Conserved Hand-over-Hand Substrate Translocation. BioRxiv 617159.

Siddiqui, S.M., Sauer, R.T., and Baker, T.A. (2004). Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. Genes Dev. *18*,

369–374.

Sivertsson, E.M., Jackson, S.E., and Itzhaki, L.S. (2019). The AAA+ protease ClpXP can easily degrade a 3 1 and a 5 2 -knotted protein. Sci. Rep. *9*, 1–14.

Stinson, B.M., Nager, A.R., Glynn, S.E., Schmitz, K.R., Baker, T.A., and Sauer, R.T. (2013). Nucleotide Binding and Conformational Switching in the Hexameric Ring of a AAA+ Machine. Cell *153*, 628–639.

Stinson, B.M., Baytshtok, V., Schmitz, K.R., Baker, T.A., and Sauer, R.T. (2015). Subunit asymmetry and roles of conformational switching in the hexameric AAA+ ring of ClpX. Nat. Struct. Mol. Biol. *22*, 411–416.

Su, M., Guo, E.Z., Ding, X., Li, Y., Tarrasch, J.T., Brooks, C.L., Xu, Z., and Skiniotis, G. (2017). Mechanism of Vps4 hexamer function revealed by cryo-EM. Sci. Adv. *3*, e1700325.

Sun, S., Li, L., Yang, F., Wang, X., Fan, F., Yang, M., Chen, C., Li, X., Wang, H.-W., and Sui, S.-F. (2017). Cryo-EM structures of the ATP-bound Vps4 E233Q hexamer and its complex with Vta1 at near-atomic resolution. Nat. Commun. *8*, 1–13.

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*, eaax1033.

Wang, J., Hartling, J.A., and Flanagan, J.M. (1997). The Structure of ClpP at 2.3 Å Resolution Suggests a Model for ATP-Dependent Proteolysis. Cell *91*, 447–456.

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.

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. *115*, E9560–E9569.

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*, 717–725.

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*, 331–332.

Chapter V

Rotation of ClpX with respect to ClpP is not required for polypeptide translocation or degradation of metastable protein substrates

This chapter has been prepared for publication as:

Bell, Tristan A., Fei, X., Baker, Tania A., and Sauer, Robert T. (2019). Rotation of ClpX with respect to ClpP is not required for polypeptide translocation or degradation of metastable protein substrates.

T.A. Bell performed crosslinking and biochemical experiments. T.A. Baker and R.T.S. oversaw research. T.A. Bell, X.F., and R.T.S. designed experiments and contributed to writing and/or revising the manuscript.

ABSTRACT

E. coli ClpXP is a AAA+ protease that harnesses the energy of ATP hydrolysis to mechanically unfold, translocate, and degrade target proteins. Ring hexamers of the ClpX unfoldase form a symmetry-mismatched interface with heptameric rings of the $ClpP_{14}$ peptidase, and cryo-EM structures coupled with some translocation models posit relative rotation of the ClpX and ClpP rings during substrate processing. We directly tested whether ClpX and ClpP must rotate relative to each other by engineering a complex in which rotation is prevented by a covalent crosslink. Preventing rotation did not substantially inhibit ATP hydrolysis, substrate binding, or degradation of unfolded substrates. It did slow degradation of several folded substrates and had a severe effect on unfolding/degradation of GFP, the most stable substrate tested. Thus, rotation of ClpX relative to ClpP is not required for degradation of unfolded and many folded substrates. We discuss the implications of these results and present translocation models in which the interface between ClpX and ClpP is static with respect to rotation.

INTRODUCTION

Members of the AAA+ (ATPases Associated with Various Cellular Activities) enzyme family convert chemical energy into cellular mechanical work (Erzberger and Berger, 2006). Within this family, AAA+ proteases use ATP to unfold, translocate, and degrade proteins that are misfolded, toxic, or unneeded by the cell (Sauer and Baker, 2011). Bacterial ClpXP, a model AAA+ protease, is composed of $ClpX_6$, a ring hexamer and ATP-dependent protein unfoldase/translocase, and $ClpP_{14}$, a double-ring ATP-independent peptidase (Figure 5.1A) (Baker and Sauer, 2012). ClpX and ClpP interact through a symmetry-mismatched interface in which six ClpX IGF loops, named for a conserved Ile-Gly-Phe motif, dock into six of the seven hydrophobic clefts on the surface of a heptameric ClpP ring (Figure 5.1B) (Chapter IV) (Amor et al., 2016, 2019; Martin et al., 2007). ClpX recognizes and binds target proteins bearing an appropriate degradation tag (degron), including the C-terminal ssrA tag (Baker and Sauer, 2006). Force is generated by ClpX conformational changes powered by ATP hydrolysis and is transmitted to a region of the bound substrate within the axial channel of the hexamer (Chapter IV) (Martin et al., 2008a, 2008b). Repeated cycles of ATP hydrolysis ultimately induce substrate unfolding, facilitating translocation of the denatured polypeptide through the axial channel and into the chamber of ClpP, where it is cleaved into peptide fragments.

Structures of multiple AAA+ proteases (including ClpXP) and protein-remodeling motors have recently been solved using cryo-electron microscopy (cryo-EM) (Chapter IV) (Cooney et al., 2019; Deville et al., 2017; Ding et al., 2017; Dong et al., 2019; Gates et al., 2017; Gatsogiannis et al., 2019; Han et al., 2017; Lee et al., 2010; Lo et al., 2019; Majumder et al., 2019; Michalska et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017, 2019; Ripstein et al.,

2017; Rizo et al., 2019; Su et al., 2017; Sun et al., 2017; Twomey et al., 2019; White et al., 2018; Yokom et al., 2016; Yu et al., 2018; Zehr et al., 2017; Zhu et al., 2018). In nearly all of these structures, the AAA+ unfolding motor adopts a spiral conformation, with a seam between the highest and lowest subunits (Figure 5.1A). In AAA+ proteases, this spiral motor structure is observed both when the motor and peptidase are symmetry mismatched (e.g., ClpXP and the 26S or PAN•20S proteasomes) and when they are genetically tethered and therefore symmetry matched (e.g., Yme1 and Lon) (Dong et al., 2019; Majumder et al., 2019; de la Peña et al., 2018; Puchades et al., 2017, 2019). In heterohexameric motors, multiple conformational states with different subunits occupying the highest and lowest spiral positions have been observed, suggesting that dynamic rearrangement of the spiral is part of the operational mechanical cycle (Dong et al., 2019; de la Peña et al., 2018).

For ClpX and other AAA+ protein-unfolding motors, several mechanisms have been proposed for coupling conformational changes driven by ATP hydrolysis into a processive mechanical cycle. In the most popular model, the second lowest subunit in the spiral hydrolyzes ATP resulting in a rearrangement that moves this subunit and higher subunits one position down in the spiral, with the lowest subunit disengaging from substrate and moving to the top of the spiral (Cooney et al., 2019; Deville et al., 2017; Ding et al., 2017; Dong et al., 2019; Gates et al., 2017; Han et al., 2017; Majumder et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017, 2019; Ripstein et al., 2017; Rizo et al., 2019; Twomey et al., 2019; Yu et al., 2018). This model, in which each subunit of the hexamer sequentially passes through each position in the spiral, nicely accounts for processive substrate translocation by positing that the post-hydrolysis state of one power stroke

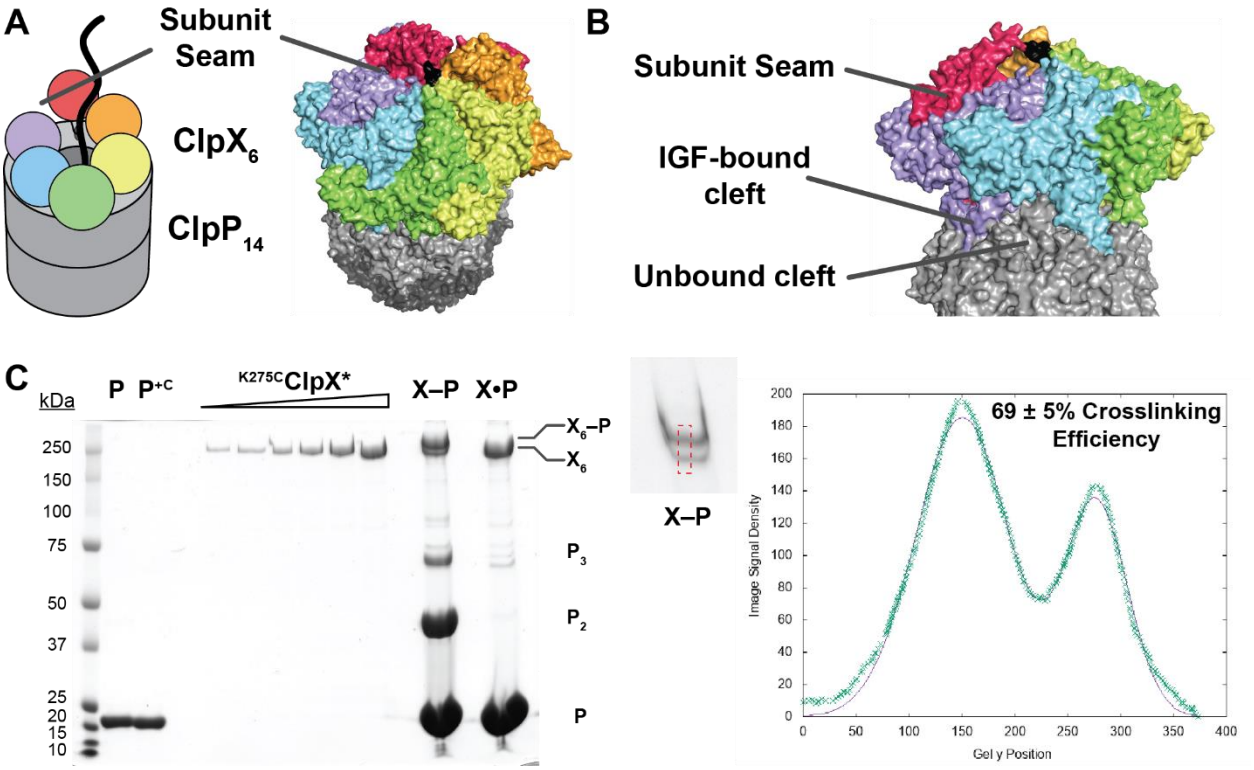


Figure 5.1 – ClpX-ClpP Crosslinking

(A) Schematic (left) and cryo-EM model (right) of ClpX–ClpP–substrate interaction. The seam between the highest (red) and lowest (purple) subunits is indicated. Molecular model adapted from Chapter IV. **(B)** In cryo-EM structures of ClpXP, the symmetry mismatched interface between the ClpX hexamer and ClpP heptamer is resolved by an empty ClpP cleft positioned between the lowest and second-lowest ClpX subunit. **(C)** SDS-PAGE (left) and crosslinking efficiency analysis (right) of ^{K275C}ClpX*, crosslinked X–P, and uncrosslinked XP samples. Extra bands in X–P sample correspond to nonspecific ClpP–ClpP crosslinking.

is positioned as the pre-hydrolysis state of the next, obviating the need for substrate handoff or rearrangement between hydrolysis events.

The ClpX spiral is able to dock with the flat ClpP ring because the IGF loops of ClpX are flexible and can expand or contract depending upon the distance between a subunit and the surface of ClpP (Chapter IV) (Gatsogiannis et al., 2019). Consistent with this flexibility, ClpX with mutations that increase or decrease the length of its IGF loops can still form a functional ClpXP complex (Amor et al., 2019). Intriguingly, in known cryo-EM structures, the ClpX spiral adopts a single orientation relative to ClpP, with the one unoccupied ClpP cleft located between the IGF loops of the lowest and second-lowest ClpX subunits (Chapter IV) (Figure 5.1B). This observation of a favored low-energy state suggests that the IGF loops of ClpX dynamically rearrange which ClpP clefts they bind in response to rearrangements in the topology of the ClpX spiral. Biochemical evidence supports a dynamic interface, as dissociation of ClpX from ClpP is normally very slow but becomes very fast in response to subsaturating doses of acyldepsipeptides (ADEPs), a class of small-molecule inhibitors that compete with the IGF loops for binding to the ClpP hydrophobic clefts (Amor et al., 2016; Martin et al., 2007).

If subunits in the ClpX hexamer pass sequentially through each position in the spiral and the empty cleft in ClpP is always between the lowest two subunits, then the ClpX ring should rotate with respect to the ClpP ring during protein translocation. Here, we test this possibility directly by crosslinking a single ClpX IGF loop to a ClpP cleft, thereby preventing rotation. Biochemical characterization of this crosslinked complex reveals only a small effect on substrate translocation, as assayed by degradation of an unfolded substrate, a modest slowing of

degradation of several native substrates with intermediate stabilities, and a severe slowing of degradation of the most stable native protein tested.

