The Role of the N Domain in Substrate Binding, Oligomerization, and Allosteric Regulation of the AAA+ Lon Protease

By

Matthew L. Wohlever

B.S. Biochemistry
The Ohio State University, 2008

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

DOCTOR OF PHILOSOPHY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

Signature of Author: _____
Department of Biology
May 14, 2013

Certified by: _____
Robert T. Sauer
Salvador E. Luria Professor Biology
Thesis Supervisor

Accepted by: _____
Amy Keating
Associate Professor of Biology
Co-Chair, Graduate Committee

The Role of the N Domain in Substrate Binding, Oligomerization, and Allosteric Regulation of the AAA+ Lon Protease

By

Matthew L. Wohlever

Submitted to the Depart of Biology on May 24, 2013 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biochemistry

ABSTRACT

For cells and organisms to survive, they must maintain protein homeostasis in varied and often harsh environments. Cells utilize proteases and chaperones to maintain their proteomes. In bacteria, most cytosolic proteolysis is performed by self-compartmentalized AAA+ proteases, which convert the chemical energy of ATP binding and hydrolysis into mechanical work to unfold and translocate substrates into an internal degradation chamber. Substrates are targeted to AAA+ proteases by degradation tags (degrons). In *E. coli*, the Lon protease is responsible for the degradation of numerous regulatory proteins, including the cell-division inhibitor SulA, but also recognizes and degrades the majority of misfolded proteins. How Lon recognizes and prioritizes such a vast array of substrates is poorly understood. Active Lon is a homohexamer in which each subunit contains an N domain, a AAA+ module that mediates ATP binding and hydrolysis, and a peptidase domain. Degron binding allosterically regulates Lon activity and can shift Lon into conformations with higher or lower protease activity, but the mechanistic basis of this regulation is unknown. The low-protease conformation of Lon may serve as a chaperone.

In Chapter 2, I describe the development and characterization of fluorescent model substrates that Lon degrades *in vitro* and *in vivo*. In Chapter 3, I describe collaborative experiments that show that Lon equilibrates between a hexamer and a dodecamer. Based on biochemical analysis and a low-resolution EM dodecamer structure, Lon appears to shift its substrate profile by changing oligomeric states and contacts between N domains appear to stabilize the dodecamer.

In Chapters 4 and 5, I identify a binding site for the sul20 degron (isolated from SulA) in the Lon N domain and demonstrate that substrate binding to this site allosterically regulates protease and ATPase activity. I also show that the E240K mutation in the N domain alters Lon activity and stabilizes dodecamers. Finally, I provide evidence that *E. coli* Lon can act as a chaperone *in vivo*. These experiments demonstrate that the N domain integrates substrate binding, oligomerization, and regulation of the catalytic activities of Lon.

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

ACKNOWLEDGEMENTS

This thesis would not be possible without the hard work and support of many people. First, I would like to thank my advisors Bob and Tania for their patient and supportive mentoring that has allowed me to grow into the scientist I am today.

I also want to thank all members of the Sauer and Baker labs, both past and present, for the countless hours spent discussing science and teaching me new techniques. A special thank you is owed to Sarah Bissonnette and Eyal Gur, who got me started working on Lon and laid the foundation for my thesis. Once I began my studies on Lon, Ellen Vieux, James Chen, and Andrew Nager were invaluable collaborators. Finally, a big thank you is owed to Joey Davis, Shankar Sundar, Vlad Baytshtok, and Dominik Barthelme who have worked in room 595 over the years and provided most of my hands-on training in the lab along with great company and conversation.

Thank you to all of my biology classmates, friends at Tech Catholic Community, IM sports teammates, running partners, and friends from Ohio State for providing a welcomed break from lab, helping me to keep my priorities straight, and making the last five years fun. Thank you for providing the great stories behind my multiple shoulder injuries, accompanying me to the ER, and helping with my recovery.

I owe my biggest thank you to my family, without your constant love and support, I never would have made it into MIT, much less finished my thesis. I especially want to thank my wife, Theresa, for putting up with all of the late dinners and busy weekends that come with graduate school. Thank you!

TABLE OF CONTENTS

LIST OF FIGURES & TABLES

Chapter 1

Chaperones, Proteases, and Proteotoxic Stress: The Role of Lon in Protein Homeostasis

Protein Quality Control: The Role of Chaperones

Proteostasis

Proteins are essential for almost every cellular process, from cell division to apoptosis. Even the relatively simple bacteria, *Escherichia coli,* contains over 4000 different proteins and over 1 million total copies of these different proteins in a volume of $\sim 10^{-15}$ L (Moran et al. 2010). In this crowded environment, the exposure of hydrophobic residues upon protein misfolding can quickly lead to protein aggregation and cell death (Maisonneuve et al. 2008). In addition to the formation of amorphous aggregates, protein unfolding can also lead to the formation of highly ordered amyloid fibrils with a characteristic cross-β structure (Sawaya et al. 2007). Amyloid formation is associated with a number of neurodegenerative diseases in humans, including Parkinson's, Huntington's, Alzheimer's, Creutzfeldt-Jakob, and familial amyloid polyneuropathy, for which few, if any, treatment options are available (Aguzzi and O'Connor 2010). The mechanism of cell death by protein aggregates and amyloids remains unclear, but it is likely caused by a combination of a toxic gain of function of the aggregates/amyloids and sequestration of otherwise viable proteins, resulting in a loss of function (Olzscha et al. 2011; Winklhofer et al. 2008). The proteases and chaperones of the protein quality control network (PQCN) are tasked with preventing these deleterious events by maintaining protein homeostasis (proteostasis) in the varied, and sometimes harsh, cellular environments that challenge the integrity of the proteome (Gottesman et al. 1997) (Figure 1.1). The proteome is highly dynamic, with constant synthesis, degradation,

and modification of proteins, all of which present unique challenges to the protein quality control network.

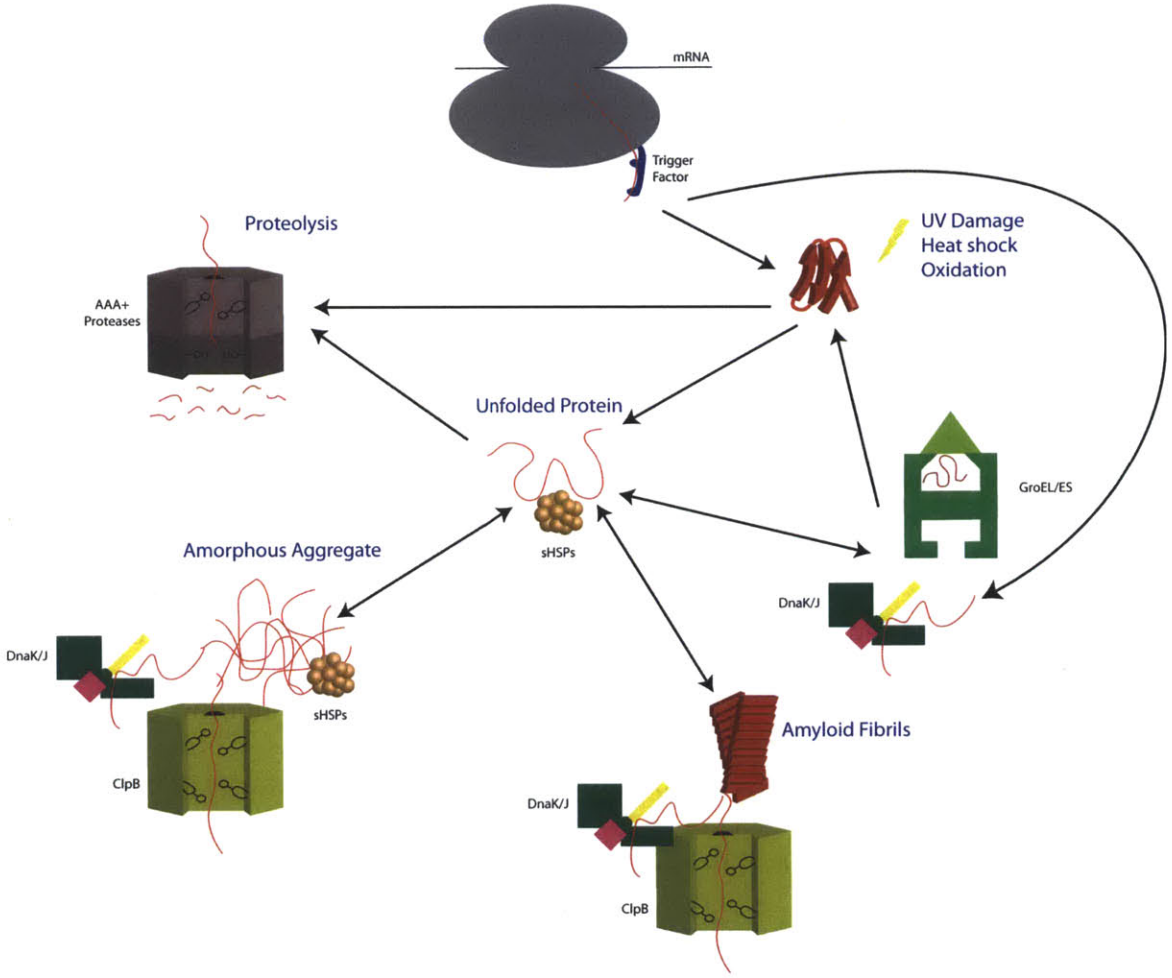

Figure 1.1. Schematic of protein quality control network in prokaryotes
Nascent proteins are engaged by the chaperone trigger factor, which maintains polypeptides in a soluble, unfolded state until translation is complete. Once translated, nascent proteins either fold spontaneously or with the assistance of chaperones. When a protein in no longer needed in the cell, it is removed by proteolysis. Mature proteins are subject to several forms of cellular stress, leading to protein unfolding and the formation of aggregates or amyloid fibrils. The sHSPs bind to unfolded proteins to maintain them in a soluble form and limit aggregation. ClpB, in conjunction with DnaK/J can resolubilize proteins that have formed amorphous aggregates or amyloid fibrils. Unfolded proteins are either degraded by proteolysis or refolded with the assistance of chaperones such as GroEL/ES, DnaK/J.

As a protein emerges from the exit tunnel of the ribosome, it is largely unfolded (Bhushan et al. 2010). Any contacts necessary for proper folding that are distant in the primary structure of a protein are unable to form until both portions of the polypeptide have been translated. Nascent proteins must therefore be maintained in a soluble, but unfolded, state until translation is complete to ensure proper formation of tertiary structure. Eukaryotes use a number of different chaperones to maintain the solubility of nascent polypeptides, but in prokaryotes, this task is primarily carried out by trigger factor, a ribosome associated chaperone (Pechmann et al. 2013).

As cells experience different environments or developmental stages, those that can best adjust their proteomes maintain a selective advantage. The proteome can be altered by transcriptional, translational, post-translational modifications, or proteolytic methods. Proteolytic regulation has the distinct advantages of being rapid and irreversible, making it the regulatory mechanism of choice for cellular processes that require directionality, such as the cell cycle (King et al. 1996). The destructive nature of proteolysis requires strict regulation, as uncontrolled proteolysis is lethal. Methods for proteolytic regulation are discussed more thoroughly in the next subsection.

A fully folded, mature protein is subject to a myriad of cellular stresses such as oxidation, UV-induced cross-linking, osmotic stress, and heat shock, all of which can cause protein misfolding and subsequent aggregation (Csermely and Yahara 2003). Although many members of the PQCN are constitutively expressed, cellular stresses frequently lead to increased expression of many PQCN proteins. In bacteria,

for example, many members of the PQCN are up regulated by the σ32 transcription factor following heat shock (Grossman et al. 1987). Irreversibly damaged proteins are often removed via proteolysis, whereas chaperones play a variety of roles such as preventing aggregation, sequestering misfolded proteins, disassembling aggregates, or assisting with protein refolding. This chapter provides a general overview of bacterial proteostasis, introducing the major cytosolic proteases and chaperones in *E. coli*, with a particular focus on the role of the Lon protease.

Holdases

One of the simplest mechanisms of chaperone action is utilized by the holdases, which prevent protein aggregation by binding misfolded proteins in an ATP-independent manner (Sun and MacRae 2005). Holdases play a prominent role in preventing formation of aggregates that scatter light in the lens of the human eye, thereby limiting cataract formation (Pande et al. 2001). The two most prominent holdases in *E. coli* are the small heat-shock proteins (sHSPs), IbpA and IbpB. The hallmark of sHSPs is a central α-crystalline domain flanked by unstructured N- and C-terminal tails that play roles in client binding and oligomerization (Sun and MacRae 2005). The basic subunit for both IbpA and IbpB is a homo-dimer, which can form large, heterologous oligomers that contain upwards of 32 subunits (Narberhaus 2002).

The mechanisms by which these chaperones function is still a matter of debate, as IbpB has been shown to be sufficient to prevent aggregation of the model substrate malate dehydrogenase, whereas IbpA, but not IbpB, is sufficient to prevent

aggregation of luciferase (Veinger et al. 1998; Ratajczak et al. 2009). Both IbpA and IbpB appear necessary for robust suppression of aggregation *in vivo* and interaction with other chaperones (Ratajczak et al. 2009). Likewise, the role of Ibp oligomerization in maintaining chaperones in a refolding competent state is poorly understood (Jiao et al. 2005). Interestingly, both IbpA and IbpB were found to be substrates of the Lon protease, suggesting a potential collaboration between these different arms of the protein quality control network (Bissonnette et al. 2010).

Refolding chaperones: Hsp60, Hsp70, and Hsp90

Although the holdases are effective at preventing protein aggregation, they are unable to catalyze protein refolding. *E. coli* contains two major chaperone systems; GroEL/ES (Hsp60/10) and DnaK/J (Hsp70/40), which promote protein refolding in an ATP-dependent manner. Prokaryotes also contain the Hsp90 chaperone, but, unlike its eukaryotic counterpart, prokaryotic Hsp90 is not essential and has only one confirmed binding partner in *E. coli* (Genest et al. 2012). As such, this section will focus on the well characterized Hsp60 and Hsp70 systems.

The GroEL chaperonin (Hsp60) works with GroES (Hsp10) to assist in refolding relatively small client proteins (< 60 kDa) by encapsulating them within a folding cage. The GroEL/ES system does not directly refold proteins, but rather provides an environment where off-pathway events, such as aggregation, are eliminated (Horwich et al. 2009). Indeed, many clients of this chaperonin are stabilized by long-range interactions and populate kinetically trapped folding intermediates (Vabulas et al. 2010). GroEL is a tetradecamer comprised of two

heptameric rings (*cis* and *trans*) stacked end to end. Each ring contains a large, central cavity, which is reversibly capped by the GroES heptamer. The *cis* and *trans*-rings interconvert in an alternating format, such that only one GroES heptamer is bound to GroEL at a time (Figure 1.2).

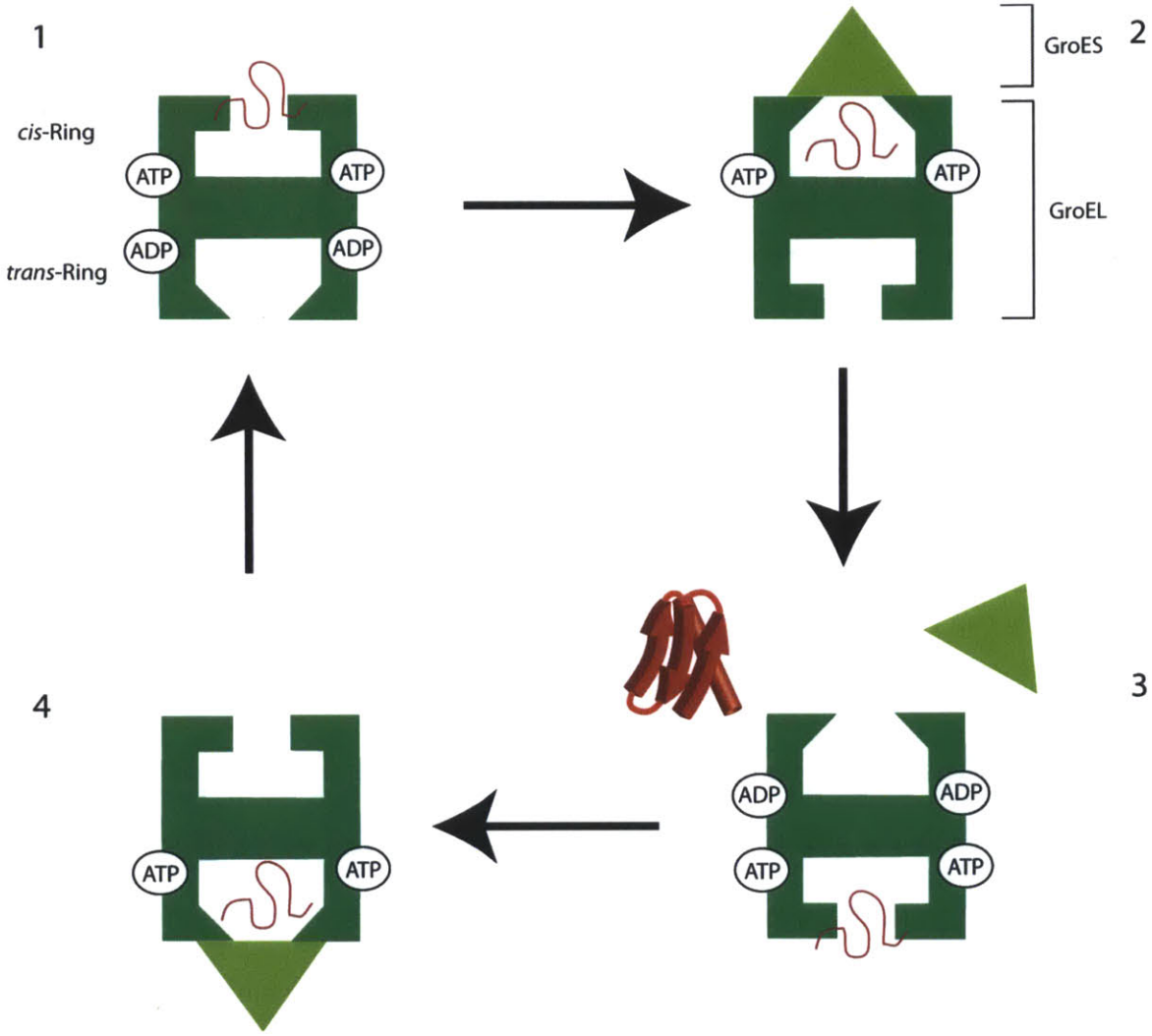

Figure 1.2 Mechanism of GroEL/ES chaperonin
In the initial stage of the reaction cycle, an unfolded client protein binds to the *cis*-ring of GroEL, causing dissociation of ADP from the *trans*-ring. GroES binding to the *cis*-ring encapsulates the client protein and causes a conformational change in GroEL, leading to formation of the folding chamber. ATP hydrolysis by the *cis* -ring allows GroES and the folded client protein to be released. ATP binding to the *trans* -ring allows for a new client protein to bind to the *trans* -ring and the cycle to begin again. Figure adapted from Lund 2009.

Misfolded client proteins bind GroEL via exposed hydrophobic patches on the *cis*-ring (Fenton et al. 1994). Encapsulation of the client by GroES leads to large conformational changes in GroEL such that the folding cage is now lined with highly hydrophilic residues that create a favorable refolding environment (Clare et al. 2012). After the *cis*-ring hydrolyzes ATP (~10-15 seconds), GroES dissociates and allows the client protein to leave the folding cage. If the client protein fails to refold properly, it can rebind and repeat the cycle.

The DnaK chaperone (Hsp70) promotes protein folding by transiently binding hydrophobic segments of client proteins, thereby shielding them from bulk solvent and preventing aggregation (Vabulas et al. 2010; Sharma et al. 2010). Upon release, client proteins can bury their hydrophobic sequences by achieving a mature fold, or they can rebind to DnaK and repeat the process. DnaK contains two domains, an N-terminal ATPase domain and a C-terminal peptide-binding domain that recognizes hydrophobic sequences that are normally buried in the core of folded proteins. In the ATP-bound state, substrates bind DnaK with a low affinity and exchange rapidly, whereas the ADP-bound state has high substrate affinity and displays slow exchange (Hartl et al. 2011). The DnaJ (Hsp40) co-chaperone targets misfolded proteins to DnaK and enhances the rate of ATP hydrolysis, leading to the stable ADP-bound state, whereas the GrpE co-chaperone promotes the exchange of ADP for ATP, which triggers substrate release (Laufen et al. 1999) (Figure 1.3). Eukaryotes posses numerous DnaK/J homologs, each with specialized functions, thereby increasing the interaction of the chaperone network (Sahi et al. 2013).

Figure 1.3

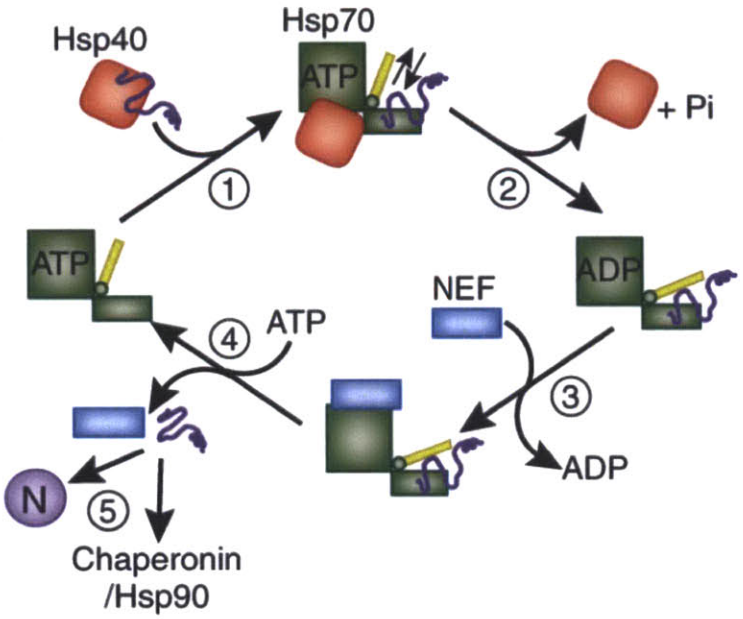

Figure 1.3 Mechanism of DnaK/J chaperone
Hsp40 (DnaJ) delivers client proteins to ATP-bound DnaK. DnaJ stimulates ATP hydrolysis by DnaK, switching the chaperone from a state with low substrate-binding affinity to one with high substrate-binding affinity. DnaK binds to hydrophobic segments of client proteins, protecting them from aggregation. GrpE serves as a nucleotide exchange factor (NEF) to stimulate the release of ADP and the client protein, which can subsequently refold to its native state or interact with other chaperones. Figure taken from Hartl and Hayer-Hartl 2009.

Disaggregases

When proteins aggregate, the ClpB disassembly chaperone (Hsp104 in yeast), in combination with the DnaK/J chaperone system, is capable of disassembling and resolubilizing these aggregates. ClpB is a hexameric, barrel shaped AAA+ unfoldase (ATPase Associated with various cellular Activities) that contains two ATPase domains connected by a coiled-coil middle domain (Desantis and Shorter 2011). ClpB undergoes ATP-dependent conformational changes that drive translocation of

client-proteins through a narrow central pore. This process can free individual polypeptide chains and destabilize the remaining aggregates. ClpB alone is a relatively poor disassembly chaperone, but activity is greatly enhanced by the DnaK/J system, which interacts with ClpB via the middle-domain (Rosenzweig et al. 2013). Order of addition experiments suggest that DnaK/J act upstream of ClpB, possibly by re-ordering the aggregates to expose segments of client proteins for translocation through the ClpB pore (Winkler et al. 2012).

Mechanisms for Regulating Proteolysis

Chaperones alone are insufficient for full maintenance of the proteome. Protein degradation is also required for removal of irreversibly damaged proteins. Due to its destructive and irreversible nature, proteolysis requires strict regulation, as uncontrolled proteolysis is lethal. Indeed, small molecules that lead to constitutive activation of the bacterial ClpP protease are being investigated as potential antibiotics (Leung et al. 2011).