RESULTS

Engineering a singly-crosslinked ClpX-ClpP interface

In cryo-EM structures of *Escherichia coli* ClpXP, the penultimate C-terminal residue of ClpP forms part of the binding cleft for the IGF loops of ClpX, and Lys^{275} in the IGF loop of a ClpX subunit should be close enough to the C-terminal region of a ClpP subunit to allow crosslinking (Chapter IV). For these studies, we introduced a K275C mutation into one IGF loop of a single-chain hexamer of *E. coli* C169S $ClpX^{\Delta N}$, a variant that is active in degrading ssrA-tagged substrates despite lacking the N-terminal domain and replacing a surface cysteine with serine (Martin et al., 2005). We refer to this variant as $^{K275C}ClpX^{*}$. We also engineered an *E. coli* ClpP variant with a cysteine between the normal C-terminal residue (Asn^{193}) and an affinity tag. We call this variant $ClpP^{+C}$. $^{K275C}ClpX^{*}$ was mixed with a 25-fold excess of $ClpP^{+C}$ in the presence of ATPγS, which supports ClpXP assembly, and disulfide crosslinking was catalyzed with copper phenanthroline. In a control reaction, $^{K275C}ClpX^{*}$ was mixed with a 25-fold excess of wild-type ClpP and treated with copper phenanthroline. Both reactions were quenched with an excess of *N*-propylmaleimide (NPM) to modify any remaining reactive cysteines.

Denaturing SDS-PAGE revealed that the $^{K275C}ClpX^{*}$ band was super-shifted only in the reaction that included $ClpP^{+C}$. Crosslinking was measured to be 69 ± 5% efficient (Figure 5.1C). Some crosslinking between ClpP protomers was also observed in the reaction containing $ClpP^{+C}$ (Figure 5.1C). We attempted to purify the major crosslinked species in the $ClpP^{+C}$ reaction away from the

minor amount of uncrosslinked species but were unable to recover sufficient amounts of homogeneous crosslinked enzyme for biochemical assays. Hence, we analyzed the crosslinked (X–P) and uncrosslinked (X•P) pools.

Crosslinked complexes hydrolyze ATP and bind protein substrate

We determined rates of ATP hydrolysis by the crosslinked X–P and control X•P samples on their own, in the presence of ^{CM}titin^{I27}-ssrA (15 μM), a substrate unfolded by carboxymethylation of normally buried cysteines (Kenniston et al., 2003), or in the presence of GFP-ssrA (15 μM), a stable native substrate. Rates of ATP hydrolysis for the X–P and X•P samples were very similar in the absence of protein substrate and were stimulated to similar extents by both ^{CM}titin^{I27}-ssrA and GFP-ssrA (Figure 5.2A). Although the crosslinked X–P pool represents a mixture of ~70% crosslinked and ~30% uncrosslinked particles, similar ATP-hydrolysis rates should only be observed if the crosslinked and uncrosslinked enzymes have similar hydrolysis activities. Thus, these results indicate that crosslinking neither impairs ATP hydrolysis nor prevents substrate binding.

Crosslinked complexes translocate unfolded ^{CM}titin^{I27}-ssrA normally

As a measure of translocation activity, we used SDS-PAGE to assay degradation of unfolded ^{CM}titin^{I27}-ssrA (10 μM) by the X–P and X•P samples in the presence of 5 mM ATP and quantified disappearance of the substrate band over time (Figure 5.2B). The X–P sample degraded this substrate at 77 ± 7% of the rate of the X•P control (Figure 5.2B). Control reactions containing ADP instead of ATP did not support degradation by either pool (Figure 5.2B). Thus, crosslinked

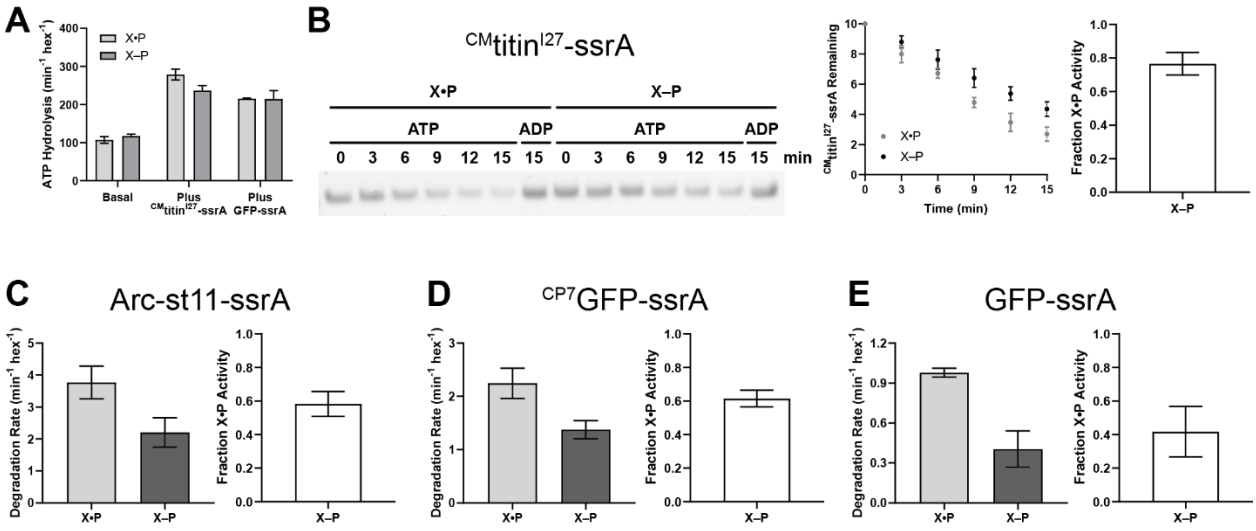


Figure 2 – Biochemical characterization of crosslinked X–P sample

(A) Basal and substrate-stimulated ATP hydrolysis activity of control (X•P) and crosslinked (X–P) samples. In all quantified biochemical experiments, values represent average ± S.D. of at least three independent replicates. **(B)** (left) SDS-PAGE analysis of ^{CM}titin^{I27}-ssrA degradation by X•P and X–P samples. (right) Quantification of ^{CM}titin^{I27}-ssrA degradation. **(C)** Degradation of Arc-st11-ssrA by X•P and X–P samples. **(D)** Degradation of ^{CP7}GFP-ssrA by X•P and X–P samples. **(E)** Degradation of GFP-ssrA by X•P and X–P samples.

X–P complexes appear to translocate unfolded substrates in an ATP-dependent manner at only modestly slower rates than uncrosslinked complexes.

Crosslinked complexes show a range of defects in degrading native substrates

We next assayed degradation of two native protein substrates (Arc-st11-ssrA and ^{CP7}GFP-ssrA). The crosslinked X–P pool degraded fluorescently labeled Arc-st11-ssrA (10 μM) at 58 ± 7% of the rate of the X•P pool and degraded ^{CP7}GFP-ssrA (10 μM) at 61 ± 5% of the control rate (Figures 5.2C–D). These levels of degradation are roughly twice the ~30% value expected if only the uncrosslinked enzyme in the X–P pool were active. Therefore, crosslinking the ClpX-ClpP interface modestly impairs but does not prevent degradation of these natively folded substrates.

GFP-ssrA is a very stable native protein and is a difficult substrate for ClpXP and other AAA+ proteases to processively degrade, in part because the GFP barrel remains intact upon enzymatic extraction of its C-terminal β-strand, and thus complete unfolding and degradation require a fast second step to successfully induce cooperative unfolding (Maillard et al., 2011; Nager et al., 2011). The X–P pool degraded GFP-ssrA (10 μM) at 40 ± 20% of the control X•P rate, a value not substantially different from the fraction of uncrosslinked ClpX in the X–P pool (Figure 5.2E). Thus, crosslinked complexes appear to degrade GFP-ssrA either very slowly or not at all.

DISCUSSION

Many aspects of protein unfolding, translocation, and degradation by ClpXP are well established (Bell et al., 2019; Cordova et al., 2014; Iosefson et al., 2015; Kenniston et al., 2003; Martin et al.,

2005; Sen et al., 2013; Siddiqui et al., 2004). It is unclear, however, whether the conformational dynamics of subunits within the ClpX spiral during substrate translocation and unfolding require reprogramming of IGF-loop contacts with ClpP clefts. All substrate-bound cryo-EM structures of ClpXP show an invariant pattern of contacts between the six ClpX-IGF loops and seven ClpP clefts in which the empty cleft is always located between the clefts occupied by the IGF loops of the lowest and second-lowest ClpP-proximal subunits in the ClpX spiral (Chapter IV). The prevailing model of AAA+ motor function also posits that each subunit in the ClpX hexamer sequentially moves through each spiral position. During the hundreds of cycles of ATP hydrolysis that are required to unfold, translocate, and degrade protein substrates (Kenniston et al., 2003), this combination of structural observation and proposed mechanism predicts that the rings of ClpX and ClpP need to rotate with respect to each other to allow degradation. In the experiments presented here, we demonstrate that crosslinking ClpX to ClpP to prevent global IGF-loop repositioning and rotation of the ClpX and ClpP rings does not substantially inhibit substrate binding or translocation. It does modestly inhibit unfolding/degradation of some stably folded substrates and completely or near-completely abolishes unfolding/degradation of GFP-ssrA, a very stable protein substrate. Despite inhibiting unfolding/degradation, ATP hydrolysis activity is stimulated in the X–P pool to similar levels as the X•P pool, indicating that crosslinked X–P undergoes futile power strokes. Our results therefore challenge either the need for an invariant pattern of IGF-cleft contacts or for a strictly sequential mechanism of ClpX function. Below, we discuss three models that are consistent with ClpXP function without rotation between the ClpX and ClpP rings.

In one model, sequential cycling of individual ClpX subunits through different spiral positions drives substrate translocation, but function does not require specific IGF loops to bind specific

ClpP clefts in the low-energy and thus preferred orientation observed in cryo-EM structures (Figure 5.3A). Genetically-tethered AAA+ proteases in the Lon and FtsH/Yme1/Agf3l2 families can unfold and degrade substrates despite the topological inability for relative rotation between the unfoldase and protease components (Baker and Sauer, 2006; Glynn, 2017). Interestingly, these tethered enzymes seem to lack the ability to degrade GFP and certain other stably folded substrates that can be degraded by ClpXP, ClpAP, and the 26S proteasome (Herman et al., 2003; Koodathingal et al., 2009). Hexameric ClpA also uses flexible loops to bind clefts in heptameric ClpP rings, whereas C-terminal tripeptides from the AAA+ Rpt_{1-6} unfolding ring dock into binding pockets on the heptameric α ring of the 20S peptidase in the 26S proteasome. These facts coupled with our finding that crosslinked X–P is less effective than uncrosslinked X•P when challenged with increasingly stable native substrates make it possible that complete or partial rotation of the unfoldase and peptidase rings is not required but facilitates higher-level force production and thus more efficient unfolding.

A second model consistent with our findings is that ClpX uses a reciprocating mechanism that maintains a static rotational conformation relative to ClpP throughout substrate processing (Figure 5.3B). In the first step in this model, the highest subunit in the ClpX spiral applies force to the bound substrate by moving to the lowest position in the spiral. The other subunits then grip the substrate as the now-lowest (previously highest) subunit reverts to its original position for the next power stroke. As our X–P sample is less active against folded than unfolded substrates, a reciprocating mechanism is unlikely to be the sole mechanism of substrate unfolding and translocation, but could represent a fundamental mechanistic mode accessible to the motor. For example, a kinetic burst of power strokes could allow multiple subunits to sequentially move from

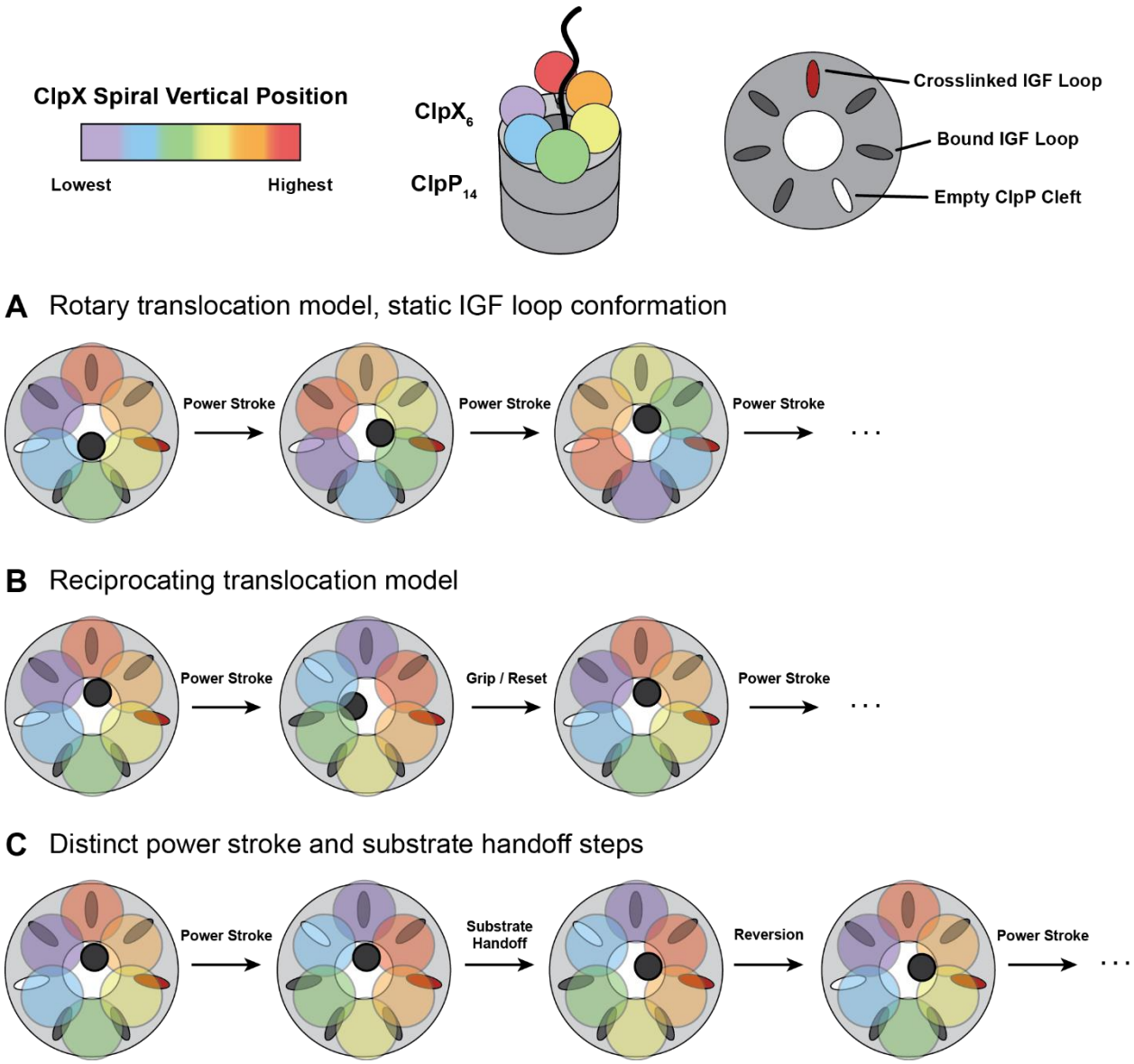


Figure 3 – Substrate translocation models

Models for substrate translocation by ClpXP consistent with at least one statically positioned IGF loop. In all models, figures look down through the axial pore of ClpX toward ClpP. **(A)** Substrate translocation by the widely-proposed rotary translocation model with static IGF loop contacts with ClpP. **(B)** Reciprocating translocation model in which a subset of IGF loops maintain static interactions with ClpP. **(C)** Model for rotary substrate translocation in which the presence of a ClpX–ClpP crosslink disrupts normal procession, but unidirectional substrate motion still occurs through distinct power stroke and substrate rearrangement steps.

the highest to the lowest position in the spiral, with accompanying rotation of the ClpX-ClpP interface, followed by a grip/reset step at the conclusion of the burst.

Finally, rotation between ClpX and ClpP may not be required for translocation because power strokes contain a distinct substrate handoff step that mitigates the antagonizing effect of the crosslink (Figure 5.3C). In any mechanism that posits rotation of ClpX relative to ClpP, the X–P crosslink would stall conformational changes that move the substrate within the channel and thus impair activity completely. However, the substrate could still advance processively if a second and distinct step occurs after force generation. For example, if the substrate rearranges its binding conformation within the axial pore after a power stroke, the motor could relax to its previous conformation without pulling the substrate backwards with it. In this substrate-handoff model, the presence of the crosslink would impair unfolding of stably folded substrates by limiting intersubunit coordination and the amount of force generated during a power stroke.

Although parsimony suggests that AAA+ motors should always use the same mechanism for substrate unfolding and translocation, this may not be true. Our observation that crosslinking ClpX to ClpP does not alter degradation of an unfolded substrate but has increasingly disruptive effects on degradation of more stable substrates might suggest that ClpXP is able to use distinct mechanisms of unfolding and translocation. For example, complete or partial rotation of ClpX relative to ClpP might be required to generate sufficient force to unfold very stable proteins like GFP, whereas a non-rotary mechanism may allow translocation of unfolded polypeptides and unfolding of metastable protein substrates.

METHODS

Crosslinking

Genes encoding *E. coli* ^{K275C}ClpX* and ClpP^{+C} were generated using PCR mutagenesis, and the corresponding proteins were purified by established protocols and stored in buffer containing 0.5 mM dithiothreitol (DTT) (Martin et al., 2005). For crosslinking reactions, ^{K275C}ClpX* (2.84 μM) and ClpP^{+C} (70.9 μM) or wild-type ClpP (70.9 μM) were mixed in a total volume of 400 μL. The reactions were desalted into Buffer PD (25 mM HEPES-KOH, pH 7.5, 200 mM KCl, 5 mM $MgCl_2$, 10% glycerol) twice over 0.5 mL Zeba Spin Desalting Columns (Thermo Scientific). Crosslinking was catalyzed by addition of copper (II) phenanthroline (50 μM) and ATPγS (2 mM), and allowed to proceed for 15 min at 30 °C. Reactions were then quenched with NPM (17.5 mM; Millipore-Sigma) and EDTA (17.5 mM) for 1 h at 37 °C and again desalted twice into Buffer PD over 0.5 mL Zeba Spin Desalting Columns.

Crosslinked species were measured by quantifying Coomassie-stained SDS-PAGE gel bands relative to ^{K275C}ClpX* standards using ImageJ (Schneider et al., 2012). Crosslinking efficiency was measured by tracing Coomassie stain density through the midpoint of a gel band using ImageJ. Individual points were recovered from the ImageJ trace by vectorizing the trace using Illustrator 6 (Adobe). The data points were then fit to a double Gaussian distribution using Gnuplot to determine the area under curve (AUC) of the crosslinked and uncrosslinked bands. Crosslinking efficiency was measured in five independent replicates, and calculated as:

$$Efficiency = \frac{AUC_{Crosslinked}}{AUC_{Crosslinked} + AUC_{Uncrosslinked}}$$

Biochemical assays

ATP hydrolysis and substrate degradation assays were performed as described with crosslinked X–P or control X•P samples (50 nM ^{K275C}ClpX* pseudohexamer) at 37 °C in Buffer PD with 5 mM ATP. ATP hydrolysis assays were performed using an NADH-coupled colorimetric assay, as described previously (Martin et al., 2005). Degradation assays included a regeneration mix composed of 32 mM creatine phosphate (Roche) and 0.08 mg/mL creatine kinase (Millipore-Sigma). For ^{CP7}GFP-ssrA and GFP-ssrA, loss of substrate fluorescence (excitation 467 nm; emission 511 nm) was measured using a SpectraMax M5 plate reader (Molecular Devices) (Nager et al., 2011). For Arc-st11-ssrA, degradation of substrate labeled with multiple fluorescein molecules was measured by tracking the increase in fluorescence (excitation 480 nm; emission 525 nm) caused by degradation-dependent decreases in fluorescein homo-quenching (Bell et al., 2018). ^{CM}titin^{I27}-ssrA was produced by unfolding and carboxymethylation of buried cysteine residues, as described previously (Kenniston et al., 2003). Degradation was visualized by SDS-PAGE, and Coomassie-stained bands were quantified using ImageJ. Fractional activity was calculated by dividing the activity of the X–P pool by the activity of the X•P pool. All experiments were performed in three independent replicates and values reported were calculated as average ± S.D. of independent replicates.

ACKNOWLEDGMENTS

This work was supported by US National Institutes of Health (NIH) grant GM-101988 (R.T.S.). T.A. Bell was supported in part by US NIH grant 5T32GM-007287. T.A. Baker is an employee of the Howard Hughes Medical Institute. We are grateful to A.J. Amor for sharing crosslinking data

from prior characterization of the ClpX-ClpP interface, and to current and past members of the Sauer and Baker labs for providing helpful feedback on the manuscript.

REFERENCES

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.

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

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

Bell, T.A., Baker, T.A., and Sauer, R.T. (2018). Hinge–Linker Elements in the AAA+ Protein Unfoldase ClpX Mediate Intersubunit Communication, Assembly, and Mechanical Activity. Biochemistry *57*, 6787–6796.

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.

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*, 502–505.

Cordova, J.C., Olivares, A.O., Shin, Y., Stinson, B.M., Calmat, S., Schmitz, K.R., Aubin-Tam, M.-E., Baker, T.A., Lang, M.J., and Sauer, R.T. (2014). Stochastic but Highly Coordinated Protein Unfolding and Translocation by the ClpXP Proteolytic Machine. Cell *158*, 647–658.

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*, e1701726.

Ding, Z., Fu, Z., Xu, C., Wang, Y., Wang, Y., Li, J., Kong, L., Chen, J., Li, N., Zhang, R., et al. (2017). High-resolution cryo-EM structure of the proteasome in complex with ADP-AlFx. Cell Res. *27*, 373–385.

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.

Erzberger, J.P., and Berger, J.M. (2006). Evolutionary Relationships and Structural Mechanisms of Aaa+ Proteins. Annu. Rev. Biophys. Biomol. Struct. *35*, 93–114.

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*, 273–279.

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.

Glynn, S.E. (2017). Multifunctional Mitochondrial AAA Proteases. Front. Mol. Biosci. *4*.

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.

Herman, C., Prakash, S., Lu, C.Z., Matouschek, A., and Gross, C.A. (2003). Lack of a Robust Unfoldase Activity Confers a Unique Level of Substrate Specificity to the Universal AAA Protease FtsH. Mol. Cell *11*, 659–669.

Iosefson, O., Nager, A.R., Baker, T.A., and Sauer, R.T. (2015). Coordinated gripping of substrate by subunits of a AAA+ proteolytic machine. Nat. Chem. Biol. *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.

Koodathingal, P., Jaffe, N.E., Kraut, D.A., Prakash, S., Fishbain, S., Herman, C., and Matouschek, A. (2009). ATP-dependent Proteases Differ Substantially in Their Ability to Unfold Globular Proteins. J. Biol. Chem. *284*, 18674–18684.

Lee, M.E., Baker, T.A., and Sauer, R.T. (2010). Control of Substrate Gating and Translocation into ClpP by Channel Residues and ClpX Binding. J. Mol. Biol. *399*, 707–718.

Lo, Y.-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*, 1–12.

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*, 459–469.

Majumder, P., Rudack, T., Beck, F., Danev, R., Pfeifer, G., Nagy, I., and Baumeister, W. (2019). Cryo-EM structures of the archaeal PAN-proteasome reveal an around-the-ring ATPase cycle. Proc. Natl. Acad. Sci. *116*, 534–539.

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

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.

Martin, A., Baker, T.A., and Sauer, R.T. (2008a). Diverse Pore Loops of the AAA+ ClpX Machine Mediate Unassisted and Adaptor-Dependent Recognition of ssrA-Tagged Substrates. Mol. Cell *29*, 441–450.

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

Michalska, K., Zhang, K., March, Z.M., Hatzos-Skintges, C., Pintilie, G., Bigelow, L., Castellano, L.M., Miles, L.J., Jackrel, M.E., Chuang, E., et al. (2019). Structure of Calcarisporiella thermophila Hsp104 Disaggregase that Antagonizes Diverse Proteotoxic Misfolding Events. Structure *27*, 449-463.e7.

Monroe, N., Han, H., Shen, P.S., Sundquist, W.I., and Hill, C.P. (2017). Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. ELife *6*, e24487.

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*, 4–16.

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*, eaav0725.

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.

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*, 1073-1085.e6.

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.

Rizo, A.N., Lin, J., 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*, 1–12.

Sauer, R.T., and Baker, T.A. (2011). AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. Annu. Rev. Biochem. *80*, 587–612.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

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*, 636–646.

Siddiqui, S.M., Sauer, R.T., and Baker, T.A. (2004). Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. Genes Dev. *18*, 369–374.

Su, M., Guo, E.Z., Ding, X., Li, Y., Tarrasch, J.T., Brooks, C.L., Xu, Z., and Skiniotis, G. (2017). Mechanism of Vps4 hexamer function revealed by cryo-EM. Sci. Adv. *3*, e1700325.

Sun, S., Li, L., Yang, F., Wang, X., Fan, F., Yang, M., Chen, C., Li, X., Wang, H.-W., and Sui, S.-F. (2017). Cryo-EM structures of the ATP-bound Vps4 E233Q hexamer and its complex with Vta1 at near-atomic resolution. Nat. Commun. *8*, 1–13.

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*, eaax1033.

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.

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*, 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. *115*, E9560–E9569.

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*, 717–725.

Zhu, Y., Wang, W.L., Yu, D., Ouyang, Q., Lu, Y., and Mao, Y. (2018). Structural mechanism for nucleotide-driven remodeling of the AAA-ATPase unfoldase in the activated human 26S proteasome. Nat. Commun. *9*, 1–12.

Chapter VI

Communication and coordination between subunits of the ClpX unfoldase

ABSTRACT

ClpXP is a model system for interrogating mechanisms of force generation and protein unfolding among the broader family of AAA+ unfoldase-proteases. In chapters II–V of this thesis, I presented four individual stories that illuminated novel aspects of ClpXP structure and function. Here, in an extended discussion, I frame these stories in the larger context of intersubunit communication and coordination by AAA+ motors. I highlight ClpX-specific structural elements and quaternary topologies that facilitate intersubunit coordination and expound on how this coordination mediates processive, rapid, and efficient substrate processing. Finally, I suggest a series of experiments and untested hypotheses that represent possible next steps in understanding the molecular mechanics of this remarkable class of molecular machines.

In an internal combustion engine, pistons work together within the constraints imposed by the engine block and fuel systems to produce the force needed to overcome load when rotating a crankshaft. Similarly, the six subunits of many AAA+ motors function together within the structural constraints of a ring hexamer topology. ClpX, a model AAA+ protein unfolding motor, exhibits positive cooperativity and coordination with respect to function. For example, the rate of ATP hydrolysis is positively cooperative, indicating that subunits sense the nucleotide-binding state of other subunits and respond by increasing their own probability of hydrolyzing ATP (Hersch et al., 2005). Mechanical activity is also coordinated, as ClpXP takes translocation steps along unfolded substrates too large to be driven by conformational changes in a single ClpX subunit (Cordova et al., 2014; Sen et al., 2013; Chapter IV). In some AAA+ motors, mechanical activity is further coordinated. For example, the double-ring motor NSF undergoes near-simultaneous conformational changes in all 12 of its subunits, sacrificing processivity to produce a massive power stroke that transmits extremely high forces to disassemble SNARE complexes (Ryu et al., 2015). Like the components of macroscale motors, individual AAA+ motor subunits must function together to ensure processivity, energetic efficiency, and high-level force production.