Localization

Like all enzymatic reactions, proteolysis requires interaction between the protease and substrate. As such, this second-order (or higher) reaction is concentration dependent. Changes in the concentration of the protease or substrate within a defined environment (localization) can be used to regulate the rate of proteolysis. For example, ClpXP degrades the cell cycle regulator CtrA in *Caulobacter crescentus* only when both proteins co-localize to the pole (Iniesta et al.

2006). Alternatively, proteases may have sub-optimal activity in certain environments. For example, the general proteases of the lysosome (cathepsins) exhibit optimal activity within the acidic environment of the lysosome, which also destabilizes native protein structures, aiding degradation (Müller et al. 2012).

Cleavage-site specificity

Within a designated environment, proteolysis is sometimes regulated by strict specificity for the polypeptide cleavage-site. Some proteases have little to no specificity for cleavage-site selection (i.e., Lon), others have strict, but relatively general requirements (trypsin cleaves after lysine and arginine), and still others show exquisite specificity (the Tobacco Etch Virus protease will only recognize the following sequence ENLYFQG/S, with cleavage between glycine and serine) (Dougherty et al. 1988).

Oligomerization

Some proteases change oligomeric states to alter substrate specificity (discussed for Lon in Chapter 3) or to enhance substrate binding. For example, the periplasmic protease DegP converts from an inactive hexamer to a proteolytically active oligomer with 12, 18, 24, or 30 subunits (Krojer et al. 2010; Kim et al. 2011). Although proteolytically competent DegP mutants have been isolated that are unable to form large oligomers, cage assembly does appear to provide a selective advantage by increasing substrate-binding affinity and providing a protective effect

against rogue proteolysis (Kim and Sauer 2012) (Seokhee Kim, personal communication).

Allostery

Allosteric regulation of proteolytic activity is exemplified by the periplasmic protease DegS, which makes the initial cleavage in the trans-membrane protein RseA to initiate the σE stress-response pathway. DegS is a trimeric protease that contains a trypsin-like protease domain and a PDZ domain. Both the DegS protease and the RseA substrate are membrane anchored and exist in close proximity, but the PDZ domain inhibits degradation activity by maintaining the protease domain in an inactive conformation (Sohn et al. 2007). When periplasmic proteostasis is compromised, outer-membrane proteins accumulate in the periplasm and bind to the PDZ domain of DegS, relieving inhibition, and allowing for rapid proteolysis of RseA (Sohn et al. 2007; 2009).

AAA+ Proteases

Most proteolysis within the bacterial cytoplasm and most non-lysosomal degradation in the eukaryotic cytosol is carried out by AAA+ proteases, which have little cleavage-site specificity. These proteases form barrel-shaped homo-oligomers with the proteolytic active site sequestered within a central degradation chamber. The degradation chamber is capped by a ring-shaped hexamer that recognizes, unfolds, and translocates substrates into the chamber through a narrow central pore (Figure 1.4A).

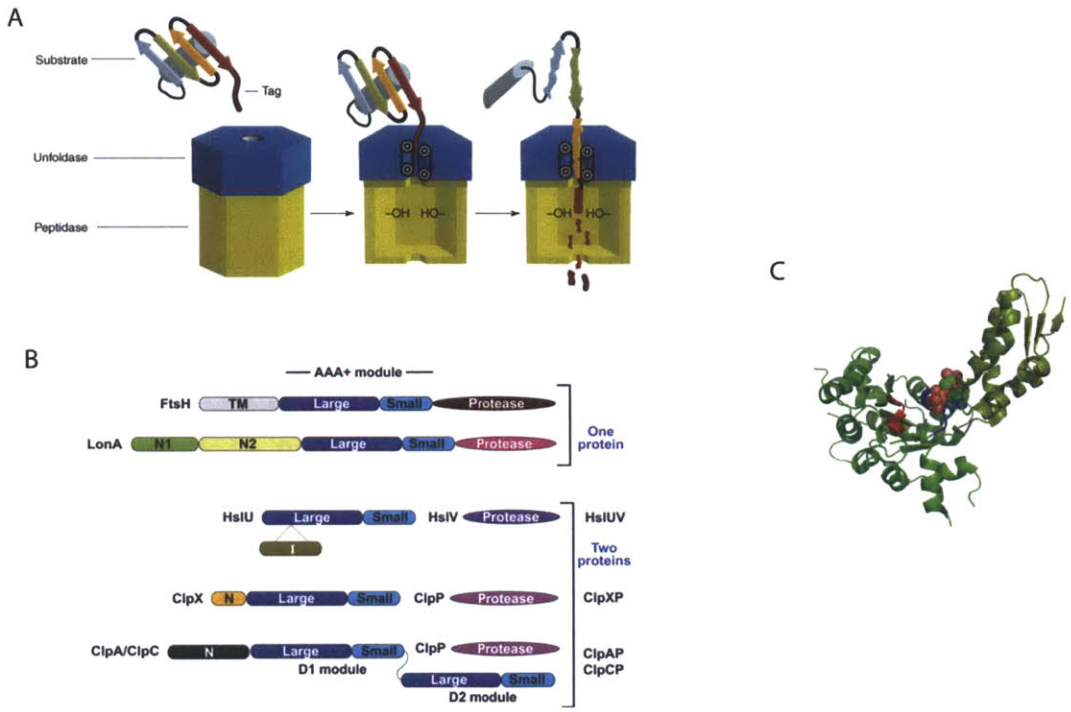

Figure 1.4 Overview of AAA+ proteases
(A) AAA+ proteases sequester the proteolytic active site within a degradation chamber formed by homo-oligomerization. The AAA+ unfoldase sits on top of the degradation chamber, and recognizes, unfolds, and translocates substrates through a pore in the center of the unfoldase and into the chamber. Figure taken from Baker and Sauer 2006. (B) Domain organization of AAA+ proteases in bacteria. FtsH and Lon have one polypeptide, whereas HslUV, ClpXP, and ClpAP have the unfoldase and protease in separate polypeptides. Figure adapted from Sauer and Baker 2011. (C) Structure of a single AAA+ module from ClpX (3HWS.pdb). The large and small AAA+ domains are different shades of green. The Walker A and B motifs are colored blue and red respectively and ADP is shown in CPK format.

The AAA+ module

The AAA+ family is a highly conserved family of proteins that convert the energy of ATP hydrolysis into mechanical work, often resulting in a change in substrate conformation. AAA+ enzymes are widely utilized in biology, playing a role in DNA replication (helicases), protein chaperones (ClpB), disassembly of protein complexes (NSF) or protein/DNA complexes (ClpX), and protein degradation (HslU)

(Enemark and Joshua-Tor 2008; Hanson and Whiteheart 2005; Sauer and Baker 2011). The AAA+ family is defined by a 200-250 residue ATP-binding module with a α/β subdomain followed by an α-helical subdomain (Hanson and Whiteheart 2005; Wendler et al. 2011) (Figure 1.4C). There are several motifs within these subdomains, the most prominent of which are the Walker A and Walker B, which are essential for nucleotide binding and hydrolysis, respectively (Walker et al. 1982).

AAA+ proteases in E. coli

E. coli utilizes a total of five AAA+ proteases, three with separate protease and ATPase modules (ClpXP, ClpAP, HslUV) and two with both modules found in a single polypeptide (FtsH, and Lon) (Figure 1.4B). The only essential AAA+ protease in *E. coli* is FtsH, which is membrane bound and responsible for degradation of misfolded membrane proteins and some cytoplasmic substrates (Bieniossek et al. 2006). Of the cytosolic proteases, ClpX and ClpA both bind to the same self-compartmentalized peptidase, ClpP. Because ClpP is a tetradecamer composed of two heptameric rings, there is a symmetry mismatch between a ClpP ring and its partner hexameric AAA+ ring (Alexopoulos et al. 2012). HslUV has no symmetry mismatch as the HslV protease is a dodecamer, composed of two hexameric rings (Sousa et al. 2002). Symmetry is also not an issue for Lon, the most widespread ATP-dependent protease, as both the protease and ATPase domains are contained in a single polypeptide, which assembles into a hexamer (Rotanova et al. 2006). The architecture of AAA+ proteases prevents uncontrolled proteolysis, but requires additional steps to allow for the proper substrates to access the degradation

chamber. Several different methods have evolved to regulate access to these self-compartmentalizing proteases.

Degrons

Substrates are targeted to AAA+ proteases by degradation tags (degrons). These degrons can be as simple as a single amino acid or as complex as post-translational attachment of several additional proteins. The N-end rule pathway has the simplest degron, consisting of a single hydrophobic amino acid (Leu, Phe, Tyr, Trp for prokaryotes) at the N-terminus of a substrate (Dougan et al. 2012). Because all proteins initially contain methionine as the N-terminal residue, which is not removed if the second residue is Leu, Phe, Tyr, or Trp, these degrons are often cryptic and only exposed after an initial endoproteolytic cleavage event (Mogk et al. 2007). In prokaryotes, N-end rule substrates are degraded by the ClpAP protease, whereas in eukaryotes they are degraded by the ubiquitin proteasome system (Dougan et al. 2012).

Degrons can also consist of several consecutive amino acids, as exemplified by the ssrA tag. When ribosomes translate truncated mRNA molecules, they translate to the end of the message and stall. To degrade aberrant translation products and free the stalled ribosomes, the tmRNA molecule enters the A-site of the ribosome and provides a template for the addition of the ssrA-tag sequence (AANDENYALAA-CO_2) to the nascent polypeptide (Moore and Sauer 2007). Once released, the ssrA-tagged polypeptide is recognized by ClpXP and ClpAP and degraded (Moore and Sauer 2005). Degrons can also be part of the endogenous

sequence of a protein. For example, attaching the first 13 residues of the Arc repressor to a model substrate is sufficient to target it for degradation by the HslUV protease (Burton et al. 2005).

In eukaryotes, most proteolysis is carried out by the ubiquitin-proteasome system. Proteins are targeted to the 26S proteasome for degradation by covalent attachment of polyubiquitin chains, but also require a long, unstructured region which is engaged by the AAA+ Rpt unfolding ring of the 19S regulatory particle (Inobe et al. 2011). Ubiquitin is a 76 amino-acid protein that is covalently attached to substrates via lysine residues. This post-translational modification is mediated by the E3 ubiquitin ligase, which in turn receives the activated ubiquitin from the E1 and E2 ubiquitin ligases. A typical mammalian genome contains 2, ~40, and ≥400 E1, E2, and E3 ligases, respectively, which allows for selective ubiquitination of thousands of different proteins (Randow and Lehner 2009).

Adaptors

AAA+ proteases typically degrade many different substrates, and often recognize several different degrons. Adaptors proteins can help prioritize the degradation of certain substrates to better adjust to cellular demands. For example, the SspB adaptor acts to tether ssrA-tagged substrates (and some additional substrates) to ClpXP, increasing affinity >10-fold (Wah et al. 2002). SspB binds to the N domain of ClpX and to the N-terminal part of the ssrA degron (AANDENY), leaving the last three residues of the ssrA tag (LAA-CO_2) free to bind in the ClpX axial pore (Flynn et al. 2001; Martin et al. 2008a). Unlike ClpX, ClpA recognizes features

of the ssrA degron that overlap with the SspB-binding site, and binding of SspB to the ssrA degron therefore enhances degradation by ClpXP while inhibiting degradation by ClpAP (Flynn et al. 2001). Adaptors can also enhance proteolysis by mechanisms other than tethering. For example, the ClpS adaptor protein enhances degradation of N-end rule substrates by ClpAP. However, ClpS truncation mutants are capable of binding ClpA and N-end rule substrates, but show no enhancement of degradation, suggesting a more active role in substrate delivery (Hou et al. 2008; Roman-Hernandez et al. 2011).

Tools for biochemical study

GFP has proven to be an invaluable model substrate for AAA+ proteases such as ClpXP. GFP fluorescence arises from a *p*-hydroxybenzylideneimidazolidinone chromophore formed by cyclization and oxidation of residues 65-67 (Heim et al. 1994). This chromophore is protected from bulk solvent by the β-barrel structure of the protein, which consists of 11 β-strands and several α-helices (Ormö et al. 1996) (Figure 1.5A). Protein unfolding exposes the chromophore to bulk solvent, which quenches fluorescence, making GFP a useful tool for dissecting the mechanism of protein unfolding and degradation.

For mechanistic studies, GFP is often targeted for degradation by the addition of a degron to the C-terminal 11th β-strand. Removal of the 11th β-strand of GFP does not expose the chromophore to bulk solvent, but still has a fluorescent fingerprint. This phenomenon arises because the phenol moiety of the chromophore can exist in either a protonated or an unprotonated state, each with

unique fluorescent properties. The unprotonated phenol moiety is excited by 467 nm light and emits 511 nm fluorescence (hereafter called 467 nm fluorescence). When the phenol moiety is in the neutral form, absorption of 400 nm light also leads to fluorescence at 511 nm (hereafter referred to as 400 nm fluorescence), but requires excited state proton transfer to remove a proton from the chromophore and transfer it to the Glu222 side chain on the 11th strand (Stoner-Ma et al. 2006; Kent et al. 2008) (Figure 1.5B). Removal of only the 11th strand of GFP eliminates the terminal proton acceptor (Glu222), but does not expose the chromophore, thus eliminating 400 nm, but not 467 nm fluorescence (Figure 1.5C). Using different changes in 400 and 467 nm fluorescence, ClpX mediated unfolding of GFP-ssrA has been shown to be a multi-step process, with the first step involving removal of the ssrA-tagged 11th β-strand from the β-barrel core of GFP (Martin et al. 2008b; Nager et al. 2011).

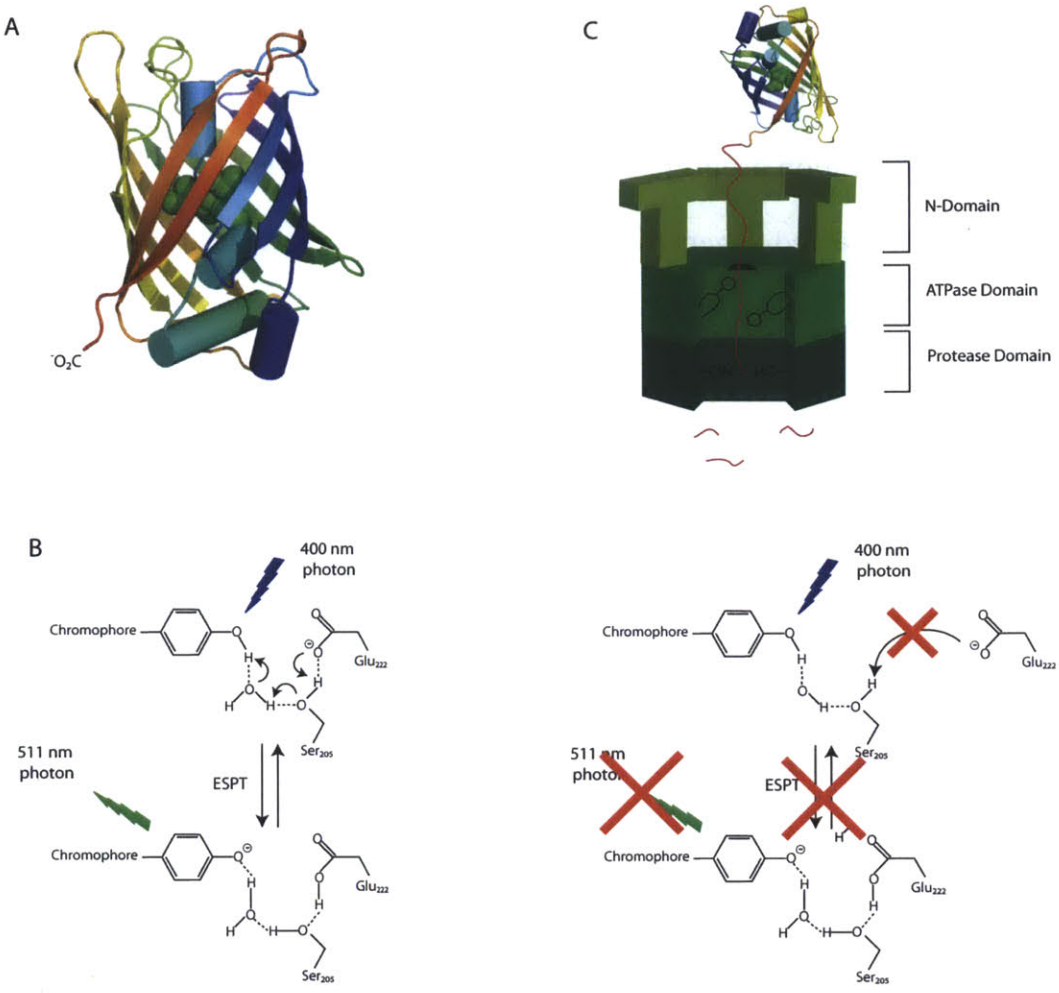

Figure 1.5 GFP as a tool for studying AAA+ proteases
(A) Crystal structure of enhanced GFP (1S6Z.pdb). The chromophore is shown in CPK format. (B) When excited with 400 nm light, GFP transfers a proton from the neutral chromophore onto glutamate 222, which is on the C-terminal 11th β-strand (red), by excited state proton transfer (ESPT). The anionic chromophore is then capable of emitting light at 511 nm. Figure adapted from Nager et al. 2011. (C) When the 11th β-strand of GFP is removed from the β-barrel core of GFP by a AAA+ protease, glutamate 222 is no longer able to serve as the terminal proton acceptor for ESPT and fluorescence from 400 nm excitation is lost.

Another useful property of GFP is the ability to create circular permutations that change the location of the N and C termini, while maintaining the overall fold and fluorescence of the protein. Because AAA+ mediated unfolding depends upon

the stability of structural motifs immediately adjacent to the degron, rather than global stability, changing the β-strand to which the degron is attached leads to different degradation properties (Baird et al. 1999; Nager et al. 2011). In Chapter 2, this feature is utilized to generate novel fluorescent substrates for the Lon protease.

Structure, Function, and Physiology of the Lon Protease

Physiological roles

Mutations in the *lon* gene were initially identified in bacteria in the 1960s as causing two distinct phenotypes, the formation of muccoid colonies and enhanced sensitivity to UV light (Markovitz 1964; Adler and Hardigree 1964). These phenotypes arise from increased cellular concentrations of two Lon substrates, RcsA and SulA, respectively. The RcsA transcription factor positively regulates expression of genes involved in the production of capsular polysaccharides, overproduction of which leads to the muccoid phenotype (Torres-Cabassa and Gottesman 1987). The cell-division inhibitor SulA is up-regulated after DNA damage as part of the SOS-response and inhibits cell division until the damage is repaired (Gottesman et al. 1981; Mizusawa and Gottesman 1983). Over the past 50 years, additional roles for the Lon protease have been identified in numerous biological processes in prokaryotes and eukaryotes.

One of the primary roles of Lon is to degrade misfolded or damaged proteins within cells, making it a member of the proteostasis network. Indeed, when *E. coli* are treated the amino-acid analog canavanine, which causes protein misfolding, there is a 50% decline in ATP-dependent protease activity in *lon*-defective strains

compared to wild-type strains (Chung and Goldberg 1981). Like many chaperones, Lon is thought to recognize misfolded proteins by binding exposed hydrophobic sequences that are normally buried in the native structure (Gur and Sauer 2008). Lon interacts with other members of the proteostasis network that also bind misfolded proteins. For example, the small heat shock proteins IbpA and IbpB are degraded by Lon (Bissonnette et al. 2010). Although Lon is not essential in *E. coli* under normal growth conditions, it becomes essential when other components of the protein quality control network are compromised (Tomoyasu et al. 2001).

Another phenotype associated with a *lon* knockout is decreased formation of bacterial persisters, which are relatively impervious to antibiotics and environmental insults and make up a tiny, slowly growing subpopulation within a culture of genetically identical, exponentially growing bacteria (Gerdes and Maisonneuve 2012). Bacterial persistence arises from genetically encoded toxin/antitoxin pairs such as RelE/B, YafO/N, YgiT/I, YoeB/YefM, and HipA/B (Christensen-Dalsgaard et al. 2010; Cherny et al. 2007; Nomura et al. 2004; Christensen et al. 2004). The uninhibited toxins inhibit cell growth but do not promote cell death, using a variety of family-specific mechanisms such as inhibition of translation, replication, or cell-wall synthesis, giving rise to the persister phenotype (Gerdes and Maisonneuve 2012). The labile antitoxin is typically a substrate for the Lon protease, but is also expressed at much higher levels than the respective toxin. Under normal conditions, all toxins are bound to and inhibited by their respective anti-toxins. Stochastic variation in the expression levels of the toxin, antitoxin, and Lon lead to a small fraction of cells with uninhibited toxins that inhibit

cell growth and give rise to persister cells (Figure 1.6A). Deletion of Lon leads to increased concentrations of antitoxins and a lower percentage of cells with uninhibited toxins, thereby reducing the number of persister cells within a population (Maisonneuve et al. 2011) (Figure 1.6B).

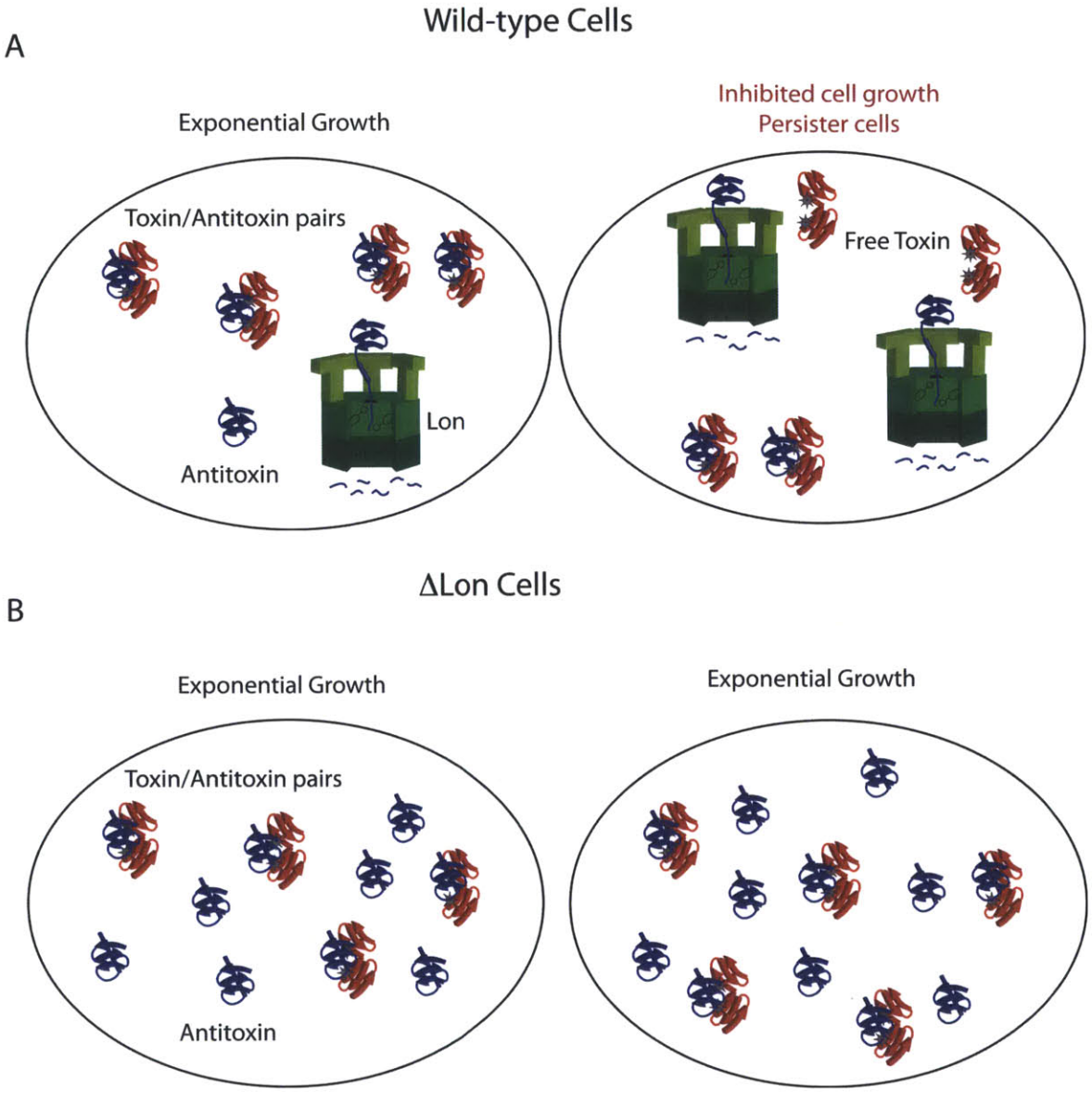

Figure 1.6 The role of Lon in persister-cell formation
(A) In wild-type cells, most toxin-proteins are associated with an antitoxin protein and are therefore inert. There is an excess of free antitoxin proteins, which are degraded by the Lon protease. Stochastic variation in expression of Lon and antitoxin can lead to cells with free toxins, which in turn inhibit cell growth and

generate persister cells. (B) In cells where the Lon protease is deleted, antitoxins are not degraded, leading to a large excess of antitoxins to toxins. All toxins in a cell are bound antitoxin and persister-cell formation is reduced.