Most mechanisms proposed for processive translocation by ClpXP and other AAA+ motors posit some degree of communication between subunits of the ring hexamer that coordinates global mechanical activity (Cooney et al., 2019; Cordova et al., 2014; Deville et al., 2017; Ding et al., 2017; Dong et al., 2019; Gates et al., 2017; Han et al., 2017; Majumder et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017, 2019; Ripstein et al., 2017; Rizo et al., 2019; Sen et al., 2013; Su et al., 2017; Twomey et al., 2019; Yu et al., 2018; Chapter IV). However,

despite deep understanding of how the structural elements within motor subunits interact with ATP, substrates, and accessory domains, little is known about how the subunits function together as a larger machine. How do subunits communicate with one another? Is communication limited to subunits that contact each other directly or is longer-range communication also possible? If longer-range communication is relevant, what is its structural basis? Does the global architecture of the ring hexamer prime different subunits to fulfill different functional roles? Among genetically untethered proteases like ClpXP, do the distinct unfoldase and protease components mutually coordinate each other's activity, and how do their interactions change during processive substrate unfolding and translocation? Finally, how are discreet power strokes coupled to produce efficient, unidirectional substrate motion? To borrow again the example of an internal combustion motor, we are much like someone who understands how pistons and valves fit together in individual cylinders, and how pistons connect to the crankshaft, but does not understand how the structure and design of the entire engine coordinates chemical combustion and mechanical activity.

In this thesis, I have explored communication between AAA+ subunits and coordination of activity within the ClpX motor. In this final chapter, I discuss newly identified mechanisms of communication and the implications of these findings for the global mechanism of processive substrate degradation. I also propose experiments that are beyond the scope of this work but have the potential to expand our understanding of coordinated motor activity in AAA+ proteases and protein-remodeling machines.

Hinge-linkers are ClpX structural elements that facilitate intersubunit communication

In Chapter II, I presented structural and biochemical characterization of the hinge-linker element of ClpX, a short sequence that facilitates rotation between ClpX subunits in the assembled ring hexamer. Previous work demonstrated that altering the length of all six hinge-linkers in a hexamer dramatically decreased the energetic efficiency of substrate degradation (Glynn et al., 2012). My work extended this finding by structurally and functionally characterizing ClpX hexamers with a single hinge-linker deleted or disrupted.

I found that deleting one ClpX hinge-linker causes the hexamer to predominantly adopt open structures instead of closed rings. This result indicates that assembly of ClpX subunits into closed-ring structures results in energetic strain that is topologically overcome by covalent connections between the large and small AAA+ domains enforced by the hinge-linkers. Conformational strain provides a mechanistic basis for long-range communication within the hexamer, as conformational changes in one subunit could alter tension felt in all other subunits. Furthermore, I found that placing a disruptive insertion into one hinge-linker per hexamer preserved the closed-ring architecture but still caused loss of cooperativity in ATP hydrolysis similar to that resulting from a hinge-linker deletion. This mutation also decreased the substrate degradation rate and mechanical efficiency of the ring hexamer, indicating an increase in ATP-hydrolysis events that are not effectively coupled to substrate motion. I interpret these results to indicate that long-range strain-mediated communication through the hinge-linkers is a functionally important mechanism that coordinates the activity of subunits within the ring hexamer of ClpX. Without this strain-mediated communication, subunits appear to function with greater independence, and ATP hydrolysis no longer results in coordinated mechanical activity.

The experiments in Chapter II also demonstrate that binding to its partner protease ClpP helps to scaffold ClpX subunits into a ring hexamer. Acting alone, ClpX hexamers with a deleted hinge-linker or one disrupted by an insertion exhibited abnormally fast ATP hydrolysis. However, binding to ClpP decreased the hydrolysis rates of both disrupted ClpX variants to levels similar to ClpP-bound wild-type ClpX. Thus, ClpX may rely on both ClpP and the hinge-linker to scaffold it into a global architecture that is primed for optimal efficiency.

The spiral structure of ClpX suggests asymmetry in subunit–subunit interactions

In Chapter IV, I presented the results of a collaborative project that solved cryo-EM structures of substrate-bound *E. coli* ClpXP. Previous structures of ClpX hexamers solved by x-ray crystallography were roughly planar, with some subunits adopting a deep twist about their hinge-linkers that precluded nucleotide binding in those subunits (Glynn et al., 2009; Stinson et al., 2013). In contrast, the cryo-EM structures of ClpXP indicate that ClpX forms a shallow, right-handed spiral with a discontinuous seam between the highest- and lowest-positioned subunits. All ClpX subunits in the cryo-EM structures adopt conformations competent to bind nucleotide, with the highest and lowest subunits bound either to ADP or ATPγS, and the middle subunits bound to ATPγS. The pore-1 and pore-2 loops of different ClpX subunits interact with substrate, and the spiral architecture mediates a network of multivalent interactions with substrate across the entire length of the ClpX axial pore. In addition, ClpX adopts a single conformation in its symmetry-mismatched interface with ClpP, with six ClpP clefts bound by the six ClpX IGF loops and an empty cleft positioned between the lowest and second-lowest subunits in the ClpX spiral.

The asymmetry inherent in the spiral topology raises the possibility that different ClpX subunits could play functionally distinct roles in substrate processing. Several previous crystal structures of ClpX ring hexamers were nearly two-fold symmetric, raising questions about what might distinguish subunits during division of functions (Glynn et al., 2009; Stinson et al., 2013). For example, many translocation models for spiral-shaped AAA+ machines propose that hydrolysis and a power stroke only occurs when a subunit occupies a defined position within the spiral (Cooney et al., 2019; Deville et al., 2017; Ding et al., 2017; Dong et al., 2019; Gates et al., 2017; Han et al., 2017; Majumder et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017; Ripstein et al., 2017; Rizo et al., 2019; Su et al., 2017; Twomey et al., 2019; Yu et al., 2018). Other tasks, such as substrate grip (discussed below) could also be delegated to specific subunits if the subunits adopted sufficiently distinct conformations.

Although the spiral structure does not inherently suggest that processive activity occurs in a rotary manner, it does provide a convenient explanation for how different subunits could coordinate their activity by cycling through different positions within the spiral architecture. This model of rearrangement could occur during individual ClpX power strokes or as part of a kinetic burst of coordinated power strokes. An experiment presented in Chapter II provides some evidence that coordinated ClpX subunit activity during substrate unfolding does exhibit some level of directional bias. In this experiment, I placed a disruptive mutation into a single hinge-linker of a ClpX hexamer alongside mutations that prevent ATP hydrolysis in subunits on either side of the disrupted hinge-linker. The hinge-linker mutation impairs communication across the disrupted subunit-subunit interface, allowing me to test whether subunits coordinate their activity asymmetrically with their neighbors. As expected if intersubunit coordination is directionally biased, I observed a strong

asymmetry in how efficiently these motors degraded protein substrate. Although several unsubstantiated assumptions would be necessary to make structural predictions about whether subunits preferentially communicate with their adjacent higher or lower neighbors in the spiral, this result does indicate that the architecture of ClpX biases the directionality of subunit–subunit communication within the hexameric ring.

ClpX subunits coordinate functionally to grip substrates multivalently

The spiral architecture of substrate-bound ClpX may provide a structural basis for differentiation of subunit function, but do substrates actually adopt distinct roles? In Chapter III, I presented a systematic characterization of substrate grip by ClpXP. Using a substrate that was difficult to unfold, I showed that only the first seven amino acids adjacent to a domain form productive gripping interactions with the ClpX axial pore during unfolding of that substrate. Cross-correlation of this window of residues with the positions of residues in the bound substrate in cryo-EM structures of ClpXP (Chapter IV) suggested that these substrate residues contact pore-1 loops in the uppermost three subunits of the ClpX spiral. The structures indicate that pore-1 loops in lower subunits contact the substrate region beyond these first seven residues, but these interactions do not appear to generate the grip needed to unfold a substrate. This finding clearly indicates that subunits higher and lower in the spiral adopt functionally distinct roles in establishing grip during substrate unfolding.

The investigation into substrate grip also established that grip arises from multivalent contacts between the bound substrate and multiple ClpX pore-1 loops. For example, when two alanines were in an otherwise polyglycine tract, positioning the alanines at any positions within the seven-

residue window synergistically enhanced grip, even when the alanines were spaced to contact different ClpX pore loops. Similar results were observed for substrates with two valine side chains. For substrates with multiple tyrosines, however, grip was enhanced more when the side chains were spaced in multiples of two residues, consistent with a binding state in which base-stacking networks form between the substrate and pore-1 loops. This effect extended beyond the seven-residue well-gripped window. This expansion of the grip region may indicate that favorable interactions with substrate can recruit additional subunits into a conformation that produces grip, even if this state would not otherwise predominate. Therefore, ClpX subunits not only adopt different functional roles depending on their global position, they may also be sufficiently functionally malleable to be recruited into different roles if it is energetically favorable.

Implications for mechanism(s) of substrate unfolding and translocation

One of the most significant unsolved problems in understanding the function of AAA+ proteases is describing how discreet power strokes are coupled together to generate processive substrate motion. Historically, sequential and stochastic models have been proposed to explain various structural and biochemical observations. The findings presented in this work challenge tenants of both models, and suggest that substrate translocation may in fact occur via a mechanism that contains elements of both models, or even through multiple distinct modes of action.

The spiral structure of ClpX is consistent with the idea that ClpXP unfolds substrates by a mechanism that has sequential components. If a power stroke occurs via a conformational transition that moves the lowest spiral subunit to the highest position or *vice versa*, then the spiral architecture predicts that the post-hydrolysis state of one power stroke is positioned as the pre-

hydrolysis state of the next power stroke. This provides a facile explanation for processive motor activity, as the constant adoption of a power stroke-competent state primes the motor to continue forward indefinitely. Some type of sequential mechanism is further supported by the observation that ClpX subunits communicate with directional bias, and may therefore be primed to promote activity in a specific neighboring subunit (Chapter II). However, the observation that ClpXP and other AAA+ motors retain activity when ATP hydrolysis is eliminated in individual subunits of the ring hexamer complicates a strictly sequential mechanism. Additionally, most published sequential models suggest that power strokes are driven by small two-residue translocation steps triggered by ATP hydrolysis and conformational changes in the second-lowest spiral subunit that pull the bound substrate down through the axial pore (Cooney et al., 2019; Deville et al., 2017; Dong et al., 2019; Gates et al., 2017; Han et al., 2017; Majumder et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017, 2019; Ripstein et al., 2017; Rizo et al., 2019; Twomey et al., 2019; Yu et al., 2018). This model is not easily reconciled with the observation that the uppermost three subunits of ClpX mediate most substrate grip during unfolding (Chapter III). One attractive possibility that reconciles these observations is that the action of individual power strokes is driven though sequential subunit action, allowing bursts of ATP hydrolysis in neighboring subunits that produce high levels of force and translocate the bound substrate long distances. Each burst of power strokes could initiate stochastically, obviating the need for ATP hydrolysis in any particular subunit.

The ClpX–ClpP crosslinking experiments presented in Chapter V indicate that inhibiting relative rotation of ClpX and ClpP selectively impairs degradation of folded versus unfolded substrates. Although this finding could be explained by a mechanism distinct from the sequential or stochastic

translocation model (i.e., a reciprocating model or a ratchet model), it raises the possibility that ClpX uses a different mechanism to unfold folded proteins than to translocate unfolded polypeptides. The sequential model describes a mechanistic mode in which low-level force is consistently applied to the substrate and all ATP-hydrolysis events are productively coupled to work, ideal for efficiently processing unfolded substrates. By contrast, the stochastic model posits kinetic bursts of ATP hydrolysis that quickly generate large amounts of force, which could be ideal for inducing unfolding of stubbornly folded domains. Employing distinct mechanisms of unfolding under different circumstances could allow ClpXP to perform well at all times, rapidly unfolding stable domains through stochastically driven bursts of activity and then relaxing into a sequential mode to maximize energetic efficiency during translocation and degradation.

Future directions for study of AAA+ motor mechanism

Much remains to be understood about the function of AAA+ proteases and protein-remodeling motors. All of the experiments presented in this thesis have been performed using *E. coli* ClpXP. It is unclear whether other AAA+ protein unfolding motors operate using similar mechanisms as ClpX, although most solved cryo-EM structures of AAA+ proteases share a similar spiral topology and substrate-binding conformation (Cooney et al., 2019; Deville et al., 2017; Ding et al., 2017; Dong et al., 2019; Gates et al., 2017; Gatsogiannis et al., 2019; Han et al., 2017; Lee et al., 2010; Lo et al., 2019; Majumder et al., 2019; Michalska et al., 2019; Monroe et al., 2017; de la Peña et al., 2018; Puchades et al., 2017, 2019; Ripstein et al., 2017; Rizo et al., 2019; Su et al., 2017; Sun et al., 2017; Twomey et al., 2019; White et al., 2018; Yokom et al., 2016; Yu et al., 2018; Zehr et al., 2017; Zhu et al., 2018). In particular, the experiments investigating substrate grip (Chapter III) are amenable to replication for any AAA+ protease, so long as the substrates contain an appropriate

degron. With careful design, it may also be possible to interrogate relative unfoldase–protease rotation (as described in Chapter V for ClpXP) for other genetically untethered proteases such as ClpAP, HslUV, PAN–20S, and the 26S proteasome.

The biochemical dissection of substrate grip by ClpXP described in Chapter III raises important questions about how the ClpX axial pore adopts a conformational state that produces maximal levels of grip. This question could be resolved by copurifying and using cryo-EM to determine the structure of ClpXP bound to a defined, low-complexity substrate modeled after the GFP substrates used previously to determine grip. A GFP substrate with a single tyrosine side chain at a single position within the gripped-residue window would manifest as a single density on an otherwise featureless polypeptide backbone. This structure would allow unambiguous correlation of ClpX axial-pore structure with biochemically characterized high-level grip. Structural comparison of pore-1 loops between this structure and those presented in Chapter IV may illuminate subtle conformational changes that facilitate substrate grip.

Prior to this work, the only reported examples of weak substrate grip were natural slippery sequences that mediate biologically important partial-processing reactions (Daskalogianni et al., 2008; Levitskaya et al., 1997; Lin and Ghosh, 1996; Vass and Chien, 2013). Although a handful of such proteins have been characterized, it is possible that many partial-processing substrates with slippery sequences remain unidentified. The experiments presented in Chapter III systematically determine the extent to which each natural amino acid contributes to grip strength, as well as how side chains work synergistically to promote strong grip. As such, it should be possible to systematically analyze all amino-acid sequences flanking folded domains in multi-domain proteins

to identify candidate poorly gripped sequences. These putative substrates could then be tested for partial processing to determine whether each putative slippery sequence facilitates a partial-processing reaction. Conveniently, as previous analyses of slippery sequences have demonstrated that similar trends hold for bacterial and eukaryotic AAA+ proteases, almost all domain-annotated proteomes are amenable to characterization by this approach (Hoyt et al., 2006; Kraut et al., 2012; Zhang and Coffino, 2004). In Appendix II, I present a prototype computational method that scans all sequences adjacent to structurally defined domains in two bacterial proteomes. More sophisticated computational approaches and rigorous analyses have potential to uncover novel biological functions of AAA+ proteases in partial substrate processing.

Finally, despite the constraints on the mechanism of processive substrate unfolding and translocation determined through characterization of crosslinked ClpX–ClpP complexes (Chapter V), describing the precise mechanical actions that drive force-generating power strokes remains a significant and unsolved problem. Modifying the described ClpX–ClpP crosslinking method to enforce discreet ClpX spiral topologies may be a powerful method to determine the functions of individual subunits during a power stroke. The cryo-EM structures presented in Chapter IV show several pairs of residues in ClpX and ClpP that are only in close proximity in the lowest ClpX subunit in the spiral, and crosslinking across this interface may allow conformational locking of a ClpX hexamer into one spiral topology that would not allow sequential movement through different spiral positions. Coupling this crosslinking framework to mutations that prevent ATP hydrolysis or inactivate individual RKH, pore-1, or pore-2 loops in topologically defined subunits could allow methodical parsing of the functions imposed on different subunits by the spiral topology during a power stroke. Though the challenge of studying intersubunit coordination within

the global architecture of AAA+ motors remains challenging, new structural and biochemical methods bring us ever closer to solving these important problems.

REFERENCES

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*, 502–505.

Cordova, J.C., Olivares, A.O., Shin, Y., Stinson, B.M., Calmat, S., Schmitz, K.R., Aubin-Tam, M.-E., Baker, T.A., Lang, M.J., and Sauer, R.T. (2014). Stochastic but Highly Coordinated Protein Unfolding and Translocation by the ClpXP Proteolytic Machine. Cell *158*, 647–658.

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*, 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*, e1701726.

Ding, Z., Fu, Z., Xu, C., Wang, Y., Wang, Y., Li, J., Kong, L., Chen, J., Li, N., Zhang, R., et al. (2017). High-resolution cryo-EM structure of the proteasome in complex with ADP-AlFx. Cell Res. *27*, 373–385.

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.

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*, 273–279.

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.

Glynn, S.E., Martin, A., Nager, A.R., Baker, T.A., and Sauer, R.T. (2009). Structures of Asymmetric ClpX Hexamers Reveal Nucleotide-Dependent Motions in a AAA+ Protein-Unfolding Machine. Cell *139*, 744–756.

Glynn, S.E., Nager, A.R., Baker, T.A., and Sauer, R.T. (2012). Dynamic and static components power unfolding in topologically closed rings of a AAA+ proteolytic machine. Nat. Struct. Mol. Biol. *19*, 616–622.

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.

Hersch, G.L., Burton, R.E., Bolon, D.N., Baker, T.A., and Sauer, R.T. (2005). Asymmetric Interactions of ATP with the AAA+ ClpX6 Unfoldase: Allosteric Control of a Protein Machine. Cell *121*, 1017–1027.

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*, 1720–1729.

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*, 1444–1453.

Lee, M.E., Baker, T.A., and Sauer, R.T. (2010). Control of Substrate Gating and Translocation into ClpP by Channel Residues and ClpX Binding. J. Mol. Biol. *399*, 707–718.

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. *94*, 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*, 2248–2254.

Lo, Y.-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*, 1–12.

Majumder, P., Rudack, T., Beck, F., Danev, R., Pfeifer, G., Nagy, I., and Baumeister, W. (2019). Cryo-EM structures of the archaeal PAN-proteasome reveal an around-the-ring ATPase cycle. Proc. Natl. Acad. Sci. *116*, 534–539.

Michalska, K., Zhang, K., March, Z.M., Hatzos-Skintges, C., Pintilie, G., Bigelow, L., Castellano, L.M., Miles, L.J., Jackrel, M.E., Chuang, E., et al. (2019). Structure of Calcarisporiella thermophila Hsp104 Disaggregase that Antagonizes Diverse Proteotoxic Misfolding Events. Structure *27*, 449-463.e7.

Monroe, N., Han, H., Shen, P.S., Sundquist, W.I., and Hill, C.P. (2017). Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. ELife *6*, e24487.

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*, eaav0725.

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.

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*, 1073-1085.e6.

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.

Rizo, A.N., Lin, J., 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*, 1–12.

Ryu, J.-K., Min, D., Rah, S.-H., Kim, S.J., Park, Y., Kim, H., Hyeon, C., Kim, H.M., Jahn, R., and Yoon, T.-Y. (2015). Spring-loaded unraveling of a single SNARE complex by NSF in one round of ATP turnover. Science *347*, 1485–1489.

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*, 636–646.

Stinson, B.M., Nager, A.R., Glynn, S.E., Schmitz, K.R., Baker, T.A., and Sauer, R.T. (2013). Nucleotide Binding and Conformational Switching in the Hexameric Ring of a AAA+ Machine. Cell *153*, 628–639.

Su, M., Guo, E.Z., Ding, X., Li, Y., Tarrasch, J.T., Brooks, C.L., Xu, Z., and Skiniotis, G. (2017). Mechanism of Vps4 hexamer function revealed by cryo-EM. Sci. Adv. *3*, e1700325.

Sun, S., Li, L., Yang, F., Wang, X., Fan, F., Yang, M., Chen, C., Li, X., Wang, H.-W., and Sui, S.-F. (2017). Cryo-EM structures of the ATP-bound Vps4 E233Q hexamer and its complex with Vta1 at near-atomic resolution. Nat. Commun. *8*, 1–13.

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*, eaax1033.

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

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.

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*, 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. *115*, E9560–E9569.

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*, 717–725.

Zhang, M., and Coffino, P. (2004). Repeat Sequence of Epstein-Barr Virus-encoded Nuclear Antigen 1 Protein Interrupts Proteasome Substrate Processing. J. Biol. Chem. *279*, 8635–8641.

Zhu, Y., Wang, W.L., Yu, D., Ouyang, Q., Lu, Y., and Mao, Y. (2018). Structural mechanism for nucleotide-driven remodeling of the AAA-ATPase unfoldase in the activated human 26S proteasome. Nat. Commun. *9*, 1–12.

Appendix I

Kinetic studies of substrate grip by ClpXP

ABSTRACT

AAA+ proteases unfold substrate proteins by coupling conformational changes driven by ATP hydrolysis to a pulling force at the substrate termini. For pulling force to be transmitted effectively, the protease must maintain grip on the substrate by stably associating with the substrate during individual power strokes. The propensity of the substrate to dissociate from the protease in the intervals between power strokes may also contribute to overall substrate grip, but the relevance of this mechanism is unclear. In this appendix, I describe a novel method to directly measure the dissociation of substrates pre-bound to the AAA+ protein unfolding motor ClpX. I find that the rate of substrate dissociation correlates with the rate of ATP hydrolysis. Although tail lengths greater than five residues do not substantially stabilize ClpX–substrate complexes, complex stability is sensitive to amino-acid type, with dissociation half-times varying up to 7-fold in a panel of low-complexity sequences. Small-residue sequences associate most stably with ClpX despite promoting poor grip during unfolding of stable domains. These findings indicate that substrate release rate is a relevant mechanistic component of ClpXP grip for many substrate sequences.

INTRODUCTION

To successfully degrade a native substrate, ClpXP must apply force to the unfolded polypeptide sequence that directly abuts the folded structure. Conformational changes driven by ATP hydrolysis in subunits of ClpX drive substrate unfolding attempts. However, not all attempts to grip and unfold a substrate are successful, as many ATP-hydrolysis events occur before a substrate unfolds, and slowly unfolded substrates can dissociate from the protease into solution (Kenniston et al., 2003, 2005). As a result, substrate grip during unfolding could arise from two separate factors: (1) the ability to transmit force to a substrate during a hydrolysis-driven conformational change, and (2) the ability to prevent substrate from dissociating during intervals between ATP-hydrolysis events.

Several studies examining degradation of substrates containing poorly gripped Gly-Ala repeat sequences (a partial-processing signal in the mammalian p105 protein) have suggested that the first of these two factors is the primary driver of grip (Lin and Ghosh, 1996). An experiment looking at degradation versus release of ligand-stabilized dihydrofolate reductase (DHFR) by the 26S proteasome concluded that the presence of a poorly gripped sequence decreased the rate of substrate degradation approximately 10-fold, while the rate of substrate release was essentially unchanged (Kraut et al., 2012). A similar (though not identical) result was observed for ClpXP, as a Gly-Ala repeat sequence adjacent to the stable titin I27 domain caused a 5-fold reduction in degradation rate and only a 3.5-fold increase in substrate release rate (Kraut, 2013; Too et al., 2013). However, each of these experiments was performed by measuring the disappearance of a degradation intermediate on SDS-PAGE gels, a method with limited sensitivity. Furthermore, as these experiments have only tested known poorly gripped sequences against complex control

sequences, a systematic understanding of how different natural amino-acid sequences impact substrate release rates is lacking.

In this appendix, I present a method to directly interrogate the kinetic stability of ClpX–substrate complexes. Using this assay, I interrogate the amino-acid sequence determinants of complex stability during active unfolding by ClpXP. These findings indicate that both the transmission of unfolding force to the substrate during unfolding (as explored in Chapter III) as well as the propensity of the substrate to maintain stable association with the ClpX axial pore during repeated unfolding attempts contribute significantly to substrate grip.

RESULTS AND DISCUSSION

Assay for ClpX–substrate dissociation

To interrogate the stability of ClpX–substrate complexes, I sought a strategy to directly measure dissociation of ClpX from substrates. I developed a fluorescence-based assay by labeling single-chain $ClpX^{\Delta N}$ hexamers (neither the N-terminal domain nor intersubunit tethers affect degradation of ssrA-tagged substrates) with a TAMRA fluorophore and labeling a panel of DHFR substrates with a Black Hole fluorescence quencher (Figure A1.1A, Martin et al., 2005). DHFR substrates all contained defined sequences directly C-terminal to the folded domain terminated with a Leu-Ala-Ala (LAA) motif, the minimal portion of the ssrA degron required for substrate recognition by ClpX (Flynn et al., 2001). Fluorescent $ClpX^{\Delta N}$ exhibited an approximately 20% decrease in fluorescence upon binding to substrate in the presence of ATP, and the degree of fluorescence quenching was hyperbolic with respect to substrate concentration, as expected for single-site binding (Figure A1.1B).