Lon appears to be the only DNA-binding AAA+ protease in *E. coli* (Gur 2013) (Charette et al. 1984). Although the significance of this DNA-binding activity is unknown, it is interesting that Lon degrades numerous DNA-binding proteins. In addition to degrading anti-toxins, which bind to DNA to regulate their own synthesis, Lon also degrades transcription factors regulating the multiple antibiotic-resistance pathway (MarA), capsular-polysaccharide synthesis (RcsA), acid-resistance genes (GadE), and the superoxide response (SoxS), giving *lon* knockouts a variety of phenotypic defects (Duval et al. 2009; Heuveling et al. 2008; Griffith et al. 2004; Torres-Cabassa and Gottesman 1987). A number of nucleoid-associated proteins are also Lon substrates, including HUβ, TrfA, and StpA, as well as the translesion polymerase subunit UmuD (Liao et al. 2010; Kubik et al. 2011; Johansson and Uhlin 1999; Gonzalez et al. 1998).

Although Lon is not normally essential in *E. coli*, it plays a vital role in many other bacteria. In *C. crescentus*, Lon activity is necessary for rapid progression through the cell cycle, as it degrades the regulatory proteins CcrM and SciP (Wright et al. 1996; Gora et al. 2013). Lon mutants have been shown to lead to reduced virulence in the pathogenic bacteria *Salmonella typhimurium*, *Pseudomonas aeruginosa*, and *Brucella abortus* (Robertson et al. 2000; Ingmer and Brøndsted 2009; Breidenstein et al. 2012b; 2012a).

In eukaryotes, Lon is found in the mitochondrial matrix and in peroxisomes.

Mitochondrial Lon (mtLon) degrades the DNA-binding protein TFAM, which regulates mtDNA copy number (Lu et al. 2012; Matsushima et al. 2010). Like its prokaryotic homolog, mtLon is also the primary protease responsible for clearance of oxidized and misfolded proteins within mitochondria (Venkatesh et al. 2011; Lee and Suzuki 2008). Interestingly, overexpression of mtLon in yeast leads to a nearly two-fold increase in organismal lifespan, whereas an RNAi knockdown of mtLon is particularly toxic to lymphoma cells (Luce and Osiewacz 2009; Bernstein et al. 2012). Cancer cells are subject to increased oxidative and proteotoxic stress, and therefore are heavily reliant on stress-response proteins, such as Lon (Luo et al. 2009). Understanding how Lon recognizes misfolded and/or damaged proteins could provide a new target for anti-cancer drugs.

Structural organization

There are two established families of the Lon protease: LonA and LonB. The cytosolic LonA family is found in bacteria and eukaryotes, whereas the membrane anchored LonB family is found predominantly in archaea (Rotanova et al. 2006). The LonA family contains a large N domain, an ATPase domain, and a C-terminal peptidase domain, whereas the LonB family lacks the N domain but has two transmembrane segments inserted between the Walker A and Walker B motifs of the ATPase domain (Figure 1.7A). LonA and LonB families contain different consensus sequences for the proteolytic active sites and the Walker A and B motifs (Rotanova et al. 2004). Many genomes also contain isolated domains with strong homology to the N domain or peptidase domain of Lon. Recently, a protein with a

LonB-like peptidase domain and a large N domain bearing no homology to other Lon domains was shown to be proteolytically active (Liao et al. 2012) . This newly classified LonC family is an area of active investigation.

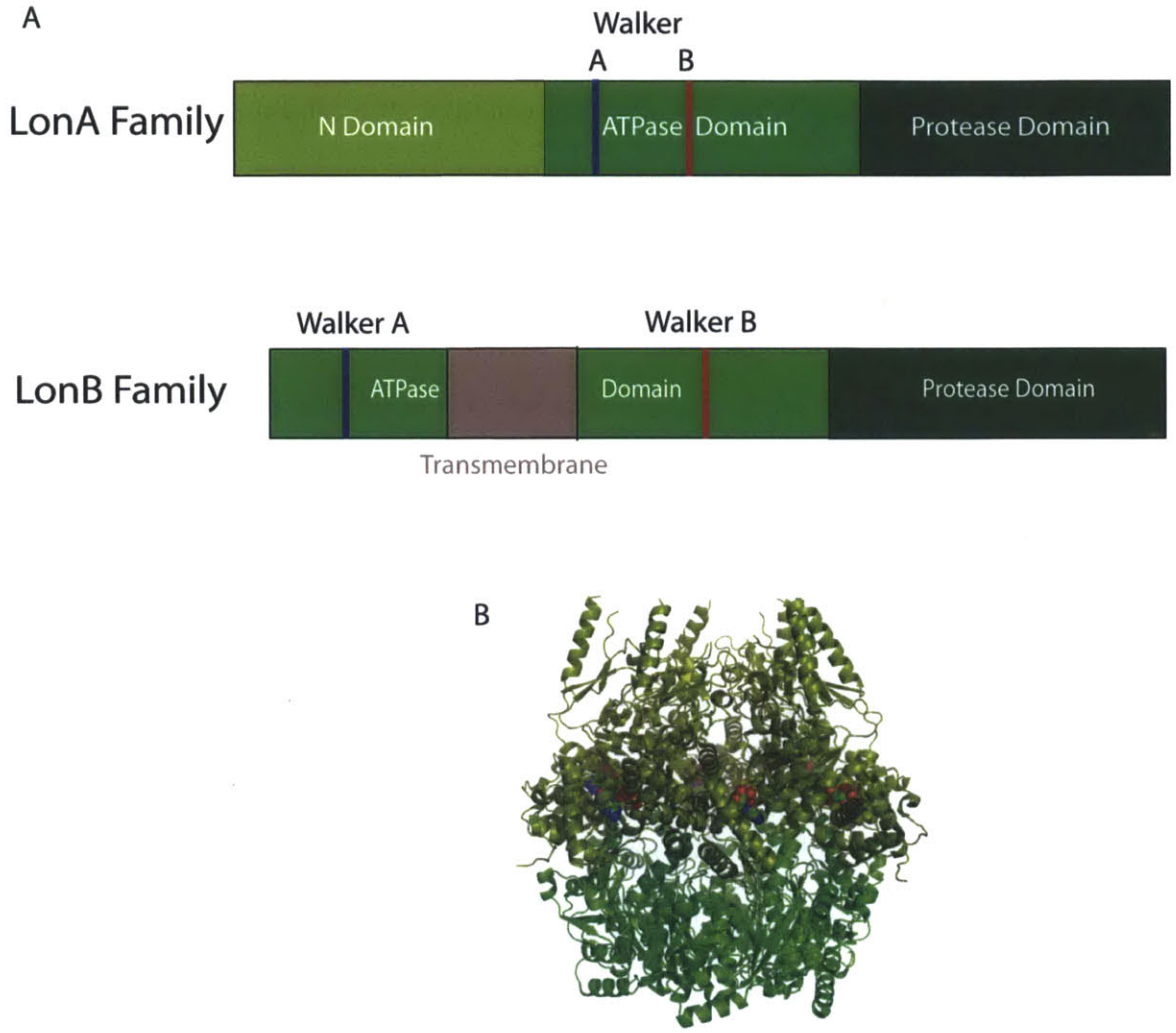

Figure 1.7 Architecture of Lon protease
(A) Domain organization of LonA and LonB families. LonA contains a large N domain that is absent in the LonB family. LonB contains a transmembrane region inserted between the Walker A (blue) and Walker B (red) motifs in the ATPase domain. (B) Crystal structure of *Thermococcus onnurineus* Lon (3K1J.pdb). The transmembrane regions of this LonB protein were removed for crystallization. The ATPase domain and protease domains are colored differently for ease of identification, and ADP molecules are shown in CPK representation.

E. coli Lon was originally thought to form a tetramer, but subsequent negative stain EM of full-length Lon and crystal structures of the isolated peptidase domain showed that Lon exists as a hexamer or a dodecamer (see Chapter 3) (Botos et al. 2004; Goldberg et al. 1994). The most informative structural insights come from the recent crystal structure of the archaea *Thermococcus onnurineus* Lon, a member of the LonB family, that was truncated to remove the transmembrane segment, but otherwise formed a functional hexamer (Cha et al. 2010) (Figure 1.7B). A full-length crystal structure of hexameric LonA has not been solved, but structures are known for overlapping segments of *Bacillus subtilis* LonA, which crystallize as helical filaments (Duman and Löwe 2010).

Lon does not contain the canonical catalytic triad found in serine proteases, but instead utilizes a Ser/Lys catalytic dyad, with lysine acting as a general base (Botos et al. 2004). Like other Ser/Lys proteases, Lon is relatively impervious to classical serine-protease inhibitors such as sulfonyl fluorides, chloromethyl ketones, and fluorophosphates (Ekici et al. 2008). Crystal structures of the LonA and LonB peptidase domains show multiple conformations of the proteolytic active site with some structures catalytically competent and others catalytically incompetent (Botos et al. 2004; Im et al. 2004; Botos et al. 2005). These results support a model in which the proteolytic active sites switch conformations, possibly in response to an allosteric signal (discussed further below).

An undefined role for the N domain

Unlike the peptidase and ATPase domains, the role of the Lon N domain remains poorly understood. The N domain of *E. coli* Lon contains two subdomains. The first subdomain (residues 1-117) contains three twisted β-sheets and one α-helix, whereas the second subdomain (residues 118-310) is predominantly α-helical and contains two predicted coiled-coil regions, possibly separated by an extended α-helix (Li et al. 2010) (Figure 1.8A). The first subdomain contains a similar fold to BPP1347 (PDB=1ZBO), a protein of unknown function from *Bordetella parapertussis* (Li et al. 2005).

Several crystal structures of parts of the N domain are available, but these have failed to yield a consensus on the quaternary structure of the domain or its location relative to the peptidase and ATPase domains. Fragments of the *E. coli* N domain corresponding to residues 1-119 and 1-245 were crystallized as an octamer and a monomer respectively (Li et al. 2005; 2010) (Figure 1.8A). Crystallization of a fragment corresponding to residues 1-209 of *B. subtilis* Lon yielded an anti-parallel domain-swapped dimer with coiled-coil interactions between the extended helices of each monomer (Duman and Löwe 2010) (Figure 1.8B). The anti-parallel nature of the domain-swapped dimer is consistent with a head-to-head interaction of two N domains rather than a planar hexameric arrangement. The significance of this observation is discussed more extensively in Chapter 3.

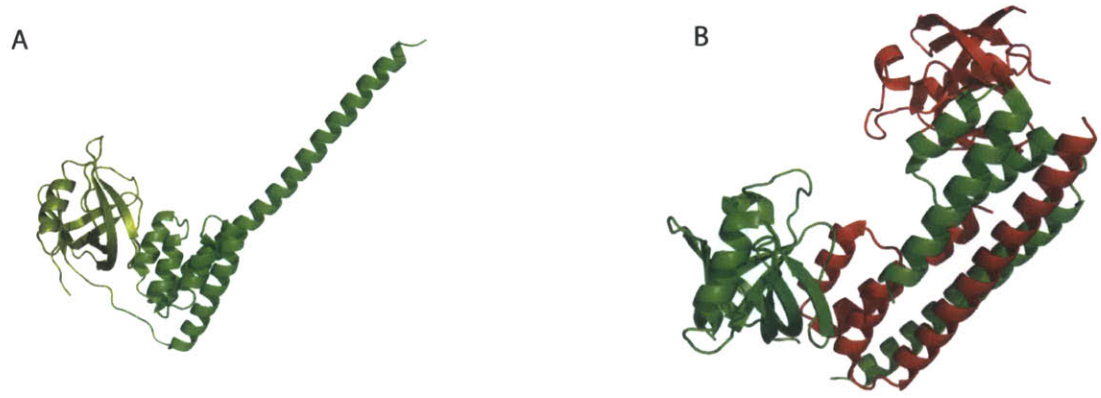

Figure 1.8 Crystal structures of the N domain
(A) Crystal structure of monomeric *E. coli* Lon N domain, residues 1-245 (3LJC.pdb). The two subdomains are colored different shades of green. (B) Crystal structure of domain-swapped dimer of *B. subtilis* Lon N domain, residues 1-209 (3M65.pdb). The two subunits are colored red and green to emphasize the head-to-head nature of the interaction between subunits.

The N domain appears to be necessary but not sufficient for Lon hexamerization as N-domain truncations in several species have resulted in altered oligomeric states, but isolated N domain fragments do not behave as hexamers (Lee et al. 2004b; Roudiak and Shrader 1998; Rudyak and Shrader 2000; Li et al. 2005). The coiled-coil region in the C-terminal part of the N domain appears to undergo ATP-dependent movements as judged by hydrogen-deuterium exchange and limited-proteolysis experiments (Cheng et al. 2012; Vasilyeva et al. 2002). Indeed, mutations in this region affect proteolytic and ATPase activity (Adam et al. 2012; Cheng et al. 2012). Several groups have suggested, but not rigorously demonstrated, that the N domain is involved in substrate binding (Ebel et al. 1999; Adam et al. 2012; Roudiak and Shrader 1998; Rudyak and Shrader 2000). The role of the N domain in substrate binding is addressed in Chapters 4 and 5.

Regulation of Lon Activity

As noted above, Lon degrades many different substrates. Although several Lon degrons have been identified, the mechanism by which many substrates are targeted for degradation remains unclear (Higashitani et al. 1997; Shah and Wolf 2006; Gur and Sauer 2008). Of particular interest is how Lon prioritizes degradation of different substrates.

No adaptor proteins have been identified that enhance Lon protease activity or shift substrate specificity. The only protein identified to date that regulates Lon activity without itself being degraded is the phage protein PinA, which inhibits protease and ATPase activity (Hilliard et al. 1998). Interestingly, interaction of Lon with DNA or polyphosphate has been shown to alter substrate selectivity, although the mechanism and physiological significance of these observations remains unclear (Kubik et al. 2011; Kuroda et al. 2001; Charette et al. 1984). Whereas other AAA+ proteases, such as ClpXP, change location in a cell-cycle dependent manner, LonA appears to be stably associated with the nucleoid throughout the cell cycle, suggesting that localization is not a prominent regulatory mechanism (Winkler et al. 2010). Two prevailing mechanisms for controlling Lon activity involve degron occlusion and allosteric regulation. Degron occlusion by proper protein folding or interaction with binding partners (i.e. toxin/anti-toxin systems) has been described above, whereas allosteric regulation is addressed below.

Allosteric regulators

The protease and ATPase activities of Lon are subject to allosteric regulation. Nucleotide binding has been shown to affect the conformation of the proteolytic active site, whereas substrate binding stimulates both ATPase and peptidase activity (Rudyak and Shrader 2000; Waxman and Goldberg 1982). Interestingly, Lon activity is influenced by the characteristics of a particular degron. When the sul20 degron (isolated from the C-terminus of SulA) or the β20 degron (isolated from an internal segment of β-galactosidase buried in the native protein) were attached to the same unfolded model substrate (CM-titin^{I27}), different V_{max} values for degradation and ATPase activity were obtained (Gur and Sauer 2009) (Figure 1.9A&B). Surprisingly, although CM-titin^{I27}-sul20 was degraded significantly faster than CM-titin^{I27}-β20, the later was a stronger stimulator of Lon ATPase activity, giving rise to a 30-fold difference in the efficiency of substrate degradation (ATP hydrolyzed per substrate degraded) (Figure 1.9C).

A

sul20	β20
ASSHATRQLSGLKIHSNLYH	QLRSLNGEWRFAWFPAPEAVE
Isolated from C-terminus of SulA	Isolated from β-galactosidase 49-68

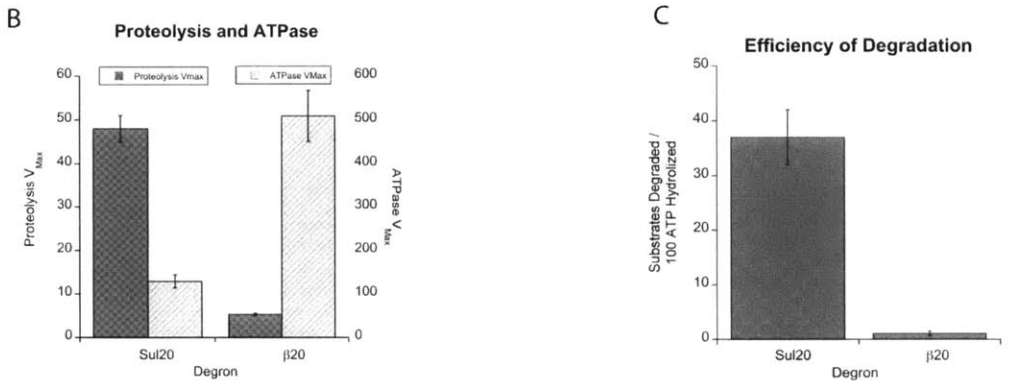

Figure 1.9 Characteristics of the sul20 and β20 degrons
(A) Origin and sequence of the sul20 and β20 degrons. Because the β20 degron was isolated from a fragment of β-galactosidase that is normally buried in the hydrophobic core, this degron is probably representative of misfolded substrates. (B) V_{max} for proteolysis and ATPase stimulation by the unfolded model substrate, CM-titin^{I27}, with a sul20 or β20 degron on the C terminus. Data taken from Gur and Sauer 2009. (C) Efficiency of degradation (substrates degraded per 100 ATP hydrolyzed) for CM-titin^{I27}-sul20 and CM-titin^{I27}-β20. Data taken from Gur and Sauer 2009.

These results led to the proposal that Lon equilibrates between three different states and that different degrons serve as allosteric regulators to bias the equilibrium between these states (Gur and Sauer 2009) (Figure 1.10A). In this model, the Lon^{off} state, which has no protease or ATPase activity, predominates in the absence of substrate. Binding of the sul20 degron biases the equilibrium to the Lon^{deg} state, which has high protease, but low ATPase activity. Binding of the β20

degron biases the equilibrium towards a chaperone-like state (Lon^{on}) with high ATPase activity, but little to no protease activity. Because there is an equilibrium between all three states, even under saturating concentrations of a β20-tagged substrate, a small population of Lon is in the Lon^{deg} state, which gives rise to slow degradation of these substrates.

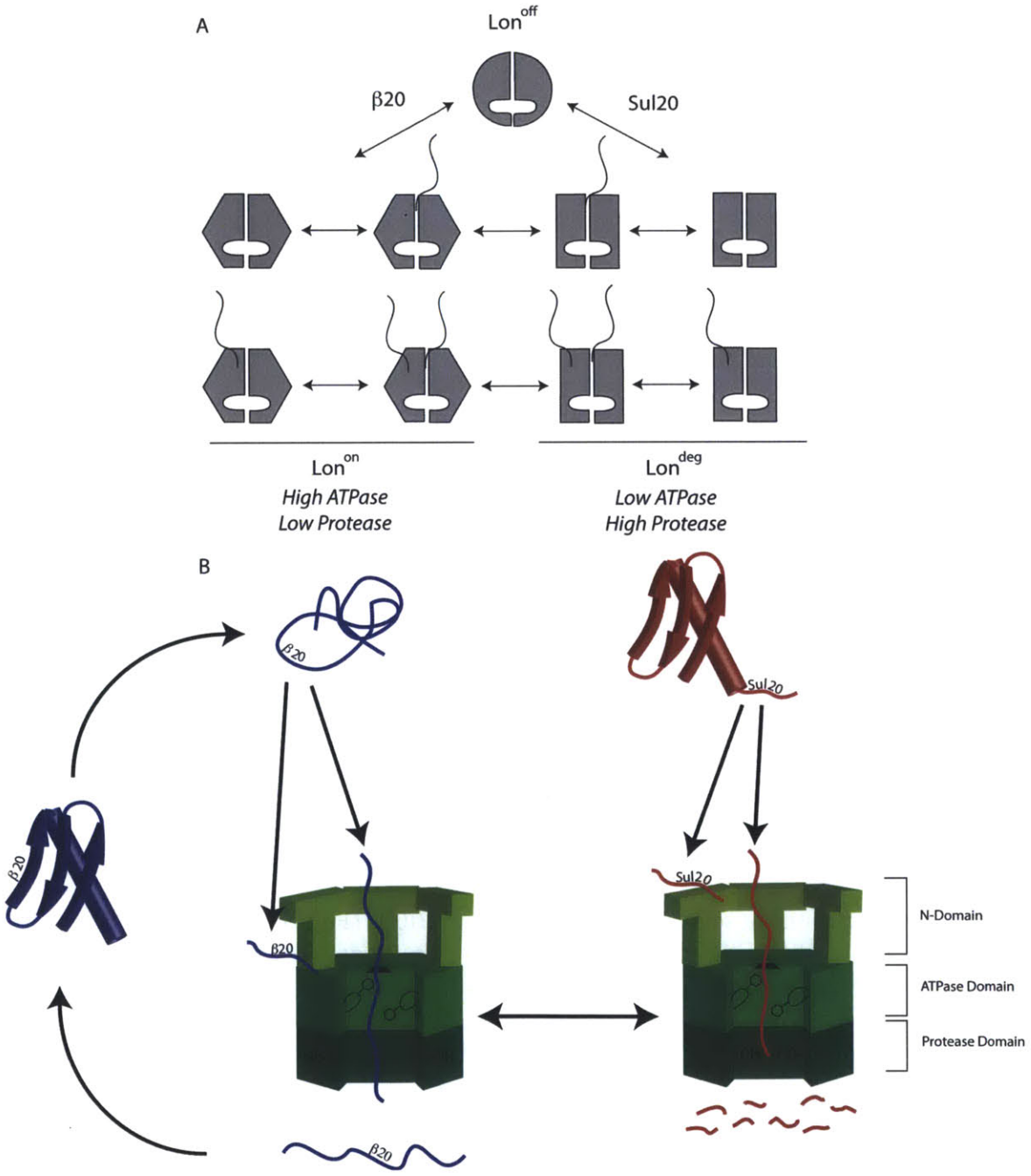

Figure 1.10 Model for degron-mediated allosteric regulation of Lon
(A) Three-state allosteric model for Lon. The Lon^{off} state (top) has no protease or ATPase activity. Binding of the sul20 degron shifts the equilibrium towards the Lon^{deg} state with low ATPase activity and high protease activity. Binding of the β20-degron biases the equilibrium towards the Lon^{on} state, which has high ATPase activity, but little to no protease activity. Figure adapted from Gur and Sauer 2009. (B) Model for Lon chaperone activity. When Lon encounters misfolded proteins (β20 degron), the proteolytic active site is malformed and inactive. Lon uses ATP hydrolysis to remodel misfolded proteins and release them back into solution without degradation, where they can subsequently refold. When Lon encounters the sul20 degron, the proteolytic active site is allosterically activated to allow for robust degradation of regulatory proteins that need to be removed quickly, such as the cell-division inhibitor SulA.

Lon as a chaperone

The allosteric model suggests that Lon is geared for the rapid removal of proteins that can be detrimental to growth (i.e. the cell-division inhibitor SulA) but may also act as a chaperone, possibly remodeling misfolded proteins without degradation. Although counterintuitive, an enzyme serving as both a protease and a chaperone is an often-repeated theme in biology. The periplasmic protease DegP is capable of increasing cell viability during extreme heat shock even when it is proteolytically inactive (Clausen et al. 2011). The AAA+ unfoldase ClpX remodels substrates in a ClpP-independent manner (Burton et al. 2001). Likewise, the membrane-bound FtsH protease has been shown to assist in the assembly of protein complexes within the plasma membrane and during protein export (Suzuki et al. 1997).

There are many hints in the literature that Lon can act as a chaperone. Variants of *Borrelia burgdorferi* and *Brevibacillus thermoruber* Lon with Ser to Ala active-site mutations in the peptidase domain have been shown to reduce insulin

aggregation *in vitro* (Coleman et al. 2009; Lee et al. 2004a). Overproduction of proteolytically inactive mtLon in *Saccharomyces cerevisiae* leads to increased assembly of the F_0F_1 ATPase and cytochrome C oxidase complex, rescuing respiration-dependent growth (Rep et al. 1996).

Lon could function as a chaperone by many mechanisms. The interaction of Lon with misfolded proteins could lead to a holdase-like function that suppresses protein aggregation. Alternatively, Lon could act like the GroEL/ES system, translocating substrates into a chamber, where protein folding can occur. Indeed, it has been proposed that proteolytically inactive Lon sequesters substrates within the protease chamber (Van Melderen and Gottesman 1999). Finally, Lon could use its unfoldase activity to remodel misfolded proteins in a manner analogous to ClpX or ClpB.

Thesis Overview

I have introduced the main components of the protein quality control network, proteases and chaperones, with a focus on the physiological role and allosteric regulation of Lon, including the potential of Lon to act as a chaperone as well as a protease. In the following chapters, I describe experiments that explore how Lon activity is regulated, with a focus on the role of the N domain. In Chapter 2, I describe the generation and characterization of GFP variants that are model substrates for Lon. Chapter 3 describes collaborative experiments that show that Lon hexamers equilibrate with dodecamers. The dodecamer is stabilized by interactions between N domains and displays altered substrate profiles. In Chapter

4, I describe the identification of the sul20-binding site within the N domain and subsequent characterization of this site in allosteric regulation of Lon activity. Chapter 4 also provides *in vivo* evidence for Lon chaperone activity. In Chapter 5, I characterize a mutant that was thought to inhibit substrate binding, but I instead show that it leads to altered oligomeric states and reduced protease and ATPase activity. Finally, in Chapter 6, I suggest future directions for investigating Lon activity.

References

Adam C, Picard M, Déquard-Chablat M, Sellem CH, Hermann-Le Denmat S, Contamine V. 2012. Biological roles of the Podospora anserina mitochondrial Lon protease and the importance of its N-domain. *PLoS ONE* **7**: e38138.

Adler HI, Hardigree AA. 1964. Cell elongation in strains of Escherichia coli. *J Bacteriol* **87**: 1240–1242.

Aguzzi A, O'Connor T. 2010. Protein aggregation diseases: pathogenicity and therapeutic perspectives. *Nat Rev Drug Discov* **9**: 237–248.

Alexopoulos JA, Guarné A, Ortega J. 2012. ClpP: A Structurally Dynamic Protease Regulated by AAA+ Proteins. *J Struct Biol.*

Baird GS, Zacharias DA, Tsien RY. 1999. Circular permutation and receptor insertion within green fluorescent proteins. *Proc Natl Acad Sci USA* **96**: 11241–11246.

Baker TA, Sauer RT. 2006. ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends Biochem Sci* **31**: 647–653.

Bernstein SH, Venkatesh S, Li M, Lee J, Lu B, Hilchey SP, Morse KM, Metcalfe HM, Skalska J, Andreeff M, et al. 2012. The mitochondrial ATP-dependent Lon protease: a novel target in lymphoma death mediated by the synthetic triterpenoid CDDO and its derivatives. *Blood.*

Bhushan S, Gartmann M, Halic M, Armache J-P, Jarasch A, Mielke T, Berninghausen O, Wilson DN, Beckmann R. 2010. alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nat Struct Mol Biol* **17**: 313–317.

Bieniossek C, Schalch T, Bumann M, Meister M, Meier R, Baumann U. 2006. The molecular architecture of the metalloprotease FtsH. *Proc Natl Acad Sci USA* **103**: 3066–3071.

Bissonnette SA, Rivera-Rivera I, Sauer RT, Baker TA. 2010. The IbpA and IbpB small heat-shock proteins are substrates of the AAA+ Lon protease. *Mol Microbiol.*

Botos I, Melnikov EE, Cherry S, Kozlov S, Makhovskaya OV, Tropea JE, Gustchina A, Rotanova TV, Wlodawer A. 2005. Atomic-resolution crystal structure of the proteolytic domain of Archaeoglobus fulgidus lon reveals the conformational variability in the active sites of lon proteases. *J Mol Biol* **351**: 144–157.

Botos I, Melnikov EE, Cherry S, Tropea JE, Khalatova AG, Rasulova F, Dauter Z, Maurizi MR, Rotanova TV, Wlodawer A, et al. 2004. The catalytic domain of Escherichia coli Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J Biol Chem* **279**: 8140–8148.

Breidenstein EBM, Bains M, Hancock REW. 2012a. Involvement of the Lon protease in the SOS response triggered by ciprofloxacin in Pseudomonas aeruginosa PAO1. *Antimicrobial Agents and Chemotherapy.*

Breidenstein EBM, Janot L, Strehmel J, Fernandez L, Taylor PK, Kukavica-Ibrulj I, Gellatly SL, Levesque RC, Overhage J, Hancock REW. 2012b. The Lon Protease Is Essential for Full Virulence in Pseudomonas aeruginosa. *PLoS ONE* **7**: e49123.

Burton BM, Williams TL, Baker TA. 2001. ClpX-mediated remodeling of mu transpososomes: selective unfolding of subunits destabilizes the entire complex. *Molecular Cell* **8**: 449–454.

Burton RE, Baker TA, Sauer RT. 2005. Nucleotide-dependent substrate recognition by the AAA+ HslUV protease. *Nat Struct Mol Biol* **12**: 245–251.

Cha S-S, An YJ, Lee C-R, Lee HS, Kim Y-G, Kim SJ, Kwon KK, De Donatis GM, Lee J-H, Maurizi MR, et al. 2010. Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *EMBO J* **29**: 3520–3530.

Charette MF, Henderson GW, Doane LL, Markovitz A. 1984. DNA-stimulated ATPase activity on the lon (CapR) protein. *J Bacteriol* **158**: 195–201.

Cheng I, Mikita N, Fishovitz J, Frase H, Wintrode P, Lee I. 2012. Identification of a Region in the N-Terminus of Escherichia coli Lon That Affects ATPase, Substrate Translocation and Proteolytic Activity. *J Mol Biol.*

Cherny I, Overgaard M, Borch J, Bram Y, Gerdes K, Gazit E. 2007. Structural and thermodynamic characterization of the Escherichia coli RelBE toxin-antitoxin system: indication for a functional role of differential stability. *Biochemistry* **46**:

12152–12163.

Christensen SK, Maenhaut-Michel G, Mine N, Gottesman S, Gerdes K, Van Melderen L. 2004. Overproduction of the Lon protease triggers inhibition of translation in Escherichia coli: involvement of the yefM-yoeB toxin-antitoxin system. *Mol Microbiol* **51**: 1705–1717.

Christensen-Dalsgaard M, Jørgensen MG, Gerdes K. 2010. Three new RelE-homologous mRNA interferases of Escherichia coli differentially induced by environmental stresses. *Mol Microbiol* **75**: 333–348.

Chung CH, Goldberg AL. 1981. The product of the lon (capR) gene in Escherichia coli is the ATP-dependent protease, protease La. *Proc Natl Acad Sci USA* **78**: 4931–4935.

Clare DK, Vasishtan D, Stagg S, Quispe J, Farr GW, Topf M, Horwich AL, Saibil HR. 2012. ATP-Triggered Conformational Changes Delineate Substrate-Binding and -Folding Mechanics of the GroEL Chaperonin. *Cell* **149**: 113–123.

Clausen T, Kaiser M, Huber R, Ehrmann M. 2011. HTRA proteases: regulated proteolysis in protein quality control. *Nat Rev Mol Cell Biol* **12**: 152–162.

Coleman JL, Katona LI, Kuhlow C, Toledo A, Okan NA, Tokarz R, Benach JL. 2009. Evidence that two ATP-dependent (Lon) proteases in Borrelia burgdorferi serve different functions. *PLoS Pathog* **5**: e1000676.

Csermely P, Yahara I. 2003. Molecular Pathomechanisms and New Trends in Drug Research - Google Books. *... and New Trends in Drug Research.*

Desantis ME, Shorter J. 2011. The elusive middle domain of Hsp104 and ClpB: Location and function. *Biochim Biophys Acta.*

Dougan DA, Micevski D, Truscott KN. 2012. The N-end rule pathway: from recognition by N-recognins, to destruction by AAA+proteases. *Biochim Biophys Acta* **1823**: 83–91.

Dougherty WG, Carrington JC, Cary SM, Parks TD. 1988. Biochemical and mutational analysis of a plant virus polyprotein cleavage site. *EMBO J* **7**: 1281–1287.

Duman RE, Löwe J. 2010. Crystal structures of Bacillus subtilis Lon protease. *J Mol Biol* **401**: 653–670.

Duval V, Nicoloff H, Levy SB. 2009. Combined Inactivation of lon and ycgE Decreases Multidrug Susceptibility by Reducing the Amount of OmpF Porin in Escherichia coli. *Antimicrobial Agents and Chemotherapy* **53**: 4944–4948.

Ebel W, Skinner MM, Dierksen KP, Scott JM, Trempy JE. 1999. A conserved domain in

Escherichia coli Lon protease is involved in substrate discriminator activity. *J Bacteriol* **181**: 2236–2243.

Ekici ÖD, Paetzel M, Dalbey RE. 2008. Unconventional serine proteases: Variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci* **17**: 2023–2037.

Enemark EJ, Joshua-Tor L. 2008. On helicases and other motor proteins. *Current Opinion in Structural Biology* **18**: 243–257.

Fenton WA, Kashi Y, Furtak K, Horwich AL. 1994. Residues in chaperonin GroEL required for polypeptide binding and release. *Nature* **371**: 614–619.

Flynn JM, Levchenko I, Seidel M, Wickner SH, Sauer RT, Baker TA. 2001. Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proc Natl Acad Sci USA* **98**: 10584–10589.

Genest O, Reidy M, Street TO, Hoskins JR, Camberg JL, Agard DA, Masison DC, Wickner S. 2012. Uncovering a Region of Heat Shock Protein 90 Important for Client Binding in E. coli and Chaperone Function in Yeast. *Molecular Cell.*

Gerdes K, Maisonneuve E. 2012. Bacterial persistence and toxin-antitoxin Loci. *Annu Rev Microbiol* **66**: 103–123.

Goldberg AL, Moerschell RP, Chung CH, Maurizi MR. 1994. ATP-dependent protease La (lon) from Escherichia coli. *Methods in Enzymology* **244**: 350–375.

Gonzalez M, Frank EG, Levine AS, Woodgate R. 1998. Lon-mediated proteolysis of the Escherichia coli UmuD mutagenesis protein: in vitro degradation and identification of residues required for proteolysis. *Genes Dev* **12**: 3889–3899.

Gora KG, Cantin A, Wohlever M, Joshi KK, Perchuk BS, Chien P, Laub MT. 2013. Regulated proteolysis of a transcription factor complex is critical to cell cycle progression in Caulobacter crescentus. *Mol Microbiol.*

Gottesman S, Halpern E, Trisler P. 1981. Role of sulA and sulB in filamentation by lon mutants of Escherichia coli K-12. *J Bacteriol* **148**: 265–273.

Gottesman S, Wickner S, Maurizi MR. 1997. Protein quality control: triage by chaperones and proteases. *Genes Dev* **11**: 815–823.

Griffith KL, Shah IM, Wolf RE. 2004. Proteolytic degradation of Escherichia coli transcription activators SoxS and MarA as the mechanism for reversing the induction of the superoxide (SoxRS) and multiple antibiotic resistance (Mar) regulons. *Mol Microbiol* **51**: 1801–1816.

Grossman AD, Straus DB, Walter WA, Gross CA. 1987. Sigma 32 synthesis can regulate the synthesis of heat shock proteins in Escherichia coli. *Genes Dev* **1**:

179–184.

Gur E. 2013. The Lon AAA+ Protease. *Subcell Biochem* **66**: 35–51.

Gur E, Sauer RT. 2009. Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine. *Proceedings of the National Academy of Sciences* **106**: 18503–18508.

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

Hanson PI, Whiteheart SW. 2005. AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol* **6**: 519–529.

Hartl FU, Bracher A, Hayer-Hartl M. 2011. Molecular chaperones in protein folding and proteostasis. *Nature* **475**: 324–332.

Hartl FU, Hayer-Hartl M. 2009. Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol* **16**: 574–581.

Heim R, Prasher DC, Tsien RY. 1994. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci USA* **91**: 12501–12504.

Heuveling J, Possling A, Hengge R. 2008. A role for Lon protease in the control of the acid resistance genes of Escherichia coli. *Mol Microbiol* **69**: 534–547.

Higashitani A, Ishii Y, Kato Y, Koriuchi K. 1997. Functional dissection of a cell-division inhibitor, SulA, of Escherichia coli and its negative regulation by Lon. *Mol Gen Genet* **254**: 351–357.

Hilliard JJ, Simon LD, Van Melderen L, Maurizi MR. 1998. PinA inhibits ATP hydrolysis and energy-dependent protein degradation by Lon protease. *J Biol Chem* **273**: 524–527.

Horwich AL, Apetri AC, Fenton WA. 2009. The GroEL/GroES cis cavity as a passive anti-aggregation device. *FEBS Lett* **583**: 2654–2662.

Hou JY, Sauer RT, Baker TA. 2008. Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. *Nat Struct Mol Biol* **15**: 288–294.

Im YJ, Na Y, Kang GB, Rho S-H, Kim M-K, Lee JH, Chung CH, Eom SH. 2004. The active site of a lon protease from Methanococcus jannaschii distinctly differs from the canonical catalytic Dyad of Lon proteases. *J Biol Chem* **279**: 53451–53457.

Ingmer H, Brøndsted L. 2009. Proteases in bacterial pathogenesis. *Res Microbiol* **160**: 704–710.

Iniesta AA, McGrath PT, Reisenauer A, McAdams HH, Shapiro L. 2006. A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. *Proc Natl Acad Sci USA* **103**: 10935–10940.

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

Jiao W, Qian M, Li P, Zhao L, Chang Z. 2005. The essential role of the flexible termini in the temperature-responsiveness of the oligomeric state and chaperone-like activity for the polydisperse small heat shock protein IbpB from Escherichia coli. *J Mol Biol* **347**: 871–884.

Johansson J, Uhlin BE. 1999. Differential protease-mediated turnover of H-NS and StpA revealed by a mutation altering protein stability and stationary-phase survival of Escherichia coli. *Proc Natl Acad Sci USA* **96**: 10776–10781.

Kent KP, Childs W, Boxer SG. 2008. Deconstructing green fluorescent protein. *J Am Chem Soc* **130**: 9664–9665.

Kim S, Grant RA, Sauer RT. 2011. Covalent linkage of distinct substrate degrons controls assembly and disassembly of DegP proteolytic cages. *Cell* **145**: 67–78.

Kim S, Sauer RT. 2012. Cage assembly of DegP protease is not required for substrate-dependent regulation of proteolytic activity or high-temperature cell survival. *Proceedings of the National Academy of Sciences* **109**: 7263–7268.

King RW, Deshaies RJ, Peters JM, Kirschner MW. 1996. How proteolysis drives the cell cycle. *Science* **274**: 1652–1659.

Krojer T, Sawa J, Huber R, Clausen T. 2010. HtrA proteases have a conserved activation mechanism that can be triggered by distinct molecular cues. *Nat Struct Mol Biol* **17**: 844–852.

Kubik S, Wegrzyn K, Pierechod M, Konieczny I. 2011. Opposing effects of DNA on proteolysis of a replication initiator. *Nucleic acids research.*

Kuroda A, Nomura K, Ohtomo R, Kato J, Ikeda T, Takiguchi N, Ohtake H, Kornberg A. 2001. Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in E. coli. *Science* **293**: 705–708.

Laufen T, Mayer MP, Beisel C, Klostermeier D, Mogk A, Reinstein J, Bukau B. 1999. Mechanism of regulation of hsp70 chaperones by DnaJ cochaperones. *Proc Natl Acad Sci USA* **96**: 5452–5457.

Lee AY-L, Chen M-Y, Wu S-H. 2004a. Identification of a gene encoding Lon protease from Brevibacillus thermoruber WR-249 and biochemical characterization of its

thermostable recombinant enzyme. *Eur J Biochem* **271**: 834–844.

Lee AY-L, Hsu C-H, Wu S-H. 2004b. Functional domains of Brevibacillus thermoruber lon protease for oligomerization and DNA binding: role of N-terminal and sensor and substrate discrimination domains. *J Biol Chem* **279**: 34903–34912.

Lee I, Suzuki CK. 2008. Functional mechanics of the ATP-dependent Lon protease-lessons from endogenous protein and synthetic peptide substrates. *Biochim Biophys Acta* **1784**: 727–735.

Leung E, Datti A, Cossette M, Goodreid J, McCaw SE, Mah M, Nakhamchik A, Ogata K, Bakkouri El M, Cheng Y-Q, et al. 2011. Activators of Cylindrical Proteases as Antimicrobials: Identification and Development of Small Molecule Activators of ClpP Protease. *Chem Biol* **18**: 1167–1178.

Li M, Gustchina A, Rasulova FS, Melnikov EE, Maurizi MR, Rotanova TV, Dauter Z, Wlodawer A. 2010. Structure of the N-terminal fragment of Escherichia coli Lon protease. *Acta Crystallogr D Biol Crystallogr* **66**: 865–873.

Li M, Rasulova F, Melnikov EE, Rotanova TV, Gustchina A, Maurizi MR, Wlodawer A. 2005. Crystal structure of the N-terminal domain of E. coli Lon protease. *Protein Sci* **14**: 2895–2900.

Liao J-H, Kuo C-I, Huang Y-Y, Lin Y-C, Lin Y-C, Yang C-Y, Wu W-L, Chang W-H, Liaw Y-C, Lin L-H, et al. 2012. A Lon-Like Protease with No ATP-Powered Unfolding Activity ed. V.N. Uversky. *PLoS ONE* **7**: e40226.

Liao J-H, Lin Y-C, Hsu J, Lee AY-L, Chen T-A, Hsu C-H, Chir J-L, Hua K-F, Wu T-H, Hong L-J, et al. 2010. Binding and cleavage of E. coli HUbeta by the E. coli Lon protease. *Biophys J* **98**: 129–137.

Lu B, Lee J, Nie X, Li M, Morozov YI, Venkatesh S, Bogenhagen DF, Temiakov D, Suzuki CK. 2012. Phosphorylation of Human TFAM in Mitochondria Impairs DNA Binding and Promotes Degradation by the AAA(+) Lon Protease. *Molecular Cell*.

Luce K, Osiewacz HD. 2009. Increasing organismal healthspan by enhancing mitochondrial protein quality control. *Nat Cell Biol* **11**: 852–858.

Lund PA. 2009. Multiple chaperonins in bacteria â" why so many? *FEMS Microbiology Reviews* **33**: 785–800.

Luo J, Solimini NL, Elledge SJ. 2009. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell* **136**: 823–837.

Maisonneuve E, Ezraty B, Dukan S. 2008. Protein aggregates: an aging factor

involved in cell death. *J Bacteriol* **190**: 6070–6075.

Maisonneuve E, Shakespeare LJ, Jørgensen MG, Gerdes K. 2011. Bacterial persistence by RNA endonucleases. *Proceedings of the National Academy of Sciences* **108**: 13206–13211.

Markovitz A. 1964. REGULATORY MECHANISMS FOR SYNTHESIS OF CAPSULAR POLYSACCHARIDE IN MUCOID MUTANTS OF ESCHERICHIA COLI K12. *Proc Natl Acad Sci USA* **51**: 239–246.

Martin A, Baker TA, Sauer RT. 2008a. Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Molecular Cell* **29**: 441–450.

Martin A, Baker TA, Sauer RT. 2008b. Protein unfolding by a AAA+ protease is dependent on ATP-hydrolysis rates and substrate energy landscapes. *Nat Struct Mol Biol* **15**: 139–145.

Matsushima Y, Goto Y-I, Kaguni LS. 2010. Mitochondrial Lon protease regulates mitochondrial DNA copy number and transcription by selective degradation of mitochondrial transcription factor A (TFAM). *Proceedings of the National Academy of Sciences* **107**: 18410–18415.

Mizusawa S, Gottesman S. 1983. Protein degradation in Escherichia coli: the lon gene controls the stability of sulA protein. *Proc Natl Acad Sci USA* **80**: 358–362.

Mogk A, Schmidt R, Bukau B. 2007. The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol* **17**: 165–172.

Moore SD, Sauer RT. 2005. Ribosome rescue: tmRNA tagging activity and capacity in Escherichia coli. *Mol Microbiol* **58**: 456–466.

Moore SD, Sauer RT. 2007. The tmRNA system for translational surveillance and ribosome rescue. *Annu Rev Biochem* **76**: 101–124.

Moran U, Phillips R, Milo R. 2010. SnapShot: key numbers in biology. *Cell* **141**: 1262–1262.e1.

Müller S, Dennemärker J, Reinheckel T. 2012. Specific functions of lysosomal proteases in endocytic and autophagic pathways. *Biochim Biophys Acta* **1824**: 34–43.

Nager AR, Baker TA, Sauer RT. 2011. Stepwise unfolding of a β barrel protein by the AAA+ ClpXP protease. *J Mol Biol* **413**: 4–16.

Narberhaus F. 2002. Alpha-crystallin-type heat shock proteins: socializing minichaperones in the context of a multichaperone network. *Microbiol Mol Biol*

Rev **66**: 64–93; table of contents.

Nomura K, Kato J, Takiguchi N, Ohtake H, Kuroda A. 2004. Effects of inorganic polyphosphate on the proteolytic and DNA-binding activities of Lon in Escherichia coli. *J Biol Chem* **279**: 34406–34410.

Olzscha H, Schermann SM, Woerner AC, Pinkert S, Hecht MH, Tartaglia GG, Vendruscolo M, Hayer-Hartl M, Hartl FU, Vabulas RM. 2011. Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* **144**: 67–78.

Ormö M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. 1996. Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273**: 1392–1395.

Pande A, Pande J, Asherie N, Lomakin A, Ogun O, King J, Benedek GB. 2001. Crystal cataracts: human genetic cataract caused by protein crystallization. *Proc Natl Acad Sci USA* **98**: 6116–6120.

Pechmann S, Willmund F, Frydman J. 2013. The ribosome as a hub for protein quality control. *Molecular Cell* **49**: 411–421.

Randow F, Lehner PJ. 2009. Viral avoidance and exploitation of the ubiquitin system. *Nat Cell Biol* **11**: 527–534.

Ratajczak E, Zietkiewicz S, Liberek K. 2009. Distinct activities of Escherichia coli small heat shock proteins IbpA and IbpB promote efficient protein disaggregation. *J Mol Biol* **386**: 178–189.

Rep M, van Dijl JM, Suda K, Schatz G, Grivell LA, Suzuki CK. 1996. Promotion of mitochondrial membrane complex assembly by a proteolytically inactive yeast Lon. *Science* **274**: 103–106.