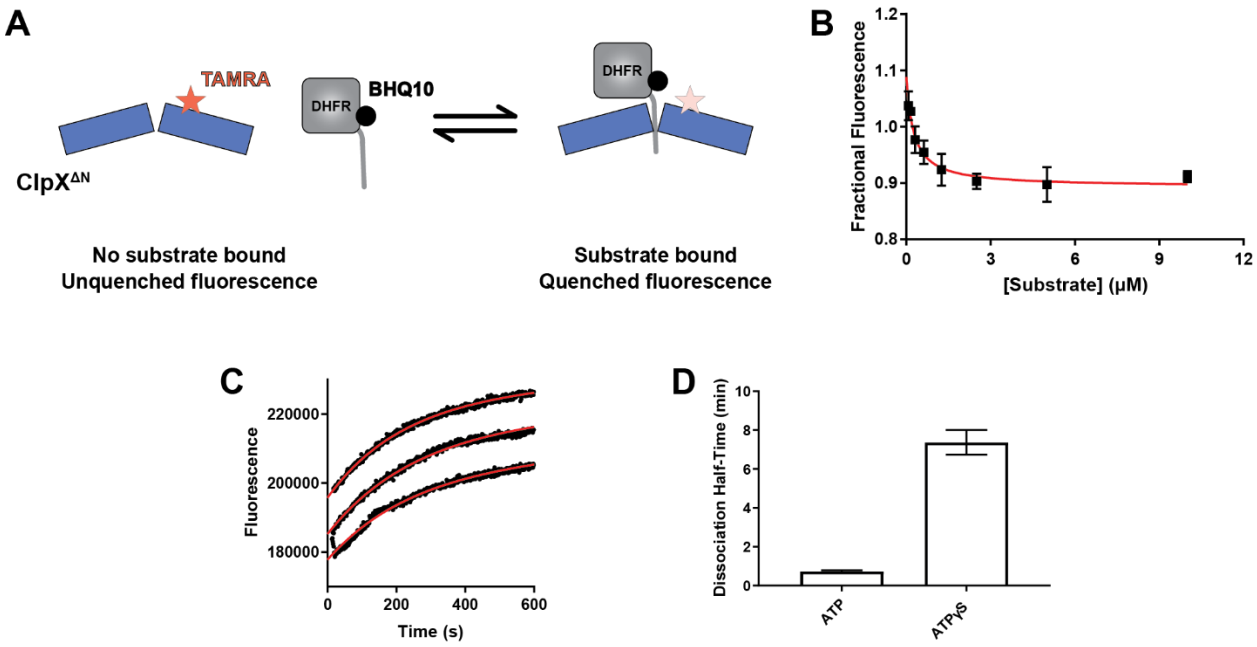


Figure A1.1 – Assay for ClpX–substrate complex stability

(A) Scheme for fluorescence quenching assay. **(B)** Fluorescence quenching is hyperbolic as a function of substrate concentration with an approximate K_{app} of 0.5 μM. **(C)** Kinetic transitions from quenched to unquenched state fit well to single exponential functions with a linear term. **(D)** ClpX–substrate complexes are substantially more stable in the presence of ATPγS than ATP.

Upon binding to a substrate and entering a partially quenched fluorescence state, quenching could be relieved either by substrate release or by unfolding and translocation of the substrate through the $ClpX^{\Delta N}$ axial pore. To isolate the kinetics of substrate release independent of unfolding, quencher-labeled DHFR substrates were pre-incubated with methotrexate (MTX), a small molecule that stabilizes DHFR structure and prevents unfolding by ClpX (Eilers and Schatz, 1986; Lee et al., 2001; Too et al., 2013). $ClpX^{\Delta N}$ and substrate were then incubated in the presence of ATP before dilution into a solution containing excess unlabeled DHFR substrate to prevent post-release reassociation of quencher-labeled substrates.

The fluorescence transition from a quenched to unquenched state fit to a single exponential function, indicating a single rate-limiting factor underlying substrate release (Figure A1.1C). The half-time of $ClpX^{\Delta N}$–substrate complexes increased approximately 7-fold when ATPγS was substituted for ATP, consistent with the slower rate of hydrolysis for ATPγS by $ClpX^{\Delta N}$ (Figure A1.1D, Burton et al., 2003). Thus, substrate release during unfolding depends on nucleotide hydrolysis rate.

Effect of substrate length on kinetic stability

Prior to unfolding an ssrA-tagged substrate, ClpX translocates forward from the C-terminal degron to the sequence abutting a folded domain. I hypothesized that the length of sequence traversed prior to unfolding could impact the rate of substrate release by ClpX because more polypeptide must diffuse back through the pore before the substrate is liberated. To investigate the impact of substrate C-terminal tail length, I designed a panel of DHFR substrates with different lengths of mixed hydrophobic and aromatic residues between the C-terminus of DHFR and the LAA degron,

with tails ranging from 5–21 amino acids in length (Figure A1.2A). Each of these substrates was purified, labeled with quencher, and dissociation kinetics in the presence of ATP were measured. All substrates exhibited single-exponential dissociation kinetics, and fits were used to determine half-times for dissociation of the $ClpX^{\Delta N}$–substrate complexes. Complex half-time did not vary between substrates with different C-terminal tail lengths, indicating that five tail residues are sufficient for $ClpX^{\Delta N}$ to maintain stable grip (Figure A1.2B). This observation converges with findings that the 5–6 residues abutting a GFP domain contribute most significantly to grip during substrate unfolding (Chapter III), and raises the possibility that a common mode of substrate interaction is required for $ClpX^{\Delta N}$ to transmit unfolding force during a power stroke and to maintain association with substrate in the intervals between power strokes.

Effect of substrate sequence on kinetic stability

I designed a second panel of DHFR substrates to interrogate how different types of amino acids contribute to $ClpX^{\Delta N}$–substrate complex stability. Each substrate contained a different 22-residue sequence tract between the C-terminus of DHFR and the LAA degron (Figure A1.3A). Fitting dissociation kinetics to single-exponential functions again facilitated calculation of complex half-times. Substrates with alternating glycine and alanine residues or an array of small amino acids maintained more stable association with $ClpX^{\Delta N}$, with an average complex half-time approximately 4-fold longer than the hydrophobic residue tracts tested previously (Figures A1.2B, A1.3B). Other sequence tracts of negatively and positively charged residues, a poly-histidine tract, and a tract of alternating proline and small residues all exhibited half-times similar to the hydrophobic-residue tracts. As small residues promote weaker grip than large residues during GFP unfolding (Chapter III), the finding that small-residue sequences maintain more stable association

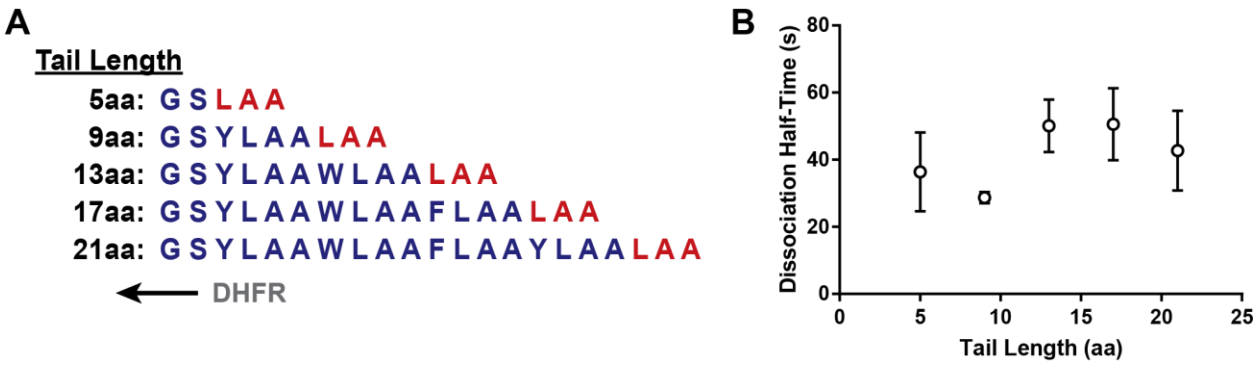


Figure A1.2 – Effects of substrate tail length on complex stability
(A) C-terminal tail sequence of substrates tested. **(B)** Half-times for single-exponential dissociation of ClpX–substrate complexes.

with ClpX$^{\Delta N}$ than better-unfolded sequences is surprising. It is possible that the increased conformational flexibility of small residues within the ClpX$^{\Delta N}$ pore or increased constriction of the pore around small residues increases the likelihood that they are re-bound following an unsuccessful unfolding attempt.

Mechanistic insights into grip by AAA+ proteases

Although previous studies found that grip defects related to slippery sequences inhibit ClpXP and the 26S proteasome more by slowing substrate unfolding than by promoting substrate release, this effect was far more pronounced for the proteasome. My finding here that different amino-acid sequences vary the kinetic stability of a ClpX$^{\Delta N}$–substrate complex by up to 7-fold further substantiates the model that substrate release is a significant factor in overall substrate grip. The 19S regulatory particle of the proteasome contains a number of ubiquitin- and substrate-binding factors in addition to a AAA+ motor, resulting in a highly multivalent substrate interaction network (Budenholzer et al., 2017; de la Peña et al., 2018). These accessory substrate-binding partners may reduce backsliding when substrate is transiently released by the Rpt1-6 motor, decreasing the relevance of substrate release for proteasomal grip. It is important to note that the experiments described above were performed using the truncated ClpX$^{\Delta N}$ variant. In the bacterial cytoplasm, the ClpX N-domains interact with adapter proteins such as SspB, which delivers ssrA-tagged substrates for degradation and substantially decreases the K_M for substrate degradation by ClpXP (Levchenko et al., 2000). It is therefore possible that when degrading a substrate in partnership with SspB, ClpXP maintains a higher-valency substrate interaction that minimizes grip destabilization through substrate backsliding.

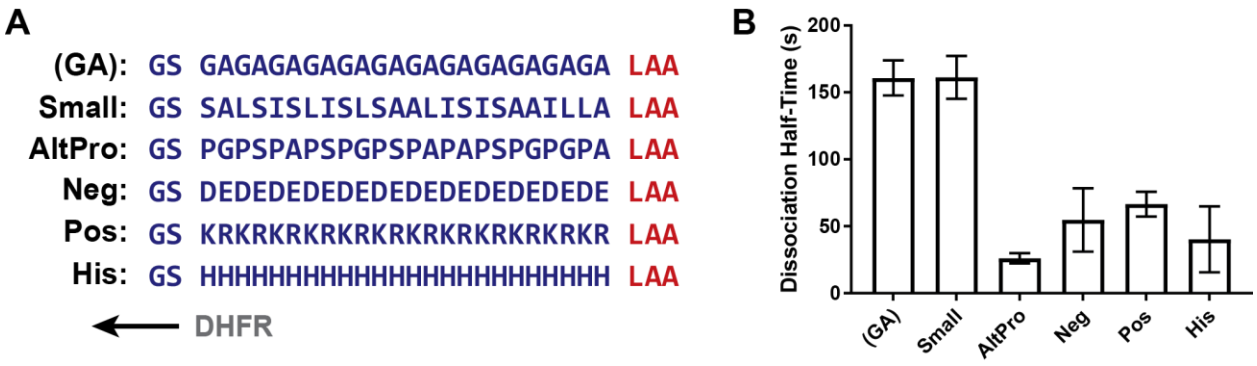


Figure A1.3 – Effects of substrate tail amino acid sequence on complex stability
(A) C-terminal tail sequence of substrates tested. **(B)** Half-times for single-exponential dissociation of ClpX–substrate complexes.

Although substrate release is likely an important factor in ClpX–substrate grip strength under many circumstances, it is notable that different amino acid types are differentially sensitive to being released. Previous kinetic analyses of substrate release for both ClpXP and the 26S proteasome were performed using known poorly gripped sequence motifs enriched in glycine and other small amino acids. My results indicate that these sequence motifs in fact form the most kinetically stable complexes with $ClpX^{\Delta N}$. This effect could arise because power strokes are highly ineffective or because increased conformational flexibility increases the likelihood of substrate rebinding after a transient failure of grip. Looking forward, it will be interesting to determine whether similar amino-acid specific trends identified here for ClpXP hold for the proteasome as well, as they may indicate increased relevance of substrate release in tuning grip strength during unfolding of many cellular substrates.

METHODS

Preparation of BHQ10-maleimide labeling reagent

Black Hole Quencher 10 (BHQ10) was purchased as a succinimide ester (SE) derivative (Biosearch Technologies). To convert the quencher to a maleimide for cysteine labeling, BHQ10-SE was resuspended in DMSO to a concentration of 25 mM. The solubilized quencher was mixed with 0.9 volumes of 25 mM *N*-(2-aminoethyl)maleimide and 0.1 volumes of DMSO. Anhydrous triethylamine was then added to a final concentration of 25 mM, and the mixture was allowed to

react for 1 h at ambient temperature, protected from light. The maleimide-derivatized product of this reaction was stored in small aliquots at –20 °C until used for labeling.

Protein purification and labeling

Mutations were introduced into one subunit of a $ClpX^{\Delta N}$ single-chain hexamer (D170C) and *E. coli* DHFR (D132C) by PCR mutagenesis. DHFR D132C variable-tail substrates were generated via PCR mutagenesis by encoding the tail sequences onto oligonucleotides. $ClpX^{\Delta N}$ and substrates were purified as described previously (Martin et al., 2005). Prior to labeling, proteins were desalted over G25 resin into buffer lacking DTT. TAMRA-maleimide ($ClpX^{\Delta N}$) or BHQ10-maleimide (DHFR) was added at a 3:1 ratio. Labeling reactions were carried out for 2 h at ambient temperature. The $ClpX^{\Delta N}$ labeling reaction was quenched by addition of 1 mM DTT and desalted over G25 resin to remove free label. DHFR labeling reactions were desalted over G25 resin without quenching. Labeled and unlabeled DHFR were then resolved by Source 15Q ion exchange chromatography, and the labeled peak was used for kinetic assays.