Robertson GT, Kovach ME, Allen CA, Ficht TA, Roop RM. 2000. The Brucella abortus Lon functions as a generalized stress response protease and is required for wild-type virulence in BALB/c mice. *Mol Microbiol* **35**: 577–588.

Roman-Hernandez G, Hou JY, Grant RA, Sauer RT, Baker TA. 2011. The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. *Molecular Cell* **43**: 217–228.

Rosenzweig R, Moradi S, Zarrine-Afsar A, Glover JR, Kay LE. 2013. Unraveling the Mechanism of Protein Disaggregation Through a ClpB-DnaK Interaction. *Science*.

Rotanova TV, Botos I, Melnikov EE, Rasulova F, Gustchina A, Maurizi MR, Wlodawer A. 2006. Slicing a protease: structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains. *Protein Sci* **15**: 1815–

1828.

Rotanova TV, Melnikov EE, Khalatova AG, Makhovskaya OV, Botos I, Wlodawer A, Gustchina A. 2004. Classification of ATP-dependent proteases Lon and comparison of the active sites of their proteolytic domains. *Eur J Biochem* **271**: 4865–4871.

Roudiak SG, Shrader TE. 1998. Functional role of the N-terminal region of the Lon protease from Mycobacterium smegmatis. *Biochemistry* **37**: 11255–11263.

Rudyak SG, Shrader TE. 2000. Polypeptide stimulators of the Ms-Lon protease. *Protein Sci* **9**: 1810–1817.

Sahi C, Kominek J, Ziegelhoffer T, Yu HY, Baranowski M, Marszalek J, Craig EA. 2013. Sequential Duplications of an Ancient Member of the DnaJ-Family Expanded the Functional Chaperone Network in the Eukaryotic Cytosol. *Mol Biol Evol.*

Sauer RT, Baker TA. 2011. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **80**: 587–612.

Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, Thompson MJ, Balbirnie M, Wiltzius JJW, McFarlane HT, et al. 2007. Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* **447**: 453–457.

Shah IM, Wolf RE. 2006. Sequence requirements for Lon-dependent degradation of the Escherichia coli transcription activator SoxS: identification of the SoxS residues critical to proteolysis and specific inhibition of in vitro degradation by a peptide comprised of the N-terminal 21 amino acid residues. *J Mol Biol* **357**: 718–731.

Sharma SK, De los Rios P, Christen P, Lustig A, Goloubinoff P. 2010. The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. *Nat Chem Biol* **6**: 914–920.

Sohn J, Grant RA, Sauer RT. 2007. Allosteric activation of DegS, a stress sensor PDZ protease. *Cell* **131**: 572–583.

Sohn J, Grant RA, Sauer RT. 2009. OMP Peptides Activate the DegS Stress-Sensor Protease by a Relief of Inhibition Mechanism. *Structure* **17**: 1411–1421.

Sousa MC, Kessler BM, Overkleeft HS, McKay DB. 2002. Crystal structure of HslUV complexed with a vinyl sulfone inhibitor: corroboration of a proposed mechanism of allosteric activation of HslV by HslU. *J Mol Biol* **318**: 779–785.

Stoner-Ma D, Melief EH, Nappa J, Ronayne KL, Tonge PJ, Meech SR. 2006. Proton relay reaction in green fluorescent protein (GFP): Polarization-resolved ultrafast

vibrational spectroscopy of isotopically edited GFP. *J Phys Chem B* **110**: 22009–22018.

Sun Y, MacRae TH. 2005. Small heat shock proteins: molecular structure and chaperone function. *Cell Mol Life Sci* **62**: 2460–2476.

Suzuki CK, Rep M, van Dijl JM, Suda K, Grivell LA, Schatz G. 1997. ATP-dependent proteases that also chaperone protein biogenesis. *Trends Biochem Sci* **22**: 118–123.

Tomoyasu T, Mogk A, Langen H, Goloubinoff P, Bukau B. 2001. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the Escherichia coli cytosol. *Mol Microbiol* **40**: 397–413.

Torres-Cabassa AS, Gottesman S. 1987. Capsule synthesis in Escherichia coli K-12 is regulated by proteolysis. *J Bacteriol* **169**: 981–989.

Vabulas RM, Raychaudhuri S, Hayer-Hartl M, Hartl FU. 2010. Protein folding in the cytoplasm and the heat shock response. *Cold Spring Harb Perspect Biol* **2**: a004390.

Van Melderen L, Gottesman S. 1999. Substrate sequestration by a proteolytically inactive Lon mutant. *Proc Natl Acad Sci USA* **96**: 6064–6071.

Vasilyeva OV, Kolygo KB, Leonova YF, Potapenko NA, Ovchinnikova TV. 2002. Domain structure and ATP-induced conformational changes in Escherichia coli protease Lon revealed by limited proteolysis and autolysis. *FEBS Lett* **526**: 66–70.

Veinger L, Diamant S, Buchner J, Goloubinoff P. 1998. The small heat-shock protein IbpB from Escherichia coli stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. *J Biol Chem* **273**: 11032–11037.

Venkatesh S, Lee J, Singh K, Lee I, Suzuki CK. 2011. Multitasking in the mitochondrion by the ATP-dependent Lon protease. *Biochim Biophys Acta.*

Wah DA, Levchenko I, Baker TA, Sauer RT. 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.

Walker JE, Saraste M, Runswick MJ, Gay NJ. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1**: 945–951.

Waxman L, Goldberg AL. 1982. Protease La from Escherichia coli hydrolyzes ATP and proteins in a linked fashion. *Proc Natl Acad Sci USA* **79**: 4883–4887.

Wendler P, Ciniawsky S, Kock M, Kube S. 2011. Structure and function of the AAA+ nucleotide binding pocket. *Biochim Biophys Acta.*

Winkler J, Seybert A, König L, Pruggnaller S, Haselmann U, Sourjik V, Weiss M, Frangakis AS, Mogk A, Bukau B. 2010. Quantitative and spatio-temporal features of protein aggregation in Escherichia coli and consequences on protein quality control and cellular ageing. *EMBO J.*

Winkler J, Tyedmers J, Bukau B, Mogk A. 2012. Chaperone networks in protein disaggregation and prion propagation. *J Struct Biol.*

Winklhofer KF, Tatzelt J, Haass C. 2008. The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases. *EMBO J* **27**: 336–349.

Wright R, Stephens C, Zweiger G, Shapiro L, Alley MR. 1996. Caulobacter Lon protease has a critical role in cell-cycle control of DNA methylation. *Genes Dev* **10**: 1532–1542.

Chapter 2

Engineering fluorescent protein substrates for the AAA+ Lon protease

This work was originally published as Matthew L.Wohlever, Andrew R. Nager, Tania A. Baker, and Robert T. Sauer. 2013 *PEDS* 26: 299-305.

M.L.W. did all experiments and wrote initial manuscript. A.R.N. cloned constructs and aided in experimental design.

Abstract

AAA+ proteases, such as *E. coli* Lon, recognize protein substrates by binding to specific peptide degrons and then unfold and translocate the protein into an internal degradation chamber for proteolysis. For some AAA+ proteases, attaching specific degrons to the N- or C-terminus of GFP generates useful substrates, whose unfolding and degradation can be monitored by loss of fluorescence, but Lon fails to degrade appropriately tagged GFP variants at a significant rate. Here, we demonstrate that Lon catalyzes robust unfolding and degradation of circularly permuted variants of GFP with a β20 degron appended to the N terminus or a sul20 degron appended to the C terminus. Lon degradation of non-permuted GFP-sul20 is very slow, in part because the enzyme cannot efficiently extract the degron-proximal C-terminal b strand to initiate denaturation. The circularly permuted GFP substrates described here allow convenient high-throughput assays of the kinetics of Lon degradation *in vitro* and also permit assays of Lon proteolysis *in vivo*.

Introduction

AAA+ proteases are molecular machines that convert the chemical energy of ATP binding and hydrolysis into mechanical work that is used to unfold and translocate a protein substrate through the axial pore of a hexameric ring and into a sequestered chamber for degradation (Baker and Sauer 2006; Sauer and Baker 2011). Architecturally, the Lon hexamer is one of the simplest AAA+ proteases, as its family-specific N-terminal domain, its AAA+ ATPase module, and its peptidase domain are all connected in a single polypeptide chain. In bacteria, archaea, and endosymbiotic organelles, Lon plays important roles in protein quality control by degrading misfolded or damaged proteins and also degrades native proteins that are no longer needed or must be removed for regulatory purposes (Torres-Cabassa and Gottesman 1987; Gonzalez et al. 1998; Striebel et al. 2009). Lon is also a promising therapeutic target, because it is required for the virulence of several pathogenic bacteria, and inhibition of the human mitochondrial enzyme causes apoptosis of lymphoma cells (Robertson et al. 2000; Ingmer and Brøndsted 2009; Yang et al. 2011; Bernstein et al. 2012).

Like all AAA+ proteases, Lon recognizes protein substrates via degrons or degradation tags, which can be as simple as specific sequences at the N or C terminus (Sauer and Baker 2011). Several degrons for *Escherichia coli* Lon have been identified, including a 20 amino-acid sequence within b-galactosidase (called β20), which becomes accessible to Lon only in the unfolded protein, and a sequence from the C-terminus of the cell-division inhibitor SulA (called sul20) (Higashitani et al. 1997; Gonzalez et al. 1998; Gur and Sauer 2008). Attachment of the sul20 and

β20 sequences to model substrates leads to their degradation by *E. coli* Lon. Interestingly, proteins bearing a C-terminal sul20 degron are degraded with a higher maximal velocity than otherwise identical substrates tagged with a C-terminal β20 degron, suggesting that degron identity regulates Lon activity in some fashion (Gur and Sauer 2009). The β20 sequence also functions as an N-terminal or internal degradation tag for Lon. Although, it is not known how Lon binds the sul20 or β20 degrons, studies with other AAA+ proteases have shown that some degrons bind in the axial pore of the hexameric ring, where they are engaged by the translocation machinery of the enzyme. Subsequent translocation of the degradation tag creates a pulling force when the attached native protein cannot enter the narrow axial channel, eventually leading to unfolding of the protein substrate. Because unfolding is an inherently mechanical process, its rate depends both on the pulling force that the enzyme can exert and on the stability of the local protein structure adjacent to the degradation tag (Lee et al. 2001; Kenniston et al. 2003; 2004).

Green fluorescent protein (GFP) has been an extremely valuable substrate for studies of many AAA+ proteases, as the fluorescent chromophore formed by cyclization and oxidation of residues 65-67 is protected from solvent quenching by the β-barrel structure of the native protein (Heim et al. 1994; Ormö et al. 1996). Enzymatic unfolding and degradation exposes the chromophore to solvent, quenching its fluorescence. Unfortunately, Lon degrades C- and N-terminally tagged variants of GFP extremely slowly (Choy et al. 2007; Gur and Sauer 2008; Gur et al. 2012), which has hindered mechanistic studies *in vitro* and *in vivo*. Here, we

demonstrate that Lon can unfold and degrade circularly permuted variants of GFP with appropriate degrons at either the N or C terminus. These substrates allow rapid, high-throughput kinetic analysis of Lon degradation *in vitro* and can also be used to monitor Lon degradation *in vivo*. We also show that Lon cannot efficiently extract the degron tagged C-terminal strand from the β barrel of non permuted GFP, a substrate that resists Lon degradation.

Materials and Methods

Protein cloning, expression, and purification

Variants of superfolder GFP (Pédelacq et al. 2006) were cloned into a pCOLADuet-1 plasmid vector with an N-terminal $MGSH_6$SLEVLFQGPGS tag that included a PreScission protease site (Nager et al. 2011). For circular permutations, the normal N and C-termini of GFP were connected by a GGTGGS linker (Reeder et al. 2010). The sul20 degron (ASSHATRQLSGLKIHSNLYH) was added to the C-terminus of superfolder GFP variants and the β20 degron (QLRSLNGEWRFAWFPAPEAV) was inserted after residues MG in the N-terminal tag by PCR cloning. The variant with the thrombin-cleavage site contained the sequence GGTEGS**LVPRGS**GESGGS (thrombin site in bold; flanking sequences introduced to limit steric hindrance) inserted between NEK^{235} and ^{236}RDH, where the residue numbers refer to the sequence of superfolder GFP-sul20. The underlined residues in the GNILGHKLEYNLEASSHAT sequence of cp6-sul20 were deleted in the cp6-Δ8-sul20 variant.

E. coli Lon was expressed from a pBAD33 overexpression vector as described (Gur and Sauer 2009). Briefly, cells were grown at 37° C until OD_{600} ~1, induced with 0.2% arabinose at 37 °C for 3.5 h, harvested, and resuspended in buffer A (100 mM potassium phosphate [pH 6.5], 1 mM DTT, 1 mM EDTA, and 10% glycerol). After lysis by sonication, insoluble material was removed by high-speed centrifugation, and 2 µL of benzonase (250 U/µL, Sigma) was added to the supernatant, which was incubated on ice for 20 min. The lysate was mixed with P11 phosphocellulose resin (Whatman) equilibrated in buffer A, and the resin was washed twice with buffer A and twice with an otherwise identical buffer containing 200 mM potassium phosphate [pH 6.5]. Lon was eluted from the P11 resin with buffer B (400 mM potassium phosphate [pH 6.5], 1 mM DTT, 1 mM EDTA, and 10% glycerol), and then chromatographed on a S300 size-exclusion column equilibrated in 50 mM HEPES [pH 7.5], 2 M NaCl, and 1 mM DTT. The peak fractions of Lon from this column were >95% pure as assayed by SDS-PAGE and were buffer exchanged into storage buffer (50 mM HEPES [pH 7.5], 150 mM NaCl, 10 µM EDTA, 1 mM DTT, 10% glycerol) and frozen at -80 °C.

Superfolder GFP variants were expressed and purified largely as described (Nager et al. 2011). Cells harboring overexpression plasmids were grown at 37° C to OD_{600} ~1, induced with 0.5 mM IPTG at room temperature for 3.5 h, harvested by centrifugation, and lysed by sonication. After an initial step of Ni^{++}-NTA (Qiagen) affinity chromatography, material was bound to a MonoQ column (GE Healthcare) equilibrated in 25 mM Tris [pH 8.0], 25 mM NaCl, 1 mM DTT, and 10 µM EDTA, and eluted using a linear gradient in the same buffer to 500 mM NaCl over 20 column

volumes. The peak fractions of degron-tagged superfolder GFP variants were >90% pure as assayed by SDS-PAGE and were buffer exchanged into 50 mM HEPES [pH 7.5], 150 mM NaCl, 10 μM EDTA, 1 mM DTT, and 10% glycerol and frozen at –80 °C.

Biochemical assays

Unless noted, biochemical assays were performed at 37 °C in 25 mM Tris [pH 8.0], 100 mM KCl, 10 mM $MgCl_2$. ATPase assays, monitored by absorbance at 340 nm, also contained 5 mM DTT, 2 mM ATP, 1 mM NADH, lactate dehydrogenase (10 U/mL) an ATP-regeneration system– 20 mM phosphoenolpyruvate (Sigma) and 10 U/ mL of rabbit muscle pyruvate kinase (Sigma) – and $MgCl_2$, which had been warmed to 37 °C, was added last to initiate the reaction (Nørby 1988). Degradation assays contained supplemental 1 mM DTT, 2 mM ATP, and an ATP-regeneration system, in addition to *E. coli* Lon and protein substrates. For degradation reactions assayed by fluorescence, reaction mixtures without ATP were incubated in 96-well flat bottom, ½ area plates (Corning) until there was no change in GFP fluorescence caused by thermal equilibration (~15 min), and degradation was initiated by addition of ATP. For degradation assays monitored by SDS PAGE, 10 μL aliquots were taken at the specified time points and quenched by addition of 3.3 μL of 8% SDS, 250 mM Tris pH 6.8, 40% glycerol, 160 mM DTT, and 0.05% bromophenol blue.

To monitor Lon-mediated tail clipping, cp6-sul20 (50 μM) was incubated with Lon (1 μM hexamer) for 2 h at 37 °C, the sample was loaded onto a nickel spin column (Qiagen), and the column was washed with 50 mM HEPES [pH 7.5], 150 mM NaCl, 20 mM imidazole, 10 μM EDTA, 1 mM DTT, and 10% glycerol and then eluted

with a similar buffer containing 250 mM imidazole. The eluted material from this experiment and purified substrate proteins were submitted for LC-MS analysis using a QSTAR Elite quadrupole-time-of-flight mass spectrometer. Approximately 3 picomoles of sample were loaded onto a reversed phase protein trap, and the sample was desalted on-line and eluted isocratically. Deconvolution of the electrospray data to generate molecular weight spectra was performed with the BioAnalyst software included with the QSTAR Elite data system.

To prepare the split GFP-sul20 substrate, treatment of the sample with 10 U/mL of thrombin (GE Healthcare) for 2 h at 37° C in storage buffer resulted in ~95% cleavage as assayed by SDS-PAGE, and the split substrate was immediately used for Lon extraction/degradation assays. For the SDS-PAGE assay of degradation, quenched time points were electrophoresed on a 4-20% gradient gel, which was subsequently stained with Coomassie blue. For strand-extraction assays monitored by fluorescence, reactions were performed in a fluorescent plate reader as described above.

Degradation in vivo

The *lon*$^{-}$ *E. coli* strain ER2566 (New England Biolabs) was transformed with a pCOLAduet vector encoding cp6-sul20 under transcriptional control of a T7 promoter and/or with pBAD33 encoding wild-type Lon under transcriptional control of the *araC* promoter. Cells containing one or both plasmids were grown in M9 minimal media with appropriate antibiotics (10 μg/mL chloramphenicol and/or 50 μg/mL kanamycin) at 37 °C until OD_{600} ~0.3, 1 mM IPTG was added to induce

cp6-sul20, and the temperature was dropped to 30° C to allow for more efficient chromophore maturation. After 2 h of CP6-sul20 expression, Lon expression was induced by adding arabinose to a final concentration of 0.2%. At this time, 100 µL of culture was transferred to a corning 96-well ½ area flat bottom plate in a SpectraMax M5 plate reader and fluorescence (467 nm excitation; 511 nm emission) was measured at 37° C with shaking of the plate between measurements. Assays were performed in triplicate.

For FACS analysis, *E. coli* strains W3110 (*lon*$^+$) and W3110 *lon::kan* (*lon*$^-$) were transformed with pCOLAduet encoding cp6-sul20, and grown at 37 °C. At OD_{600} ~0.3, 1 mM IPTG was added to induce cp6-sul20 expression and the temperature was lowered to 30° C to facilitate more efficient chromophore maturation. After 90 min, a 100 µL aliquot of cells was taken, pelleted in a microcentrifuge, and then resuspended in 70% cold ethanol to fix the cells. Samples were stored at −20° C until FACS analysis, and then pelleted, resuspended in phosphate buffered saline (Boston Bioproducts) at a final concentration of 10^7 cells/mL, and subjected to FACS on a BD FACScan instrument. Data were collected as phycoerythrin fluorescence versus GFP fluorescence to eliminate auto-fluorescence. ER2566 cells without GFP were used as a negative control to determine the lower limit for the GFP-positive gate.

Results

Lon degrades circularly permuted GFP substrates

For all of the studies presented here, substrates were variants of superfolder GFP, some of which were circularly permuted to place b-strand 6 (cp6) or b-strand 7 (cp7) at the C-terminal end of the 11-stranded b barrel (Figure 2.1A). Each variant had an N-terminal His_6 tag and either a β20 degron preceding the His_6 sequence, a β20 degron at the C terminus, or a sul20 degron at the C-terminus of the protein. Because prior studies showed that circular permutation of superfolder GFP enhanced degradation by ClpXP (Nager et al. 2011), we sought to determine if circular permutation would also facilitate degradation of GFP by *E. coli* Lon.

We constructed and purified nine substrates: β20-GFP, β20-cp6, β20-cp7, GFP-β20, cp6-β20, cp7-β20, GFP-sul20, cp6-sul20, and cp7-sul20. As assayed by loss of native fluorescence, 0.3 μM Lon catalyzed robust degradation of 10 μM concentrations of β20-cp6, β20-cp7, cp6-sul20, cp7-sul20 and slow degradation of the same concentrations of cp6-β20 and cp7-β20 (Figure 2.1B). Under the same assay conditions, we observed no change in the fluorescence of β20-GFP, GFP-β20, or GFP-sul20 (not shown), indicating that these non-permuted proteins are extremely poor substrates for Lon degradation. For the four best substrates, Lon degradation of β20-cp6 and β20-cp7 proceeded to a greater extent than degradation of cp6-sul20 and cp7-sul20, which effectively ceased with ~30% of the initial fluorescence remaining. An SDS-PAGE assay also showed incomplete degradation of cp6-sul20 (Figure 2.1C), which was not improved by adding additional ATP, additional Lon, or overnight incubation (not shown). We return to the issue of incomplete degradation below.

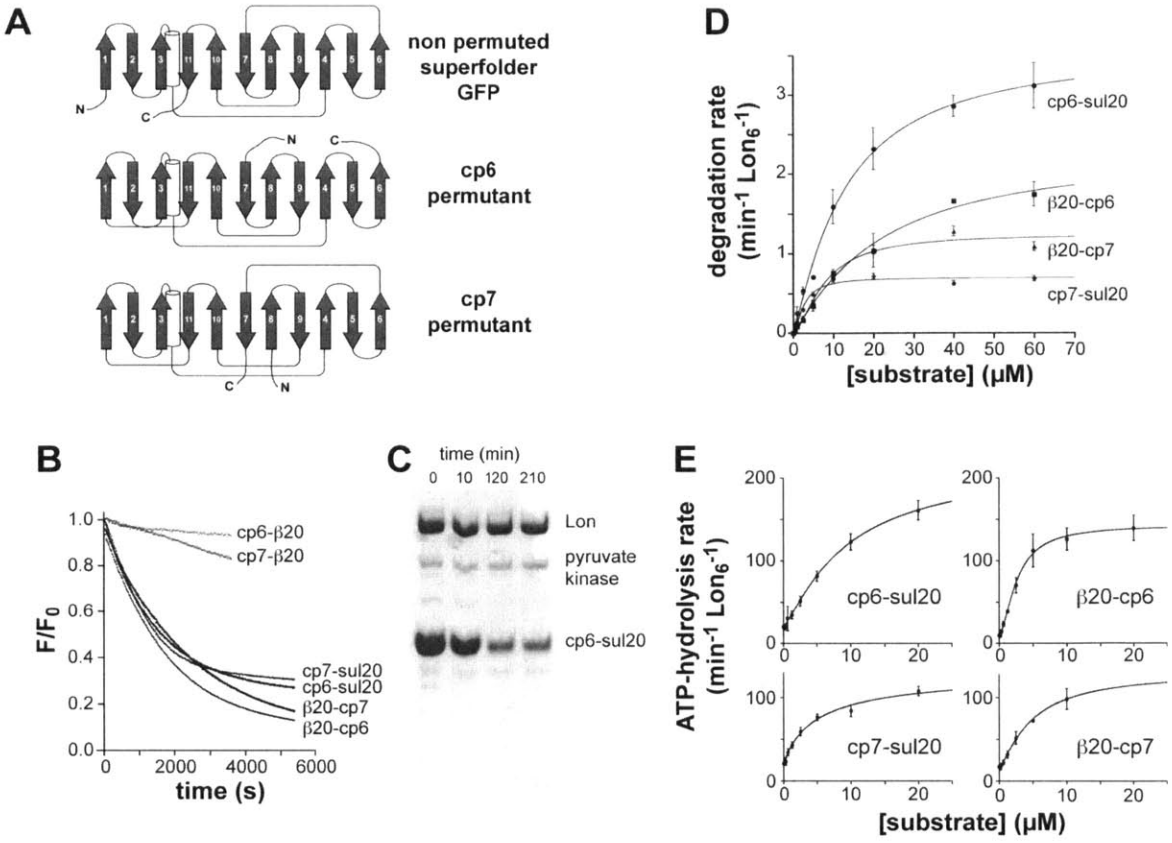

Figure 2.1 Design and characterization of GFP circular permutations
(A) Cartoon representation of the order of secondary structure elements in superfolder GFP and the circularly permuted cp6 and cp7 variants. The N terminus (N) and C terminus (C) of each protein is marked. (B) Proteolysis of degron-tagged variants of the cp6 and cp7 substrates (10 μM) by Lon_6 (0.3 μM) was monitored by changes in fluorescence (excitation 467 nm; emission 511 nm). Fluorescence values were normalized by dividing by the time-zero fluorescence. (C) SDS-PAGE assay of degradation of cp6-sul20 (20 μM) by Lon_6 (0.6 μM). Following electrophoresis, the gel was stained with Coomassie Blue. (D) Substrate dependence of Lon_6 (0.15 μM) degradation of the cp6-sul20, cp7-sul20, b20-cp6, and b20-cp7 proteins. The solid lines are fits to the Hill form of the Michaelis-Menten equation (rate = $V_{max}/(1+(K_M/[S])^n)$. Values are shown as means ± SEM ($N = 3$). (E) Dependence on the concentration of substrate proteins of the rate of hydrolysis of ATP (2 mM) by Lon_6 (0.15 μM). The solid lines are fits to a Hill equation, rate = basal + amp/$(1+(K_M/[S])^n)$, where V_{max} = basal + amp. Values are shown as means ± SEM ($N \geq 3$).

To determine steady-state kinetic parameters for Lon degradation of the four substrates that were degraded most rapidly, we determined initial rates as a function of the concentration of the β20-cp6, β20-cp7, cp6-sul20, or cp7-sul20

substrates and fit them to the Hill form of the Michaelis-Menten equation (Figure 2.1D). K_M, V_{max}, and n (the Hill constant) values for each substrate are listed in Table 1. In terms of V_{max} values, the substrate rank from highest to lowest was cp6-sul20 > β20-cp6 > β20-cp7 > cp7-sul20 with an ~5-fold difference between the fastest (~3.6 min^{-1} Lon_6^{-1}) and slowest (~0.70 min^{-1} Lon_6^{-1}) rates. For the substrates with C-terminal degrons, the GuHCl concentrations required for 50% unfolding were ~1.6 M for cp6-GFP-sul20, ~1.8 M for cp7-GFP-sul20, and ~3.8 M for GFP-sul20 (not shown). Thus, the least stable substrates are most susceptible to Lon degradation. The K_M's for cp6 or cp7 substrates with the same degron varied, in some cases by as much as 4.5-fold (Table 2.1), indicating that factors in addition to simple degron binding by Lon determine this kinetic parameter.

We also varied the concentration of each GFP variant and determined the effects on the steady-state rate of ATP hydrolysis by Lon (Figure 2.1E, Table 2.1). For the β20-cp6 substrate, the concentration at which half-maximal ATPase stimulation was observed was substantially lower than the K_M for degradation, as previously observed for some other Lon substrates (Gur and Sauer 2009). The energetic efficiency of Lon degradation of the circularly permuted substrates at saturating concentrations varied, with degradation of cp6-sul20 and β20-cp6 requiring hydrolysis of an average of ~60 ATPs, β20-cp7 requiring ~110 ATPs, and cp7-sul20 requiring ~190 ATPs (Table 2.1). Because the energetic cost of translocation of the sul20-tagged substrates should be similar to each other, the ~3-fold higher energetic cost of degrading cp7-sul20 compared to cp6-sul20 is likely to reflect an increased energetic cost of Lon unfolding. There was no clear correlation

between maximum ATPase stimulation and degron identity, unlike the situation observed previously for unfolded substrates (Gur and Sauer 2009), but this result could simply reflect a more complicated dependence on a combination of degron identity and other features of our substrates.

substrate	proteolysis V_{max} (min^{-1} Lon_6^{-1})	proteolysis apparent K_M (μM)	proteolysis Hill constant	ATPase V_{max} (min^{-1} Lon_6^{-1})	ATPase apparent K_M (μM)	ATPase Hill constant	ATP per substrate
cp6-sul20	3.6 ± 0.2	13 ± 2	1.2 ± 0.1	217 ± 16	9.3 ± 1.3	1.2 ± 0.1	60
β20-cp6	2.3 ± 0.3	21 ± 6	1.2 ± 0.2	143 ± 5	2.6 ± 0.2	1.7 ± 0.1	62
cp7-sul20	0.70 ± 0.03	2.8 ± 0.3	1.6 ± 0.3	130 ± 22	5.2 ± 2.4	0.9 ± 0.2	186
β20-cp7	1.2 ± 0.1	8 ± 1	1.6 ± 0.3	130 ± 14	4.8 ± 1.0	1.3 ± 0.2	108
cp6-Δ8-sul20	3.3 ± 0.1	13 ± 1	1.4 ± 0.1	215 ± 50	6.0 ± 3.0	1.1 ± 0.3	65

Table 2.1 Steady-state kinetic parameters for protein degradation and ATP hydrolysis by *E. coli* Lon

The error is that of non-linear-least-squares fitting. ATP per substrate was calculated by dividing V_{max} for ATP hydrolysis by V_{max} for protein degradation.

Tail clipping during Lon proteolysis prevents processive degradation of some molecules

The C-terminal residues of the sul20 degron are known to be important determinants of Lon recognition and degradation (Ishii and Amano 2001; Gur and Sauer 2009). It is possible, therefore, that proteolytic clipping of the sul20 degron accounts for the ~30% of CP6-sul20 and cp7-sul20 proteins that are not degraded.

In principle, tail clipping could occur during purification or during the degradation assay in a Lon-dependent manner. Electrospray mass spectrometry of the purified substrates showed 10-15% quantities of tail-clipped species for the cp7-sul20 substrate but effectively no tail-clipped species for the cp6-sul20 substrate (Figure 2.2A). To test if tail clipping of cp6-sul20 occurred during Lon degradation, we allowed degradation of this substrate to proceed for 2 h and purified the undegraded His_6-tagged substrate by Ni^{++}-NTA chromatography (Figure 2.2B). Electrospray mass spectrometry revealed relatively little remaining full-length substrate but a mixture of species with masses expected for removal of 2, 3, or 11 C-terminal amino acids (Figure 2.2C). Thus, it appears that Lon-dependent clipping of these residues in some substrates precludes further efficient degradation.

Tail clipping of the cp6-sul20 substrate by Lon supports a model in which the sul20 degron is the first part of this substrate that is translocated into the degradation chamber, proteolytic removal of 2 to 11 residues then occurs, and the clipped substrate is subsequently released rather than processively degraded. Release of partially degraded substrates has been observed when AAA+ proteases, including Lon, encounter a domain that is difficult to unfold in multi-domain substrates (Kenniston et al. 2005; Lee et al. 2001). For the cp6-sul20 substrate, the intact β-barrel of GFP could resist unfolding and result in release of some partially degraded substrates. In an attempt to minimize clipping and partial degradation, we deleted 8 residues preceding the sul20 degron to generate cp6-Δ8-sul20, reasoning that degradation might be more processive if a shorter linker between the degron and the β-barrel core necessitated core denaturation before the degron reached the

peptidase active sites. Indeed, Lon degraded cp6-Δ8-sul20 to ~90% completion and cp6-sul20 to ~70% completion under single-turnover conditions of enzyme excess (Figure 2.2D). In this experiment, cp6-Δ8-sul20 degradation was also ~2-fold faster than cp6-sul20 degradation (Figure 2.2D), but the steady-state V_{max} and K_M kinetic parameters for Lon degradation of cp6-Δ8-sul20 and cp6-sul20 were similar (Table 2.1). These results suggest that unfolding of cp6-Δ8-sul20 by Lon may not be the rate-limiting step under steady-state conditions.

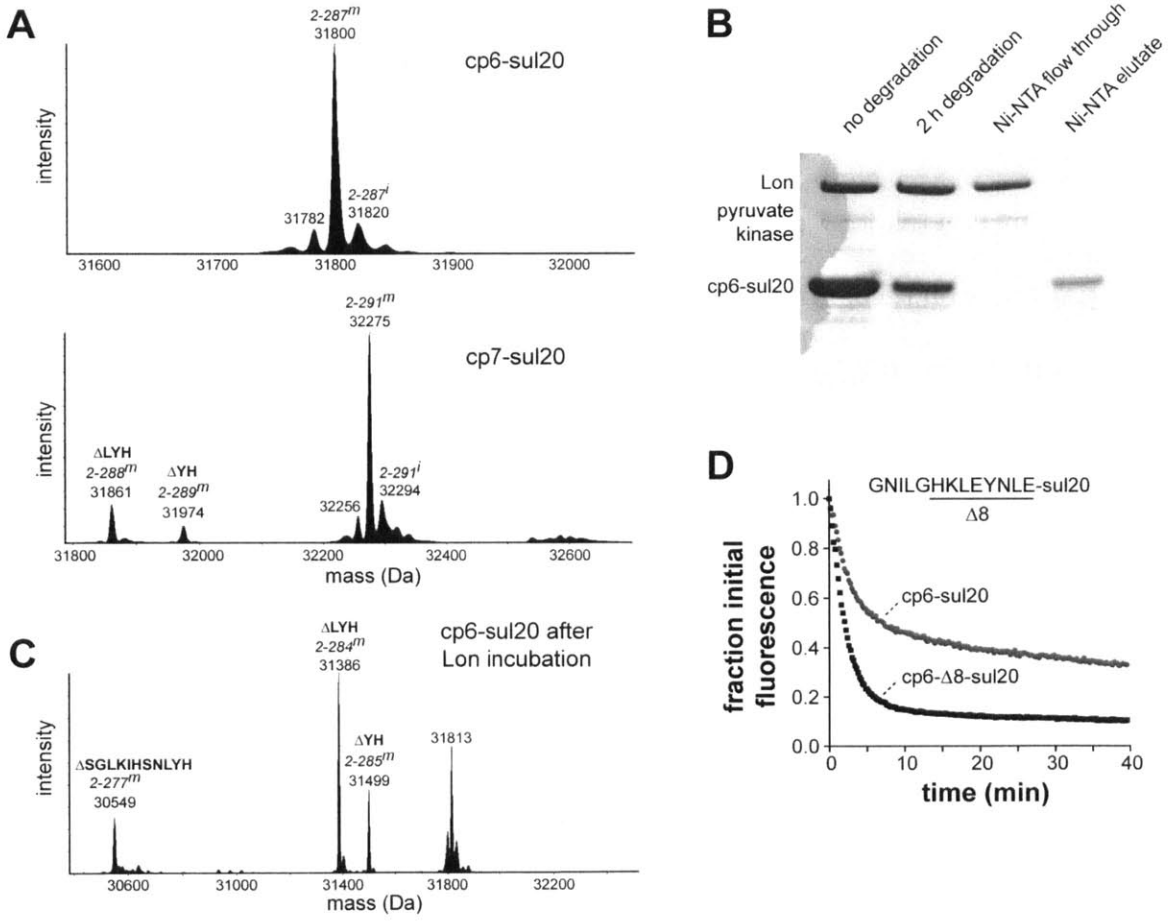

Figure 2.2 Degron clipping prevents complete degradation of fluorescent substrates

(A) Deconvoluted electrospray ionization mass spectra of the purified cp6-sul20 (top) and cp7-sul20 (bottom) proteins. Full-length cp6-sul20 with a mature chromophore but lacking the N-terminal methionine (residue 2-287^m) has an expected mass of 31,800.8 Da. Prior to chromophore maturation (2-287^i), the

expected mass is 20 Da higher. Full-length cp7-sul20 with a mature chromophore but lacking the N-terminal methionine (residues 2-291^{m}) has an expected mass of 32,274.3 Da. Small amounts of proteins missing two (ΔYH) or three residues (ΔLYH) from the C-terminal end of the sul20 degron were observed in the cp7-sul20 but not the cp6-sul20 preparation. (B) The cp6-sul20 protein (50 μM), which contains an N-terminal His_6 tag, was incubated with Lon_6 (1 μM) for 2 h at 37 °C and then purified by Ni^{++}-affinity chromatography. (C) Electrospray ionization mass spectrometry of cp6-sul20 after 2 h of Lon degradation and Ni^{++}-affinity purification revealed substantial tail clipping that removes two (ΔYH), three (ΔLYH), or eleven (ΔSGLKIHSNLYH) C-terminal residues from the sul20 degron. (D) Degradation of cp6-sul20 or cp6-Δ8-sul 20 (0.5 μM each) by Lon_6 (2 μM) at 37° C. Degradation was monitored by loss of 511 nm fluorescence after excitation at 467 nm. Fluorescence was normalized to an initial value of 1.

Inefficient Lon extraction of the degron-tagged C-terminal β-strand of GFP-sul20

Why is GFP-sul20 degraded so slowly by Lon? This protein stimulated ATP hydrolysis by Lon (K_{app} ~7 μM; not shown), whereas untagged GFP did not, indicating that Lon recognizes the sul20 degron in GFP-sul20. Previous studies demonstrated that ClpXP degrades superfolder GFP-ssrA in a two-step reaction in which the degron-tagged C-terminal β-strand is initially extracted and the resulting 10-stranded native intermediate (GFP^{1-10}) is subsequently unfolded and degraded (Martin et al. 2008; Nager et al. 2011). In principle, Lon degradation of GFP-sul20 might be very slow because the extraction and/or global unfolding step is inefficient. To test for potential effects on the extraction step, we inserted a thrombin site in the loop between strands 10 and 11 of GFP-sul20 and produced a split substrate by thrombin cleavage **(**Figure 2.3A**)**. We then incubated the split substrate (5 μM) with Lon (0.3 μM) and observed slow degradation of the sul20-tagged C-terminal fragment but no degradation of the untagged N-terminal fragment over the course of 30 min as monitored by SDS-PAGE (Figure 2.3B). This result suggests that Lon

extraction of the sul20-tagged fragment from the split substrate is inefficient. However, the C-terminal fragment did not run as a sharp band precluding quantification. As a more rigorous test of Lon's ability to extract the sul20-tagged C-terminal fragment, we exploited the fact that native GFP^{1-10} has the same fluorescence as GFP when excited at 467 nm but no fluorescence when excited at 400 nm, because excitation at the lower wavelength requires excited state transfer of a proton from the chromophore to Glu^{222} in the 11^{th} b-strand (Stoner-Ma et al. 2006; Kent et al. 2008). When thrombin-split GFP-sul20 (0.5 μM) was incubated with excess Lon (2 μM), an approximate 20% reduction in 400-nm fluorescence with little change in 467-nm fluorescence was detected over 30 min (Figure 2.3C). Thus, Lon extracts only a small fraction of the C-terminal fragments in the population of substrates. However the slow rate at which Lon extracted the degron-tagged C-terminal strand was substantially faster then the rate of degradation of the GFP-sul20 substrate, suggesting that inefficient strand extraction and slow global unfolding of the GFP^{1-10} intermediate combine to account for the resistance of GFP-sul20 to Lon degradation.

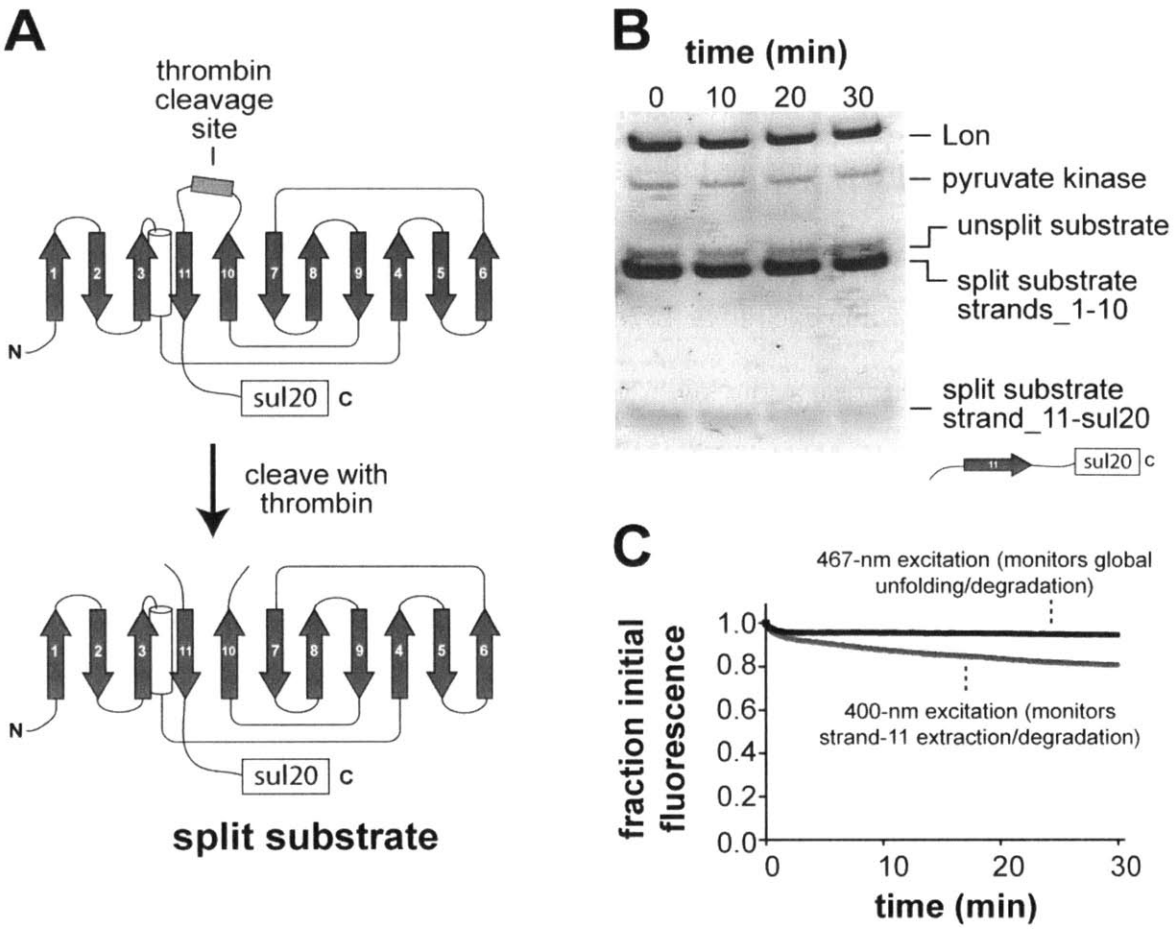

Figure 2.3 Extraction of 11th β-strand of GFP-sul20 is inefficient
(A) Cartoon showing the thrombin-cleavage site inserted between b-strands 10 and 11 of GFP-sul20 and the generation of the split substrate by thrombin cleavage. (B) The thrombin-split GFP-sul20 protein (5 μM) was incubated with Lon_6 (0.3 μM) and the reaction was monitored by SDS-PAGE, followed by staining with Coomassie Blue. (C) The thrombin-split GFP-sul20 protein (0.5 μM) was incubated with Lon_6 (2 μM) and the reaction was monitored by changes in fluorescence at 511 nm after excitation at 400 nm (extraction/degradation of β-strand 11) or 467 nm (global unfolding/degradation).

Fluorescence detection of Lon degradation in vivo

Changes in cellular fluorescence linked directly to Lon degradation would provide a potentially powerful way to monitor proteolysis *in vivo*. To test for intracellular degradation, we expressed plasmid-borne cp6-sul20 from an IPTG-inducible promoter in an *E. coli* strain that lacks a functional chromosomal *lon* gene

(ER2566), and after ~2 h induced expression of plasmid-borne Lon from an arabinose-inducible promoter. Cellular fluorescence began to decrease ~15 min after Lon induction and, after 60 min, approached the level in an experiment in which cp6-sul20 expression was not induced (Figure 2.4A). By contrast, cellular fluorescence was ~3-fold higher when cp6-sul20 expression was induced but Lon expression was not induced (Figure 2.4A). These results demonstrate that cp6-sul20 fluorescence can be used as a reporter of intracellular Lon degradation.

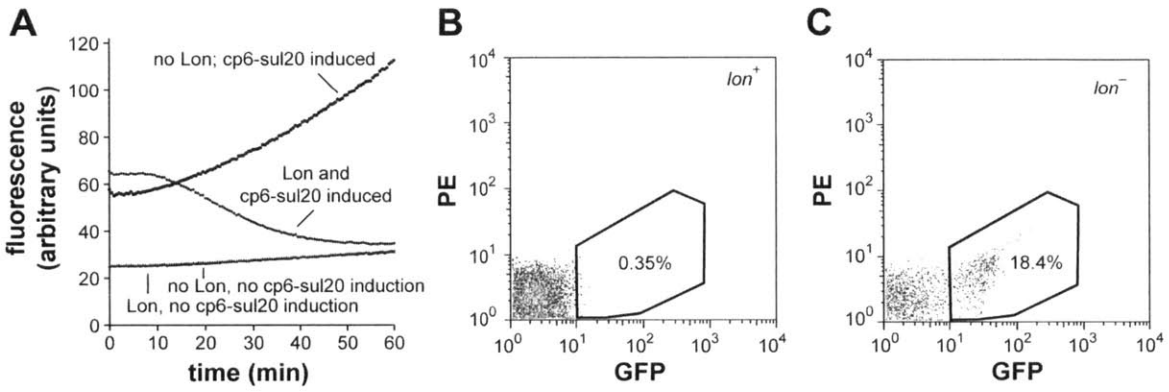

Figure 2.4 *In vivo* degradation of cp6-sul20
(A) Fluorescence of cells expressing different combinations of the cp6-sul20 GFP variant and Lon protease. At time = 0, arabinose was added to induce expression of Lon from a pBAD33 vector or to *lon*⁻ cells containing an empty pBAD33 plasmid. Two hours before, IPTG was added to induce expression of plasmid-borne cp6-sul20 or water was added to the un-induced control. (B) FACS analysis of *lon*⁺ W3110 cells expressing plasmid-borne cp6-sul20. The x-axis shows GFP fluorescence (excitation 488 nm; emission 530 ± 15 nm); the y-axis shows phycoerythrin (PE) fluorescence (excitation 488 nm; emission 585 ± 21 nm) as an auto-fluorescence control. ~0.4% of the cells fall within the GFP-positive area. (C) FACS analysis of *lon::kan* W3110 cells expressing plasmid-borne cp6-sul20. ~18% of the cells fall within the GFP-positive area.

To determine if degradation of cp6-sul20 by normal cellular concentrations of Lon could be detected, we transformed *lon*⁺ and *lon*⁻ cells with an over-expressing plasmid, induced cp6-sul20 expression for 90 min, and then performed fluorescence

activated cell sorting (FACS). Only ~0.4% of the *lon*$^+$ cells were present in the high-fluorescence gating region (Fig. 4B), whereas ~18% of the *lon*$^-$ cells were recovered in this region (Fig. 4C). Thus, normal cellular Lon levels are sufficient for reliable detection of cp6-sul20 degradation *in vivo* by FACS analysis.

Discussion

We have shown that *E. coli* Lon efficiently degrades circularly permuted variants of superfolder GFP with specific N-terminal or C-terminal degrons. These and related substrates should be useful for characterizing how Lon structure/function mutations alter the steady-state kinetics of protein degradation, for assessing the relationship between degron sequences and their activities in promoting Lon degradation, and for performing high-throughput screens for small-molecule inhibitors of degradation. Importantly, Lon degradation of these circularly permuted GFP substrates is also easily monitored *in vivo*, facilitating screens for cellular mutations that influence Lon degradation or GFP mutations that enhance or suppress degradation. Fluorescent assays for Lon degradation of peptides or non-native proteins have been reported (Rudyak and Shrader 2000; Lee and Berdis 2001; Gur and Sauer 2008; 2009), but these assays do not require substrate unfolding by Lon, are often independent of the addition of a defined degron, and cannot be used *in vivo*.