Substrate dissociation measurement and analysis

TAMRA-$ClpX^{\Delta N}$ (0.2 μM hexamer) and BHQ10-DHFR-MTX substrates (5 μM monomer) were mixed in PD buffer (25 mM HEPES-KOH pH 7.5, 200 mM KCl, 5 mM $MgCl_2$, 10% glycerol) in the presence of 5 mM nucleotide in a total volume of 20 μL. For reactions containing ATP, a regeneration system (32 mM phosphocreatine and 0.08 mg/mL creatine kinase) was also included. Reactions were incubated for 5 min at ambient temperature to allow complex formation. At the start of measurement, 20 μL of a competitor mixture containing 500 μM unlabeled DHFR-MTX-ssrA prepared in the same buffer and with appropriate nucleotide was added and quickly mixed by

pipetting in a 1.5 mm quartz cuvette. TAMRA fluorescence (excitation 530 nm, emission 578 nm) was measured once per s for 60 min or until fluorescence signal saturated on a fluorometer with 2 mm slit width and recorded using FeliX32 software (PTI). Fluorescence levels were fit to single exponentials with an additional linear term included to correct for instrument signal drift using Prism 7 software (GraphPad). Complex half-times were calculated from the exponential fits as $t_{1/2} = \ln(2) / k$.

Substrate–ClpX equilibrium binding was measured using 25 nM TAMRA-ClpX$^{\Delta N}$ and 80–20,000 nM BHQ10-DHFR substrate. After a 2 min incubation with or without 4 mM ATP and regeneration mix at ambient temperature, a fluorescence emission spectrum (excitation 530 nm, emission 560–590 nm) was recorded using an M5 plate reader (SpectraMax). The emission spectrum was fit to a polynomial function using a custom Python script and the area under the curve (AUC) from 560–590 nm was integrated. Quenching was calculated as:

$$Quenching = 1 - \left(\frac{AUC\ with\ ATP}{AUC\ without\ ATP}\right)$$

ACKNOWLEDGMENTS

I am grateful to my former labmates Ben Stinson and Reuben Saunders, who identified a suitable fluorophore labeling position on ClpX$^{\Delta N}$ and optimized purification and labeling protocols.

REFERENCES

Budenholzer, L., Cheng, C.L., Li, Y., and Hochstrasser, M. (2017). Proteasome Structure and Assembly. J. Mol. Biol. *429*, 3500–3524.

Burton, R.E., Baker, T.A., and Sauer, R.T. (2003). Energy-dependent degradation: Linkage

between ClpX-catalyzed nucleotide hydrolysis and protein-substrate processing. Protein Sci. *12*, 893–902.

Eilers, M., and Schatz, G. (1986). Binding of a specific ligand inhibits import of a purified precursor protein into mitochondria. Nature *322*, 228–232.

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*, 10584–10589.

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.

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. *102*, 1390–1395.

Kraut, D.A. (2013). Slippery Substrates Impair ATP-dependent Protease Function by Slowing Unfolding. J. Biol. Chem. *288*, 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*, 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*, 627–637.

Levchenko, I., Seidel, M., Sauer, R.T., and Baker, T.A. (2000). A Specificity-Enhancing Factor for the ClpXP Degradation Machine. Science *289*, 2354–2356.

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*, 2248–2254.

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

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*, eaav0725.

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*, 13243–13257.

Appendix II

Computational prediction of partial-processing signals for AAA+ proteases

ABSTRACT

During unfolding of a substrate domain, AAA+ proteases grip the unfolded polypeptide sequence that directly neighbors the domain being unfolded. The systematic characterization of substrate grip by ClpXP presented in Chapter III allows prediction of the maximum grip that can be applied to any protein domain based only on the amino acids present in the adjoining sequence. Here, I describe a high-throughput computational method to determine the gripped amino-acid sequence adjacent to structurally characterized domains in *E. coli* and *B. subtilis*. I use these sequences to predict how well each can be gripped by ClpXP and present protein domains that may be poorly gripped. This collection of sequences likely includes those that induce slipping and prevent processive degradation, causing release of partially processed products. I also discuss potential weaknesses of this high-throughput discovery approach and provide suggestions for how the method could be modified to expand the number of proteins analyzed and improve its accuracy in predicting partial-processing signals.

INTRODUCTION

AAA+ proteases apply grip to their substrates during unfolding of stably folded domains, and the identities of the amino acids directly adjacent to the folded domain determine the extent of grip that the motor can produce (Siddiqui et al., 2004; Chapter III). In cases where the level of grip is insufficient relative to the stability of the folded substrate domain, unfolding and degradation stall, leading to release of a partially processed substrate fragment (Kenniston et al., 2005). Several natural examples of low-grip sequences have been identified in eukaryotes, bacteria, and viruses, suggesting a conserved mechanism of partial substrate processing by AAA+ proteases (Daskalogianni et al., 2008; Levitskaya et al., 1997; Lin and Ghosh, 1996; Vass and Chien, 2013).

In Chapter III, I characterized the sequence determinants of grip by ClpXP, a bacterial AAA+ protease. This work revealed that the first six residues directly adjacent to a folded domain are most important for grip, and showed that residues with small side chains are gripped poorly. This finding is in good agreement with characterization of naturally occurring glycine- and alanine-rich poorly gripped sequences (Hoyt et al., 2006; Kraut, 2013; Kraut et al., 2012; Sharipo et al., 2001; Tian et al., 2005; Too et al., 2013; Zhang and Coffino, 2004). ClpX grips polar and charged residues more weakly than hydrophobic side chains of similar size. Additionally, I established that multiple substrate side chains of a substrate can synergistically enhance grip strength.

With this information, it is theoretically possible to score any amino-acid sequence adjacent to a folded protein domain for its propensity to induce slipping and partial processing by a AAA+ protease. Here, I describe a computational method that predicts poorly gripped partial-processing signatures and discuss its performance in analyzing two bacterial proteomes.

RESULTS AND DISCUSSION

Identification of amino acid sequences that directly abut folded domains

To ensure that my analysis of gripped sequences only included amino acid stretches directly adjacent to folded domains, I restricted the analysis to proteins with reported structures in the Protein Data Bank (PDB). Because over 95% of structures in the PDB are solved at a reported resolution of less than 4 Å, this filtering step ensured that analysis of the residues directly flanking the terminus of the folded domain (gripped-sequence windows) would not be more than one amino acid (~3.5 Å length) offset from the terminus (Berman et al., 2000). The CATH Database is an annotation of all folded domains within published PDB structures (Dawson et al., 2017). I aligned the sequences contained within PDB structures against a reference proteome to identify the gripped-sequence windows on the N- and C-termini of each domain.

Structures of proteins within the PDB are not necessarily identical to the sequences of their native counterparts. Structurally characterized proteins are often truncated, mutated, or fused to other sequences to facilitate purification and structure determination. Furthermore, flexible regions may be left unbuilt in structural models and therefore excluded from the PDB file. These discrepancies make alignment of PDB and reference proteome sequences nontrivial. To address this challenge, sequences taken from PDB files were first truncated to individual domains using domain boundary annotations from the CATH database. The domain sequences were then permissively aligned onto an appropriate reference proteome with tolerance for mutations, gaps, and insertions in sequence. Sequences that could not be unambiguously aligned were excluded from downstream analysis. In cases where unambiguous alignment to a reference protein sequence was possible but gaps or mismatches between the PDB and reference sequences were present within ten residues of the

CATH-annotated domain boundary, the terminal residues of the PDB domain sequence were threaded across the reference sequence to obtain the best possible alignment. If the terminal alignment was ambiguous, the domain was excluded from analysis. Upon successful alignment of a domain with a reference-protein sequence, the six amino acids from the reference sequence directly N-terminal or C-terminal to the boundaries of the PDB domain (i.e., the gripped-residue window at either terminus) were recorded and stored. ClpXP can degrade substrates from the N- or C-terminus, so both sets of gripped-residue windows were analyzed (Olivares et al., 2017).

Scoring sequences by predicted grip strength

Gripped-residue windows six amino acids in length were then assigned scores using the scoring matrix shown in Figure A2.1. This scoring function groups residues by how well they support unfolding and degradation of GFP-ssrA and awards higher scores to better-gripped residues (Chapter III). In addition, to account for synergy in grip, additional points were awarded to residues with multiple well gripped or poorly gripped amino acids.

Figure A2.2 shows histograms of grip score for N- and C-terminal gripped-residue windows for domains that aligned unambiguously to the *E. coli* and *B. subtilis* reference proteomes. Although gripped-residue windows exhibited a wide range of scores on a scale from 0–16, a small number of sequences with scores less than 1 were present in each proteome. These sequences, which are shown in Table A2.1, are candidate poorly gripped sequences.

Condition	Score
Weakly-gripped residues Asp, Asn, Gly, Pro, Ser, Thr	+0 points
Modestly-gripped residues Ala, Arg, Cys, Glu, Gln, His, Lys,Trp	+0.5 points
Well-gripped residues Ile, Leu, Met, Phe, Tyr, Val	+1.5 points
2 well-gripped	+3 points
3+ well-gripped	+6 points
2 modestly-gripped	+1 point
3+ modestly-gripped	+2 points
Modestly-gripped + well-gripped	+2 points

Figure A2.1 – Grip scoring matrix

Scoring rules used to calculate grip score for gripped-sequence windows of six residues. Each sequence window receives a score ranging from 0 (gripped poorly) to 16 (gripped well).

Caveats in interpreting grip scores

Not all domains identified in this analysis are necessarily gripped poorly by AAA+ proteases *in vivo*, and some poorly gripped sequences likely evade detection by this method. One major caveat relates to accurate determination of domain boundaries. Folded proteins are flexible macromolecules that sample a range of subtly different conformational states at cellular temperatures (Fraser et al., 2011). As a result, sequences that appear to flank a folded domain in published structures are not necessarily the sequence gripped during unfolding – if a terminal structural element adopts a structured fold under structure determination conditions but normally is in equilibrium with an unfolded state, it may in fact constitute the gripped sequence. Because the worst-gripped sequences tend to be enriched in glycine (which has low propensity to adopt ordered secondary structures), this caveat may be ameliorated to some degree.

Another weakness of this method is that it assumes that all amino acids within the gripped sequence window contribute equally to grip during unfolding. In fact, grip propensity is roughly normally distributed with respect to residue distance from the domain boundary (Chapter III). A more nuanced algorithm would apply a Gaussian smoothing function during calculation of grip score to better simulate the different extent to which substrate residues determine grip strength.

Finally, regardless of the accuracy of grip prediction, the impact of amino-acid identity in the gripped-residue window on domain unfolding also depends heavily on the stability of the domain being unfolded. More grip is required to unfold and degrade more stably folded domains (Kenniston et al., 2003). At the opposite extreme, ClpXP quickly degrades unfolded polypeptides with little sequence preference (Barkow et al., 2009). A predicted poorly gripped sequence may

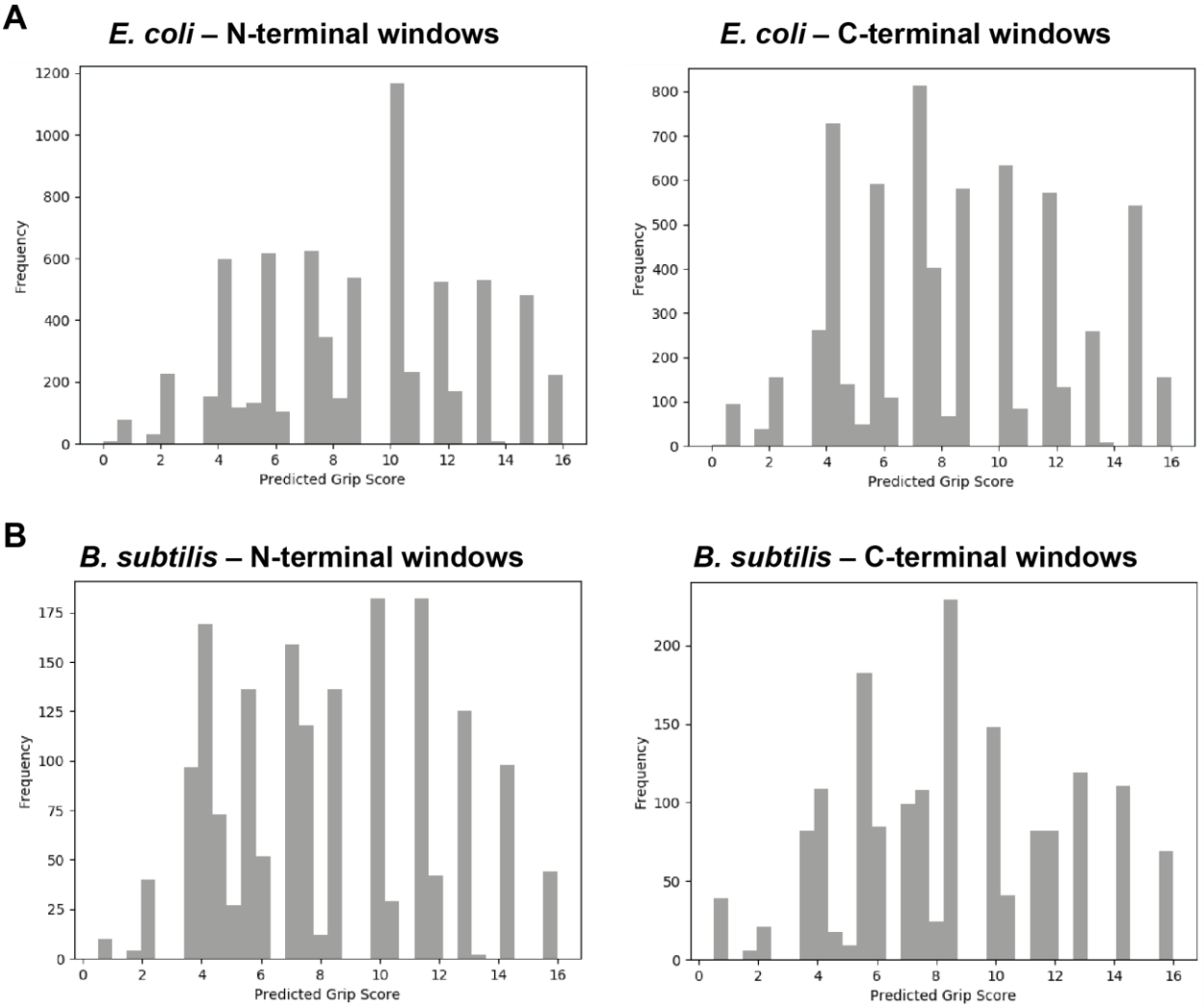


Figure A2.2 – Grip score histograms among bacterial proteins with reported structures
Distribution of grip scores for N- and C-terminal gripped sequence windows in **(A)** *E. coli* and **(B)** *B. subtilis*.

have no biological phenotype if it sits adjacent to a metastable folded domain that does not require high-level force application for unfolding. As such, the list of sequences in Table A2.1 is likely to contain many false positives (i.e., proteins which can be unfolded and degraded by AAA+ proteases).

Future directions to improve and expand computational grip prediction

The putative poorly gripped sequences identified in this analysis could be biologically interesting if they represent uncharacterized partial-processing signals for AAA+ proteases. Unfortunately, because the final output is likely to contain many false positives, empirical validation will be necessary to identify any new biological functions. There is, however, potential to reduce the model's false positive rate. For example, domains that contain elements such as β-barrels and large β-sheets tend to be difficult for AAA+ proteases to mechanically degrade (Kenniston et al., 2003; Li et al., 2000; Martin et al., 2008). Correlation of grip score from this method with the predicted stability of the adjacent domain could limit empirical analysis to a subset of proteins that are difficult to both grip and unfold.

As presented, this method only analyzes grip for substrates with published structures. However, the analysis could be extended to most proteins if reference sequences were threaded onto existing structures to predict domain boundaries *de novo*. Although the domain boundaries of these structures would undoubtedly be less precise than those for determined structures, using a rolling window across the predicted domain boundaries to determine grip scores could minimize this problem.

Organism	Terminus	Protein (PDB ID)	Domain(s) (PDB Chain)	Gripped Sequence Window	Grip Score
E. coli	N	1XXB	A	PTTSSP	0.0
E. coli	N	1XXA	A	PTTSSP	0.0
E. coli	N	3CSU	A	GDGSNQ	0.5
E. coli	N	4UHT	A	QNNDNG	0.5
E. coli	N	3MBT	A	SPTPPR	0.5
E. coli	N	2FPO	A	PNHSGS	0.5
E. coli	N	3A2Z	A	SKGTTS	0.5
E. coli	N	3NPM	A	GDGSNQ	0.5
E. coli	N	3RGN	A	QDTSPD	0.5
E. coli	N	4LRX	C	NNDGRG	0.5
E. coli	N	3RGM	A	QDTSPD	0.5
E. coli	N	1RAB	A	GDGSNQ	0.5
E. coli	N	1RAC	A	GDGSNQ	0.5
E. coli	N	3QDR	A	GSGNTK	0.5
E. coli	N	1RAA	A	GDGSNQ	0.5
E. coli	N	1RAF	A	GDGSNQ	0.5
E. coli	N	3A30	A	SKGTTS	0.5
E. coli	N	1RAD	A	GDGSNQ	0.5
E. coli	N	1RAE	A	GDGSNQ	0.5
E. coli	N	4F04	A	GDGSNQ	0.5
E. coli	N	1RAI	A	GDGSNQ	0.5
E. coli	N	3098	A	GTTSQD	0.5
E. coli	N	1GQ3	A	GDGSNQ	0.5
E. coli	N	3DMY	A	GSGSSQ	0.5
E. coli	N	3MPU	A	GDGSNQ	0.5
E. coli	N	3D7S	A	GDGSNQ	0.5
E. coli	N	4P6I	D	SGQPGG	0.5
E. coli	N	1A3A	A	DDSSAN	0.5
E. coli	N	3NKE	A	SGQPGG	0.5
E. coli	N	4V0B	A	PSESNG	0.5
E. coli	N	1RAG	A	GDGSNQ	0.5
E. coli	N	4KX6	A	TAGNGN	0.5
E. coli	N	4E2F	A	GDGSNQ	0.5
E. coli	N	4FYW	A	GDGSNQ	0.5
E. coli	N	4FYV	A	GDGSNQ	0.5
E. coli	N	4FYY	A	GDGSNQ	0.5
E. coli	N	4FYX	A	GDGSNQ	0.5
E. coli	N	1RAH	A	GDGSNQ	0.5
E. coli	N	5DQZ	C	SGQPGG	0.5
E. coli	C	1AA3	A	PNSTPD	0.0
E. coli	C	3Q8D	A	PGSDGP	0.0
E. coli	C	3JZF	A	PGSDGP	0.0

E. coli	C	3W7S	A	NPDATP	0.5
E. coli	C	3W7W	A	NPDATP	0.5
E. coli	C	4XMZ	A	TDTATD	0.5
E. coli	C	3W7U	A	NPDATP	0.5
E. coli	C	3W7T	A	NPDATP	0.5
E. coli	C	5CA3	A	NPDATP	0.5
E. coli	C	4XN2	A	TDTATD	0.5
E. coli	C	3W7X	A	NPDATP	0.5
E. coli	C	5CA4	A	NPDATP	0.5
E. coli	C	2HPT	A	TDTATD	0.5
E. coli	C	4PYD	E	GGKSGD	0.5
E. coli	C	3PUU	A	TDTATD	0.5
E. coli	C	4HQ0	A	DDSPQP	0.5
E. coli	C	4XNB	A	TDTATD	0.5
E. coli	C	4XMW	A	TDTATD	0.5
E. coli	C	4XNA	A	TDTATD	0.5
E. coli	C	3I6P	A	KGDSSN	0.5
E. coli	C	4XND	A	TDTATD	0.5
E. coli	C	4XN5	A	TDTATD	0.5
E. coli	C	4XN1	A	TDTATD	0.5
E. coli	C	4IQZ	A	SGGTKD	0.5
E. coli	C	4Q4E	A	TDTATD	0.5
E. coli	C	4Q4I	A	TDTATD	0.5
E. coli	C	4XO3	A	TDTATD	0.5
E. coli	C	2HPO	A	TDTATD	0.5
E. coli	C	3B3B	A	TDTATD	0.5
E. coli	C	3B37	A	TDTATD	0.5
E. coli	C	4XN4	A	TDTATD	0.5
E. coli	C	3QHS	A	SNNAGG	0.5
E. coli	C	3QJX	A	TDTATD	0.5
E. coli	C	3KED	A	TDTATD	0.5
E. coli	C	4XO5	A	TDTATD	0.5
E. coli	C	4V2S	A	SNNAGG	0.5
E. coli	C	2Y5Y	A	TDAPSS	0.5
E. coli	C	4ZYR	B	TDAPSS	0.5
E. coli	C	4KNY	B	TTAPDP	0.5
E. coli	C	4RCB	A	SNNAGG	0.5
E. coli	C	2ZXG	A	TDTATD	0.5
E. coli	C	3V8V	B	TPDSKP	0.5
E. coli	C	5GW7	A	NPDATP	0.5
E. coli	C	3V97	B	TPDSKP	0.5
E. coli	C	3STJ	A	TSSSAS	0.5
E. coli	C	4XMV	A	TDTATD	0.5
E. coli	C	4XMU	A	TDTATD	0.5
E. coli	C	4XMT	A	TDTATD	0.5

E. coli	C	4XN7	A	TDTATD	0.5
E. coli	C	3B34	A	TDTATD	0.5
E. coli	C	3WYJ	B	GTEPGD	0.5
E. coli	C	4XO4	A	TDTATD	0.5
E. coli	C	3D3I	A	NPDATP	0.5
E. coli	C	4XN8	A	TDTATD	0.5
E. coli	C	4XN9	A	TDTATD	0.5
E. coli	C	3B2P	A	TDTATD	0.5
E. coli	C	4XMX	A	TDTATD	0.5
B. subtilis	N	3PPR	A	GGASDD	0.5
B. subtilis	N	1OR4	A	SDSNGQ	0.5
B. subtilis	N	3PPQ	B	GGASDD	0.5
B. subtilis	N	2Z3X	A	SRSNNN	0.5
B. subtilis	N	2IEE	B	DSKDTG	0.5
B. subtilis	N	2R11	B	SNHSSS	0.5
B. subtilis	C	3HMA	A	GGGSQT	0.5
B. subtilis	C	1SBR	B	GDTQGD	0.5
B. subtilis	C	2Y1R	A	GSNETG	0.5
B. subtilis	C	1S7H	A	GDTQGD	0.5
B. subtilis	C	2RCV	A	PNGGGE	0.5
B. subtilis	C	1SBR	A	GDTQGD	0.5
B. subtilis	C	1TYG	C	TASSPG	0.5
B. subtilis	C	1RZN	A	SPSSGA	0.5
B. subtilis	C	3PXG	A	GSNETG	0.5
B. subtilis	C	3HMB	A	GGGSQT	0.5
B. subtilis	C	2Y1R	A	GSNETG	0.5
B. subtilis	C	1S99	A	GDTQGD	0.5
B. subtilis	C	3RDR	A	GGGSQT	0.5

Table A2.1 – Predicted poorly gripped sequences

Protein domains predicted to be poorly gripped by ClpXP during unfolding. This table does not include duplicate domains listed under a single PDB ID, but does include duplicates arising from multiple published structures of the same protein.

Finally, because poorly gripped sequences for ClpXP and the eukaryotic 26S proteasome are strikingly similar, this grip-scoring algorithm may be extensible to eukaryotic proteomes (Hoyt et al., 2006; Kraut et al., 2012; Zhang and Coffino, 2004). One significant roadblock to this advance is that the gripped-residue window for the 26S proteasome has not been well determined. Preliminary experiments that placed a poorly gripped sequence motif at different distances relative to a folded domain suggested that there is greater spacing between the folded substrate domain and the gripped residue window during degradation by the 26S proteasome than by ClpXP (Hoyt et al., 2006). As such, the gripped-residue window will need to be better defined before a computational search could be meaningfully performed.

METHODS

Input files and protein sequences

PDB sequences and domain annotations were obtained from the CATH Database (version 4.2.0). Reference proteomes for *E. coli* (taxonomy ID 83333) and *B. subtilis* (taxonomy ID 1423) were obtained from Uniprot (reference identifications UP000000625 and UP000001570, respectively).

Identification and scoring of domain-adjacent gripped sequences

PDB domains were aligned to appropriate reference proteomes as described above using a custom Python script that included sequence manipulation functions from the Biopython package (Cock et al., 2009). After extraction of gripped-residue windows, sequences were passed to a subsequent custom Python script for scoring according to the rules described above. Grip score histograms were generated using the Python package Matplotlib (Hunter, 2007).

ACKNOWLEDGMENTS

I am grateful to S.B. Hari, J.P. Morehouse, and X. Fei for helpful conversations during the development of this method, and to the CATH Database project for making their well-maintained annotations of PDB domains freely accessible.

REFERENCES

Barkow, S.R., Levchenko, I., Baker, T.A., and Sauer, R.T. (2009). Polypeptide Translocation by the AAA+ ClpXP Protease Machine. Chem. Biol. *16*, 605–612.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423.

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*, 30090–30100.

Dawson, N.L., Lewis, T.E., Das, S., Lees, J.G., Lee, D., Ashford, P., Orengo, C.A., and Sillitoe, I. (2017). CATH: an expanded resource to predict protein function through structure and sequence. Nucleic Acids Res. *45*, D289–D295.

Fraser, J.S., Bedem, H. van den, Samelson, A.J., Lang, P.T., Holton, J.M., Echols, N., and Alber, T. (2011). Accessing protein conformational ensembles using room-temperature X-ray

crystallography. Proc. Natl. Acad. Sci. *108*, 16247–16252.

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*, 1720–1729.

Hunter, J.D. (2007). Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. *9*, 90–95.

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.

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. *102*, 1390–1395.

Kraut, D.A. (2013). Slippery Substrates Impair ATP-dependent Protease Function by Slowing Unfolding. J. Biol. Chem. *288*, 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*, 1444–1453.

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. *94*, 12616–12621.

Li, H., Carrion-Vazquez, M., Oberhauser, A.F., Marszalek, P.E., and Fernandez, J.M. (2000). Point mutations alter the mechanical stability of immunoglobulin modules. Nat. Struct. Biol. *7*,

1117–1120.

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*, 2248–2254.

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*, 139–145.

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.

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*, 137–142.

Siddiqui, S.M., Sauer, R.T., and Baker, T.A. (2004). Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. Genes Dev. *18*, 369–374.

Tian, L., Holmgren, R.A., and Matouschek, A. (2005). A conserved processing mechanism regulates the activity of transcription factors Cubitus interruptus and NF-κB. Nat. Struct. Mol. Biol. *12*, 1045–1053.

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*, 13243–13257.

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

Zhang, M., and Coffino, P. (2004). Repeat Sequence of Epstein-Barr Virus-encoded Nuclear Antigen 1 Protein Interrupts Proteasome Substrate Processing. J. Biol. Chem. *279*, 8635–8641.