E. coli Lon degrades GFP-ssrA and GFP-b20 but at extremely slow rates ($\leq$ 0.06 min^{-1} Lon_6^{-1}) (Choy et al. 2007; Gur and Sauer 2008), precluding the use of these substrates for many assays. Consistently, we found that Lon degrades

superfolder variants of GFP-β20, GFP-sul20, and β20-GFP very slowly. In combination, these results suggest that Lon has a very difficult time unfolding the b barrel of non-permuted GFP, either when pulling on the protein from the N-terminus or the C-terminus. Indeed, we find that Lon inefficiently dislodges even the C-terminal b strand of a split variant of superfolder GFP-sul20. By contrast, Lon degrades the circularly permuted cp6-sul20 and β20-cp6 variants of superfolder GFP at rates ~50-fold faster than the best non-permuted variants. From the perspective of protein stability, the b barrels of cp6 and cp7 are less thermodynamically stable than the barrel of non permuted superfolder GFP (Nager et al. 2011). The detailed rates at which Lon degrades the circularly permuted GFP substrates probably depend on the mechanical stabilities of degron-proximal elements of secondary structure in these proteins (Lee et al. 2001).

The relative unfolding capabilities of Lon and other AAA+ proteases are highly substrate specific. For example, Lon is a very weak unfoldase for some degron-tagged substrates (Koodathingal et al. 2009) but a robust unfoldase for others (Gur et al. 2012). Our current and recent results (Nager et al. 2011) show that circular permutation of GFP can dramatically alter its susceptibility to degradation by different AAA+ proteases. For example, Lon can barely degrade non-permuted GFP but robustly degrades the cp6 variant, whereas ClpXP degrades non-permuted GFP at a much faster rate than Lon but degrades the cp6 variant at a slower rate. We anticipate that circularly permuted variants of GFP will also be useful substrates for different classes of AAA+ proteases, protein-remodeling enzymes, and protein-secretion machines.

Lon clipping of the terminal residues of the degron in some cp6-sul20 substrates prevents complete degradation and probably accounts for incomplete proteolysis of cp7-sul20 as well. For such tail clipping to occur, the sul20 degron must enter the degradation chamber but the substrate must not be committed to processive degradation. The build-up of clipped cp6-sul20 substrates to levels higher than the Lon hexamer concentration indicates that clipped substrates can be released from the enzyme, probably as a consequence of an unsuccessful unfolding attempt. Indeed, it seems likely that engaged but unclipped substrates are also released when unfolding fails, accounting for the higher energetic cost of degrading cp7-sul20 compared to cp6-sul20. Previous experiments showed that the sul20 degron can also act as a tethering sequence, facilitating degradation of a substrate by Lon, without itself being degraded (Gur et al. 2012). It will be important to determine the location of the tethering site in the structure and to determine how degron binding to different sites in Lon is coordinated.

Acknowledgements

We thank I. Papayannopoulos and Eyal Gur for help and advice. This work was supported by National Institutes of Health grant AI-16982, an National Science Foundation Graduate Research Fellowship to M.L.W., and NCI grant P30CCA14051, supporting the FACS Core of the Koch Institute for Cancer Research.

References

Baker TA, Sauer RT. 2006. ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends Biochem Sci* **31**: 647–653.

Bernstein SH, Venkatesh S, Li M, Lee J, Lu B, Hilchey SP, Morse KM, Metcalfe HM, Skalska J, Andreeff M, et al. 2012. The mitochondrial ATP-dependent Lon protease: a novel target in lymphoma death mediated by the synthetic triterpenoid CDDO and its derivatives. *Blood*.

Choy JS, Aung LL, Karzai AW. 2007. Lon protease degrades transfer-messenger RNA-tagged proteins. *J Bacteriol* **189**: 6564–6571.

Gonzalez M, Frank EG, Levine AS, Woodgate R. 1998. Lon-mediated proteolysis of the Escherichia coli UmuD mutagenesis protein: in vitro degradation and identification of residues required for proteolysis. *Genes Dev* **12**: 3889–3899.

Gur E, Sauer RT. 2009. Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine. *Proceedings of the National Academy of Sciences* **106**: 18503–18508.

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

Gur E, Vishkautzan M, Sauer RT. 2012. Protein unfolding and degradation by the AAA+ Lon protease. *Protein Sci* **21**: 268–278.

Heim R, Prasher DC, Tsien RY. 1994. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci USA* **91**: 12501–12504.

Higashitani A, Ishii Y, Kato Y, Koriuchi K. 1997. Functional dissection of a cell-division inhibitor, SulA, of Escherichia coli and its negative regulation by Lon. *Mol Gen Genet* **254**: 351–357.

Ingmer H, Brøndsted L. 2009. Proteases in bacterial pathogenesis. *Res Microbiol* **160**: 704–710.

Ishii Y, Amano F. 2001. Regulation of SulA cleavage by Lon protease by the C-terminal amino acid of SulA, histidine. *Biochem J* **358**: 473–480.

Kenniston JA, Baker TA, Fernandez JM, Sauer RT. 2003. Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* **114**: 511–520.

Kenniston JA, Baker TA, Sauer RT. 2005. Partitioning between unfolding and release of native domains during ClpXP degradation determines substrate selectivity and partial processing. *Proc Natl Acad Sci USA* **102**: 1390–1395.

Kenniston JA, Burton RE, Siddiqui SM, Baker TA, Sauer RT. 2004. Effects of local protein stability and the geometric position of the substrate degradation tag on the efficiency of ClpXP denaturation and degradation. *J Struct Biol* **146**: 130–140.

Kent KP, Childs W, Boxer SG. 2008. Deconstructing green fluorescent protein. *J Am Chem Soc* **130**: 9664–9665.

Koodathingal P, Jaffe NE, Kraut DA, Prakash S, Fishbain S, Herman C, Matouschek A. 2009. ATP-dependent proteases differ substantially in their ability to unfold globular proteins. *J Biol Chem* **284**: 18674–18684.

Lee C, Schwartz MP, Prakash S, Iwakura M, Matouschek A. 2001. ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Molecular Cell* **7**: 627–637.

Lee I, Berdis AJ. 2001. Adenosine triphosphate-dependent degradation of a fluorescent lambda N substrate mimic by Lon protease. *Anal Biochem* **291**: 74–83.

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

Nager AR, Baker TA, Sauer RT. 2011. Stepwise unfolding of a β barrel protein by the AAA+ ClpXP protease. *J Mol Biol* **413**: 4–16.

Nørby JG. 1988. Coupled assay of Na+,K+-ATPase activity. *Methods in Enzymology* **156**: 116–119.

Ormö M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. 1996. Crystal structure of the Aequorea victoria green fluorescent protein. *Science* **273**: 1392–1395.

Pédelacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS. 2006. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* **24**: 79–88.

Reeder PJ, Huang Y-M, Dordick JS, Bystroff C. 2010. A rewired green fluorescent protein: folding and function in a nonsequential, noncircular GFP permutant. *Biochemistry* **49**: 10773–10779.

Robertson GT, Kovach ME, Allen CA, Ficht TA, Roop RM. 2000. The Brucella abortus Lon functions as a generalized stress response protease and is required for wild-type virulence in BALB/c mice. *Mol Microbiol* **35**: 577–588.

Rudyak SG, Shrader TE. 2000. Polypeptide stimulators of the Ms-Lon protease. *Protein Sci* **9**: 1810–1817.

Sauer RT, Baker TA. 2011. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **80**: 587–612.

Stoner-Ma D, Melief EH, Nappa J, Ronayne KL, Tonge PJ, Meech SR. 2006. Proton relay reaction in green fluorescent protein (GFP): Polarization-resolved ultrafast vibrational spectroscopy of isotopically edited GFP. *J Phys Chem B* **110**: 22009–22018.

Striebel F, Kress W, Weber-Ban E. 2009. Controlled destruction: AAA+ ATPases in protein degradation from bacteria to eukaryotes. *Current Opinion in Structural Biology* **19**: 209–217.

Torres-Cabassa AS, Gottesman S. 1987. Capsule synthesis in Escherichia coli K-12 is regulated by proteolysis. *J Bacteriol* **169**: 981–989.

Yang HJ, Lee JS, Cha JY, Baik HS. 2011. Negative regulation of pathogenesis in Pseudomonas syringae pv. tabaci 11528 by ATP-dependent Lon protease. *Mol Cells* **32**: 317–323.

Chapter 3

Distinct quaternary structures of the AAA+ Lon protease control substrate degradation

This work has been submitted to *PNAS* as Ellen F. Vieux, Matthew L. Wohlever, James Z. Chen, Robert T. Sauer, and Tania A. Baker.

E.F.V. and M.L.W. performed biophysical and biochemical characterization. J.Z.C. and E.F.V. performed structural studies. E.F.V. and T.A.B. wrote the initial manuscript.

Abstract

Lon is a AAA+ protease that controls cell division in response to stress and also degrades misfolded and damaged proteins. Subunits of Lon are known to assemble into ring-shaped homohexamers that enclose an internal degradation chamber. Here, we demonstrate that hexamers of *E. coli* Lon also interact to form a dodecamer at physiological protein concentrations. Electron microscopy of this dodecamer reveals a prolate structure with the protease chambers at the distal ends and a matrix of N domains forming an equatorial hexamer-hexamer interface, with portals of ~45 Å providing access to the enzyme lumen. Compared to hexamers, Lon dodecamers are much less active in degrading large substrates but equally active in degrading small substrates. Our results support a novel gating mechanism that allows the repertoire of Lon substrates to be tuned by its assembly state.

Introduction

Protein quality control is vital under stress conditions that promote protein unfolding and aggregation. *Escherichia coli* Lon degrades many unfolded proteins (Fredriksson et al. 2005; Kowit and Goldberg 1977; Shineberg and Zipser 1973) and also degrades folded proteins, including SulA, IbpA, and IbpB (Gottesman et al. 1981; Mizusawa and Gottesman 1983; Bissonnette et al. 2010). In *E. coli* and many other bacteria, Lon is up-regulated under numerous stress conditions (Goff et al. 1984; Phillips et al. 1984; Goff and Goldberg 1985; Van Melderen and Aertsen 2009). In mitochondria, Lon helps combat oxidative stress (Bender et al. 2011; Venkatesh et al. 2011; Ngo and Davies 2009; Ngo et al. 2011), and human mitochondrial Lon was recently identified as a potential anti-lymphoma target (Bernstein et al. 2012). It is widely believed that a major role of Lon in all organisms is to degrade misfolded proteins (Kowit and Goldberg 1977; Van Melderen and Aertsen 2009; Rosen et al. 2002).

Lon subunits consist of an N domain, a central AAA+ ATPase module, and a C-terminal peptidase domain. Although early reports suggested that Lon might be a tetramer (Goldberg et al. 1994), it is now clear that six subunits of the *E. coli* enzyme assemble into a hexamer with an internal degradation chamber accessible via an axial pore in the AAA+ ring (Botos et al. 2004; Park et al. 2006). Lon substrates are recognized, unfolded if necessary by ATP-dependent reactions mediated by the AAA+ ring, and then translocated through the pore and into the peptidase chamber for degradation (Sauer and Baker 2011).

In many families of ATP-dependent proteases, the AAA+ unfolding/ translocation ring and the self-compartmentalized peptidase are encoded by distinct polypeptides, which assemble into independent oligomers prior to interacting to form the functional protease (Sauer et al. 2004; Joshi et al. 2004). For example, the ClpXP protease consists of AAA+ ClpX hexamers, which dock with the self-compartmentalized ClpP peptidase. This interaction suppresses the ATPase rate of ClpX and enhances the peptidase activity of ClpP (Joshi et al. 2004). Lon activity cannot be controlled in this way because the ATPase and protease domains are always physically attached. Little is currently known about how Lon activity is regulated, although mutational studies show that the AAA+ and peptidase domains influence each other's activities (Roudiak et al. 1998; Starkova et al. 1998; van Dijl et al. 1998). In some cases, the function of the two domains also appears to be linked via allosteric communication mediated by substrate binding (Waxman and Goldberg 1986; Gur and Sauer 2009).

Here, we demonstrate that Lon forms dodecamers that equilibrate with hexamers at physiological concentrations. A structure determined by electron microscopy (EM) at low resolution reveals a novel protease architecture with the degradation chambers of each hexamer at opposite ends of a prolate ellipsoid. Near the equator of this structure, the arrangement of N domains creates portals, which could serve as entry sites for protein substrates. Formation of the dodecamer suppresses proteolysis of large but not small protein substrates, suggesting that the

dodecamer uses a gating mechanism that allows the repertoire of Lon substrates to be tuned by its state of assembly.

Experimental Procedures

Protein purification

Wild-type *E. coli* Lon (Goff and Goldberg 1985) and the Lon^{S679A} variant (Van Melderen and Gottesman 1999) were purified as described (Gur and Sauer 2008) with minor changes. After lysis by sonication or French press, lysates were incubated with benzonase (Merck KGaA, Damstadt, Germany) for 1 h at 4° C. Following elution from P11 phosphocellulose (Whatman, GE Healthcare, Piscataway, NJ), the sample was concentrated and buffer exchanged to 2 mL either in 25 mM HEPES-KOH (pH 7.5), 2 M NaCl, 1 mM EDTA, 0.1 mM TCEP, and 10% glycerol, or 50 mM HEPES-KOH (pH 7.5), 2 M NaCl, and 0.1 mM TCEP, passed through a 0.45 μm filter, and purified on a HR 10/300 Superose 6 gel filtration column (GE Healthcare, Piscataway, NJ). Fractions containing Lon at >95% purity, as judged by SDS-PAGE and 280/260 nm absorbance, were combined, dialyzed against Lon storage buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM EDTA, and 0.1 mM TCEP), concentrated, flash frozen in aliquots, and stored at -80 °C.

E. coli IbpB was purified as described (Bissonnette et al. 2010), dialyzed against Ibp storage buffer (50 mM HEPES-KOH (pH 8), 600 mM potassium glutamate, 20% sucrose, and 0.1 mM TCEP), flash frozen in aliquots, and stored at -

80° C. His_6-tagged titin-I27 proteins with sul20 or β20 degrons were purified as described (Gur and Sauer 2008).

SEC-MALS

Size exclusion chromatography was performed on a WTC-030S5 size-exclusion column using an Agilent HPLC. Multi-angle laser light scattering was measured in line using a Wyatt DAWN-HELEOS instrument; concentrations were determined using an Optilab rEX instrument. Standard Zimm-plot analysis was preformed with the ASTRA software 5.3.4 (Wyatt Technology Corporation, Santa Barbara, CA). Lon^{S679A} samples (loading concentrations 24, 12, and 6 μM) were run in 50 mM HEPES-KOH (pH 7.6), 150 mM NaCl, 20 mM $MgCl_2$, 10% glycerol, and 0.1 mM TCEP at room temperature.

Ultracentrifugation

SV-AUC experiments for Lon^{S679A} were preformed using a Beckman OptimaXL-I analytical ultracentrifuge (Biophysical Instrumentation Facility, MIT). Samples were dialyzed overnight against 50 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 0.01 mM EDTA, and 0.1 mM TCEP. Before loading the cells, 1 mM $MgCl_2$ and 0.1 mM ATPγS were added to the samples. Samples were loaded in dual-sector charcoal-filled epon centerpieces and centrifuged at 16,000 rpm in a An50-Ti rotor at 20° C. SEDFIT (Brown and Schuck 2006) was used to calculate the continuous distribution of sedimentation coefficients from 1*S* to 60*S* at resolutions of 200 or 100 scans per

concentration with a confidence level (F-ratio) of 0.95. Calculations were performed using a density of 1.00831, a viscosity of 0.010475, and a Lon partial specific volume of 0.7431 (SEDNTERP; J. Philo; http://www.jphilo.mailway.com).

Western blotting

E. coli W3110 cells were grown at 30° C in M9 medium supplemented with 0.4% glucose, 100 μM $CaCl_2$, 2 mM $MgSO_4$, 0.2% w/v thiamin, and 0.2% casamino acids. At an OD_{600} of 0.3, cultures were split, additional medium at 54° C was added to the heat-shock sample, and additional medium at 30° C was added to the control. The final temperature of the heat-shock sample was 42° C. Aliquots of 1 mL were taken at each time point from both samples, OD_{600} was recorded, cells were pelleted by centrifugation, and the supernatant was removed. Pellets were stored at −20° C until resuspended to 2.5 OD_{600} equivalents with 5X SDS loading dye. Resuspended samples were heated at 99° C for 10 min while shaking, and then cooled. The samples (10 μL) were loaded on Mini-PROTEAN TGX 4-20% precast gels (Bio-Rad, Hercules, CA). The gels were transferred onto filter paper using a wet-transfer apparatus (BioRad, Hercules, CA), probed with anti-Lon polyclonal antibody (produced by Covance Research Products) at a 1:2,000 dilution for 4 h at room temperature, incubated with goat anti-rabbit IgG-AP conjugate (BioRad, Hercules, CA) at a 1:10,000 dilution for 1 h at room temperature, and developed with alkaline phosphatase dephosphorylates ECF Substrate (GE Healthcare, Piscataway, NJ). The

blots were exposed with a blue laser and quantified with ImageQuant software (GE Healthcare, Piscataway, NJ).

The intracellular concentration of Lon was calculated from the western blot using purified Lon^{S679A} as a standard. 1.9 ± 0.36 ng of Lon was present in a sample containing 3.55 ± 0.15 x 10^8 cells per OD (as determined by counting colony-forming units under the conditions of the experiment), giving a value of 2.2 ± 0.6 x 10^{-16} g/cell. Errors represent ± 1 SEM (*N*=4) and were propagated through all calculations. The intracellular Lon concentration was calculated using a subunit M_R of 87.5 kDa and a cell volume of 10^{-15} L (Ali Azam et al. 1999). The fold increase in Lon concentration following a temperature increase from 30 to 42° C was determined by normalizing against a nonspecific band at the bottom of the Western to correct for any changes in cell density over the time course and dividing the intensity of the bands from cells grown at 42° C by the intensity of the bands from cells grown at 30° C.

Single-particle EM data collection and analysis

Wild-type Lon (24 μM subunit equivalents in storage buffer) was incubated at 37° C with 0.1 mM ATP, and 10 mM $MgCl_2$ for 4 min, sul20 peptide (synthesized and purified in house) was added to a final concentration of 200 μM and incubated for 1 min, ATPγS was added to a final concentration of 5 mM, the sample was diluted ~100-fold with storage buffer without glycerol, and immediately negatively stained with uranyl acetate (1%) on continuous carbon-film grids. Electron micrographs of single particles were recorded with a 2Kx2K CCD camera on a TF20 electron

microscope at 29,000x nominal magnification. A total of 4,235 oligomeric Lon particles were then boxed into a single stack, from which the particle images were 2x-binned to 7.92 Å per pixel. The dataset was subjected to "direct classification" using PARTICLE (www.image-analysis.net/EM), which is free of alignment error or reference bias.

For the dodecamer three-dimensional reconstruction, a set of single-particle tomographic volumes were first collected and averaged to establish an unbiased initial model, which served as the reference to align the class averages in single-particle reconstruction and refinement. The model was further validated by the tilt-pair technique (Henderson et al. 2011), utilizing 86 pairs at 30°-tilt separation, in which the average angular deviation from the controlled tilting angle was less than 7°. Data analysis (particle screening, classification, single-particle tomography, three-dimensional reconstruction, and model validation) was performed in the PARTICLE software package.

The 3M6A.pdb crystal structure includes the protease and ATPase domains of *B. subtilis* Lon in an indefinite spiral conformation (Duman and Löwe 2010). To make a planar hexameric model for fitting EM density, we aligned six copies of chain A from 3M6A.pdb to the protease domains in the hexameric structure of the *E. coli* Lon protease domain (1RR9.pdb). To fit the six equatorial bridges, we placed two *E. coli* Lon N domains (3LCJ.pdb) in each bridge with the coiled-coil regions crossing and rotated the globular domains to fit the density.

ATPase assays

ATP-hydrolysis rates were measured at 37 °C in a plate reader using an NADH enzyme-linked assay (Nørby 1988; Lindsley 2001) in buffer containing 4 mM ATP, 5 mM $MgCl_2$, 5 mM KCl, 2% DMSO, 360 mM potassium glutamate, 12% sucrose, and 50 mM HEPES-KOH (pH 8).

Degradation assays

IbpB degradation reactions contained 60% Ibp storage buffer, 5% Lon buffer, 5 mM $MgCl_2$, 5 mM KCl, and 2% DMSO. An ATP-regeneration system, containing a final concentration of 4 mM ATP, 100 mg ml^{-1} creatine kinase, and 10 mM creatine phosphate, was added to initiate the reaction. Degradation was monitored 37° C by the formation of radioactive peptides soluble in trichloroacetic acid as described (Bissonnette et al. 2010; Gottesman et al. 1998). Fluorescein isothiocyanate (FITC) casein type III (Sigma-Aldrich, St. Louis) was resuspended in Ibp storage buffer. Degradation was monitored by the increase of fluorescence at 525 nm with excitation at 365 nm in a plate reader at 37° C using final conditions as described for IbpB degradation.

Degradation of titin-I27 constructs was carried out at 37° C in buffer containing 25 mM Tris-HCl (pH 8.0), 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 2 mM ATP, 20 mM phosphoenolpyruvate, and 10 U/mL pyruvate kinase. Kinetics were determined using a mixture of 5% ^{35}S-labeled substrate and 95% unlabeled substrate as described (Gottesman et al. 1998).

Results

Lon exists in multiple oligomeric forms

In the process of characterizing *E. coli* Lon by size-exclusion chromatography (SEC) and multi-angle laser light scattering (MALS), we observed enzyme complexes with properties similar to those expected for hexamers (calculated M_R 525 kDa) and dodecamers (calculated M_R 1050 kDa). For example, SEC-MALS of the Lon^{S679A} variant, which had an active-site mutation in the peptidase domain to prevent auto-degradation (Van Melderen and Gottesman 1999), revealed two major species corresponding to molecular weights of 565 ± 13 and 930 ± 5 kDa (Figure 3.1A). Based on previous characterization of Lon as a hexamer (Botos et al. 2004; Park et al. 2006), the simplest interpretation of these results is that the larger species is a Lon dodecamer, which may dissociate to some extent during the SEC run.

To characterize assembly further, we used sedimentation velocity analytical ultracentrifugation (SV-AUC) at multiple concentrations of Lon^{S679A} in the presence of 100 μM ATPγS (Figure 3.1B). Both the large (dodecamer) and smaller (hexamer) assemblies were clearly detectable at multiple Lon^{S679A} concentrations, as was a smaller species, which appeared to be a monomer. As expected, the dodecamer was more highly populated at higher concentrations, and the hexamer and presumed monomer populations increased at lower concentrations. The hexamer and dodecamer were both populated at concentrations that are physiologically relevant (see below).

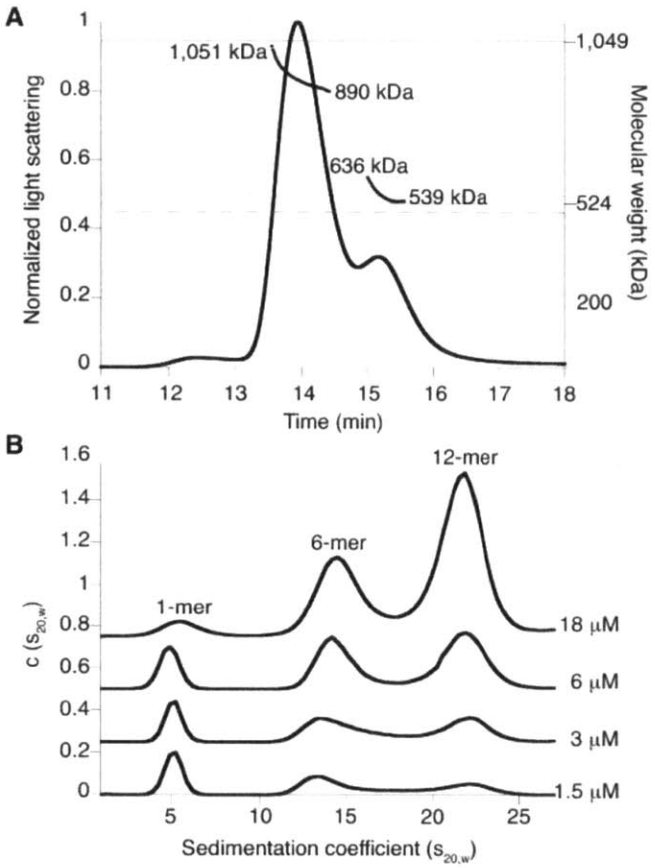

Figure 3.1 Lon assembles into dodecamers as well as hexamers
(A) Catalytically inactive Lon^{S679A} (24 μM loading concentration) formed hexamers (expected M_R ~525 kDa) and dodecamers (expected M_R ~1,050 kDa) in SEC-MALS experiments. The variation in apparent M_R over both peaks suggests that dodecamers and hexamers are in equilibrium. Dashed lines represent an error of 5% in measurement of molecular weight. Chromatography was performed at room temperature in 50 mM HEPES (pH 7.6), 150 mM NaCl, 20 mM $MgCl_2$, 10% glycerol, and 0.1 mM TCEP. (B) Concentration-dependent changes in the population of Lon^{S679A} dodecamers, hexamers, and monomers in SV-AUC $c(s_{20,w})$ distributions. Traces at each concentration were offset on the y-axis for clarity. Experiments were performed at 20° C in 50 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 0.01 mM EDTA, 0.1 mM TCEP, 1 mM $MgCl_2$, and 0.1 mM ATPγS.

To confirm that wild-type Lon also formed dodecamers, we used analytical gel filtration (Figure 3.2), which is rapid and minimizes auto-proteolysis. In addition, because nucleotide can affect Lon assembly, experiments were performed with ATPγS or without nucleotide. Under the conditions tested, Lon appeared to

chromatograph as a mixture of hexamers and dodecamers, and nucleotide had little effect on the distribution of these species (Figure 3.2).

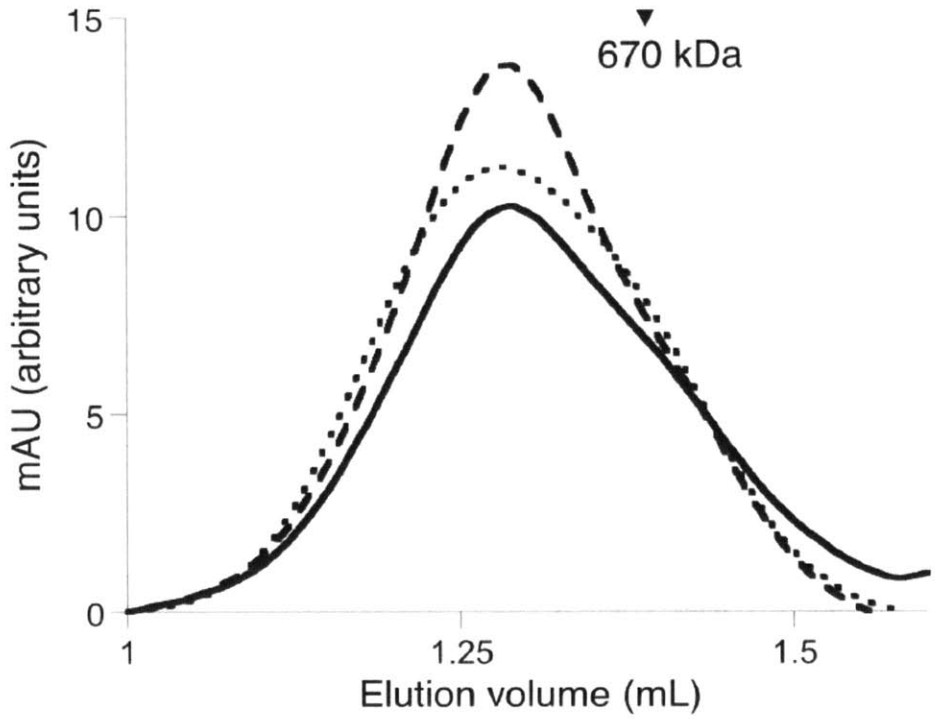

Figure 3.2 Relative distributions of Lon hexamers and dodecamers were similar under various conditions

(A) Analytical gel filtration of Lon^{S679A} (solid line), Lon^{S679A} with ATPγS (dashed line), and wild-type Lon with ATPγS (dotted line) gave similar results. In all experiments, the loading concentration of Lon was 10 mM and the column buffer was 50 mM HEPES-KOH (pH 7.6), 150 mM NaCl, 0.01 mM EDTA, 0.1 mM TCEP, 1 mM $MgCl_2$, with or without 100 mM ATPγS. Proteins were chromatographed at room temperature on a Superose-6 column using an Ettan system (GE HealthCare, Uppsala, Sweden). Protein standards (Bio-Rad, Hercules, CA) were run under the same conditions for comparison. The arrow marks the position of the 670 kDa standard.

Dodecamers should exist at intracellular concentrations

To investigate the potential for Lon dodecamers to form *in vivo*, we determined intracellular concentrations using quantitative western blots. A dilution series of purified Lon was analyzed on the same membrane as Lon from cells grown

at 30° C (Figure 3.3A). The concentration of Lon in monomer equivalents ranged from 1.7 to 3.7 μM over four measurements and averaged 2.5 ± 0.5 μM (SEM) for cells grown at 30° C. Lon forms hexamers and dodecamers at these concentrations *in vitro*. Following a temperature increase to 42° C, a modest increase in Lon levels was observed by western blots (~1.2 fold; Figure 3.3B).

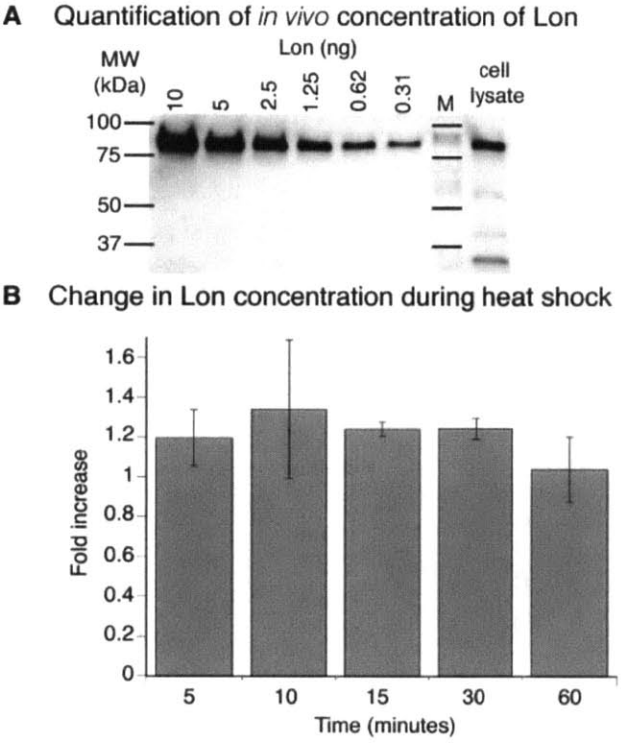

Figure 3.3 Quantitative western blots of Lon levels in *E. coli* are consistent with dodecamer formation

(A) Western blot following SDS-PAGE of different concentrations of purified Lon and an *E. coli* lysate before heat shock. The intracellular concentration of Lon was 2.5 ± 0.5 μM or ~1500 ± 300 Lon monomers per cell at 30° C (N=4). (B) Lon concentrations increased slightly in cells grown after a temperature increase to at 42° C compared to cells grown at 30° C (N = 3). In all panels values are averages ± 1 SEM.

EM dodecamer structure

EM images of negatively stained Lon complexes showed two major populations (Figure 3.4A). Classification of over 4000 particles revealed that one major species was a dodecamer with roughly six-fold symmetry (Figure 3.4B). These dodecamers preferentially assumed a side-view orientation on the grid (Figure 3.4B). Multiple views of hexamers were also observed. The side view of the hexamer corresponded to roughly half of the density observed in the side view of the dodecamer (Figure 3.4B).

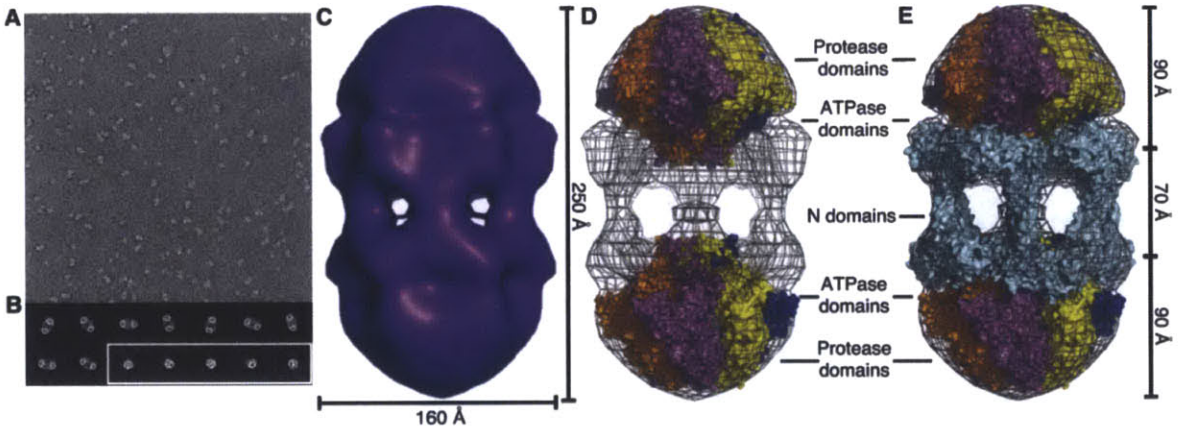

Figure 3.4 Lon dodecamers and hexamers as visualized by negative-stain EM

(A) Representative field of view of electron micrographs of single wild-type Lon particles shows two major populations, differing by roughly two-fold in size. (B) Representative class images of dodecamers (top and bottom left) and hexamers (bottom right, enclosed in white box). Samples were preincubated at 37° C and following dilution onto grids for imaging contained 0.24 μM Lon (subunit equivalents), 2 μM sul20 peptide, and 50 μM ATPγS. (C) Three-dimensional reconstruction of the Lon dodecamer. In the orientation shown, the dodecamer (displayed in surface representation) is ~250 Å high and ~160 Å wide. (D) The top and bottom portions of the electron-density map were fit well by hexameric models of the peptidase and ATPase domains from a crystal structure of *B. subtilis* Lon (3M6A.pdb). (E) Same as panel D except the equatorial density was fit using a dimeric model of the N domain from *E. coli* Lon (3LJC.pbd). The equatorial portals between N domains are ~45 Å in diameter. These figures were generated with PyMOL (version 1.2r3pre, Schrödinger, LLC.) using mesh level 4 and surface level 3.

Following six-fold (C6) averaging, a three-dimensional reconstruction of the dodecamer classes revealed face-to-face hexameric rings connected by six strands of density (Figure 3.4C). The dodecamer was ~250 Å long and ~160 Å wide and roughly the shape of a prolate ellipsoid. We created planar hexameric models of the ATPase and the protease domains from a crystal structure including these portions of a *B. subtilis* Lon subunit ((Duman and Löwe 2010), see Experimental Procedures) and manually placed them into the density map (Figures 3.4D & E). The protease domains fit well into the density at the distal ends, with the adjacent ATPase domains closer to the equator. There were some clashes between the protease and ATPase domains, but the low resolution of the structure and uncertainty about the quality of our hexameric models precluded better fitting. The extra density near the equator was fit as a matrix of interacting *E. coli* Lon N domains, arranged as overlapping dimers through a coiled-coil region, which create bridges between the two halves of the structure (Figure 3.4E). Strikingly, portals with diameters of ~45 Å were clearly visible between the N domains. As discussed below, these portals may exclude entry of large substrates into the lumen of the enzyme, where the degradation machinery resides.

Hexamers have higher basal ATPase activity than dodecamers

Because equilibration precluded isolation of pure hexamers or dodecamers for functional assays, we measured rates of ATP hydrolysis over the same range of Lon concentrations that altered the dodecamer/hexamer ratio in the SV-AUC studies.

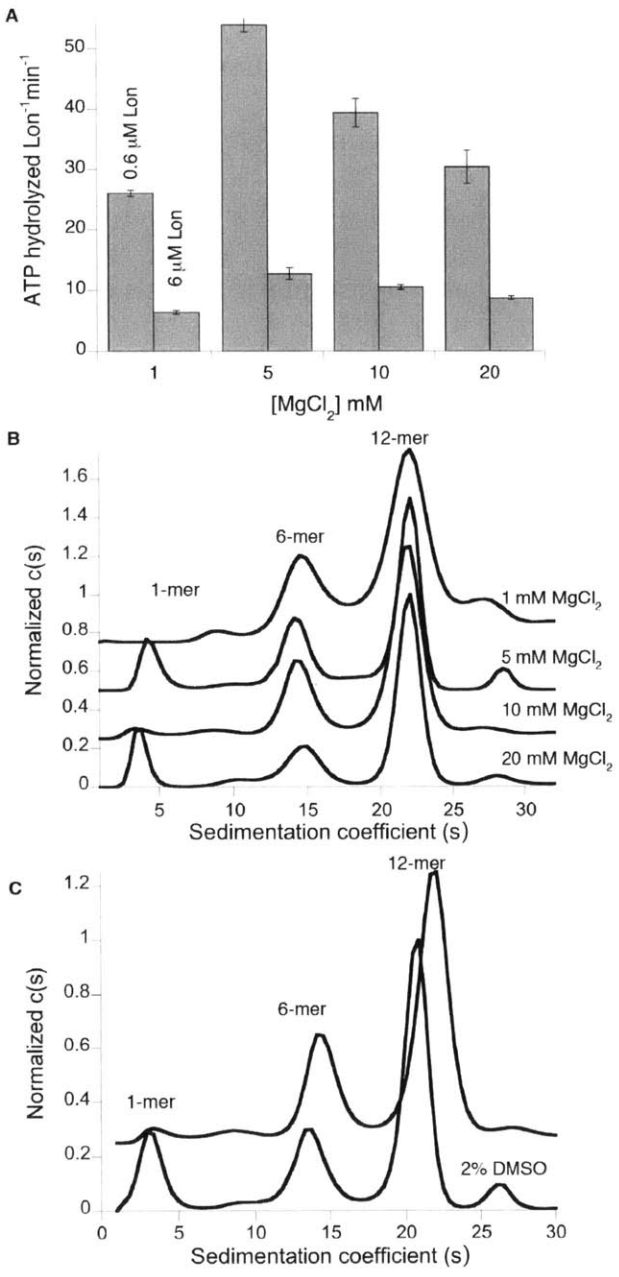

Figure 3.5 Relative distributions of Lon hexamers and dodecamers were similar under various conditions

(A) The rate of ATP hydrolysis remained dependent on Lon concentration at different concentrations of $MgCl_2$.The maximum rate of ATP hydrolysis was observed at 5 mM $MgCl_2$ for high and low concentrations of Lon. (B) Changes in $MgCl_2$ concentration did not markedly alter the distribution of dodecamers and hexamers of wild-type Lon (1.5 mM initial concentration) in SV-AUC c($s_{20,w}$) distributions. Experiments were performed in 50 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 0.01 mM EDTA, 0.1 mM TCEP, and 0.1 mM ATPγS at 20 krpm, 20 °C. For fitting the c(s) distributions the confidence level (F-ratio) was set to 0.68, and for 1 mM

$MgCl_2$ the frictional ratio was fixed at 1.48. (C) Effect of addition of 2% DMSO on SV-AUC c($s_{20,w}$) distributions for wild-type Lon (1.5 mM initial concentration) in in 50 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 0.01 mM EDTA, 0.1 mM TCEP, 10 mM $MgCl_2$ and 0.1 mM ATPγS. Other conditions were as in panel b. Calculation of buffer density and viscosity did not take changes due to DMSO addition into account and may explain the slight shift in s values.

In control experiments, we confirmed that hexamers and dodecamers were still present as the major species in buffers and under conditions that mimicked our enzyme-assay conditions (Figure 3.5). When rates of ATP hydrolysis were normalized for the total number of Lon subunits in each reaction, basal hydrolysis slowed substantially as the dodecamer/hexamer ratio increased (Figure 3.6 & 3.7A).

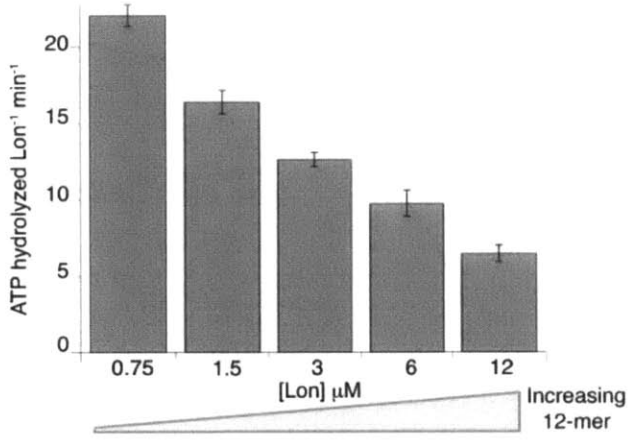

Figure 3.6 Lon hydrolyzes ATP more slowly at higher concentrations

Lon concentrations are calculated in subunit equivalents. As Lon concentration increases, the dodecamer/hexamer ratio increases. Data are plotted as averages ± 1 SEM (N = 3). Reactions were performed at 37° C and contained 4 mM ATP, 5 mM $MgCl_2$, 5 mM KCl, 2% DMSO, 360 mM potassium glutamate, 12% sucrose, and 50 mM HEPES-KOH (pH 8).

The concentration dependence of the ATPase activity was fitted best by a K_D of 3.3 ± 1.5 μM for the hexamer-dodecamer interaction, a hydrolysis rate of 23 ± 3.6 subunit^{-1} min^{-1} for the hexamer, and a hydrolysis rate of 1.8 ± 1.3 subunit^{-1} min^{-1} for the

dodecamer (Figure 3.7A). Thus, the hexamer hydrolyzes ATP ~10-fold faster than the dodecamer in the absence of protein substrates.

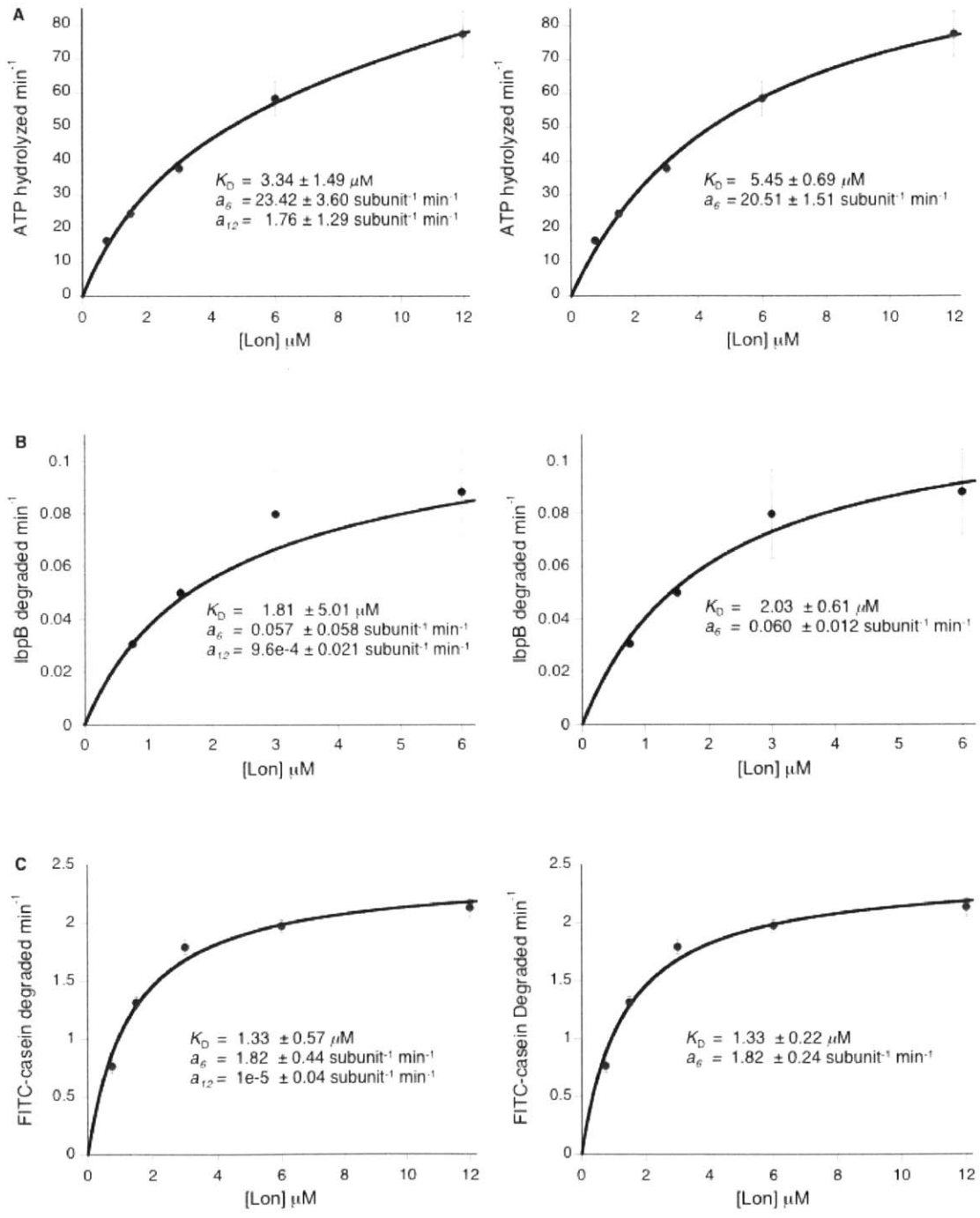

Figure 3.7 The concentration dependence of Lon activity

Data was fitted using the nonlinear least-squares tool of Kaleidagraph 3.4 (Synergy Software, Reading, PA) to the equation:

$$\text{activity} = a_6*[\text{Lon}]/(1 + [\text{Lon}]/K_D) + a_{12}*[\text{Lon}]/(1 + K_D/[\text{Lon}])$$

where a_6 and a_{12} are specific activities (units of subunit^{-1} min^{-1}) of the hexamer and dodecamer, respectively, and K_D is the equilibrium constant for dodecamer-hexamer dissociation. (A) Left panel. Fitting the concentration dependence of ATP hydrolysis gave a specific activity for the Lon hexamer ~10-fold greater than for the Lon dodecamer (R = 0.998). Right panel. Setting the dodecamer ATP-hydrolysis rate to zero also gave a satisfactory fit (R = 0.997). (B) Left panel. Fitting the concentration dependence of IbpB-degradation rates gave a specific activity for the Lon hexamer ~60-fold greater than for the Lon dodecamer (R = 0.95). The fitted a_{12} was constrained between values of 0 and 0.1 subunit^{-1} min^{-1}.. Right panel. Setting the dodecamer degradation rate to zero also gave a satisfactory fit (R = 0.98). (C) Left panel. Fits (R = 0.989) and parameters of the concentration dependence of degradation of FITC-casein degradation. The fitted a_{12} was constrained between values of 0 and 2.5 subunit^{-1} min^{-1}. Right panel. Setting the dodecamer degradation rate to zero also gave a satisfactory fit (R = 0.989). In all panels, parameters errors are those on non-linear least squares fitting and data values are averages ± 1 SEM (N = 3).

Dodecamers degrade "large" substrates poorly and "small" substrates well

To evaluate degradation of different substrates by the dodecamer, we assayed proteolysis using a range of Lon concentrations. The first substrate was an inclusion-body binding protein, specifically *E. coli* IbpB, which contains a native a-crystallin domain that is recognized by Lon (Bissonnette et al. 2010). Although IbpB monomers are relatively small (~16 kDa), they assemble into large cage-like oligomers (Jiao et al. 2005; Shearstone and Baneyx 1999). For example, IbpB runs at an apparent M_R > 670 kDa in gel-filtration chromatography (Figure 3.8 and (Bissonnette et al. 2010)).

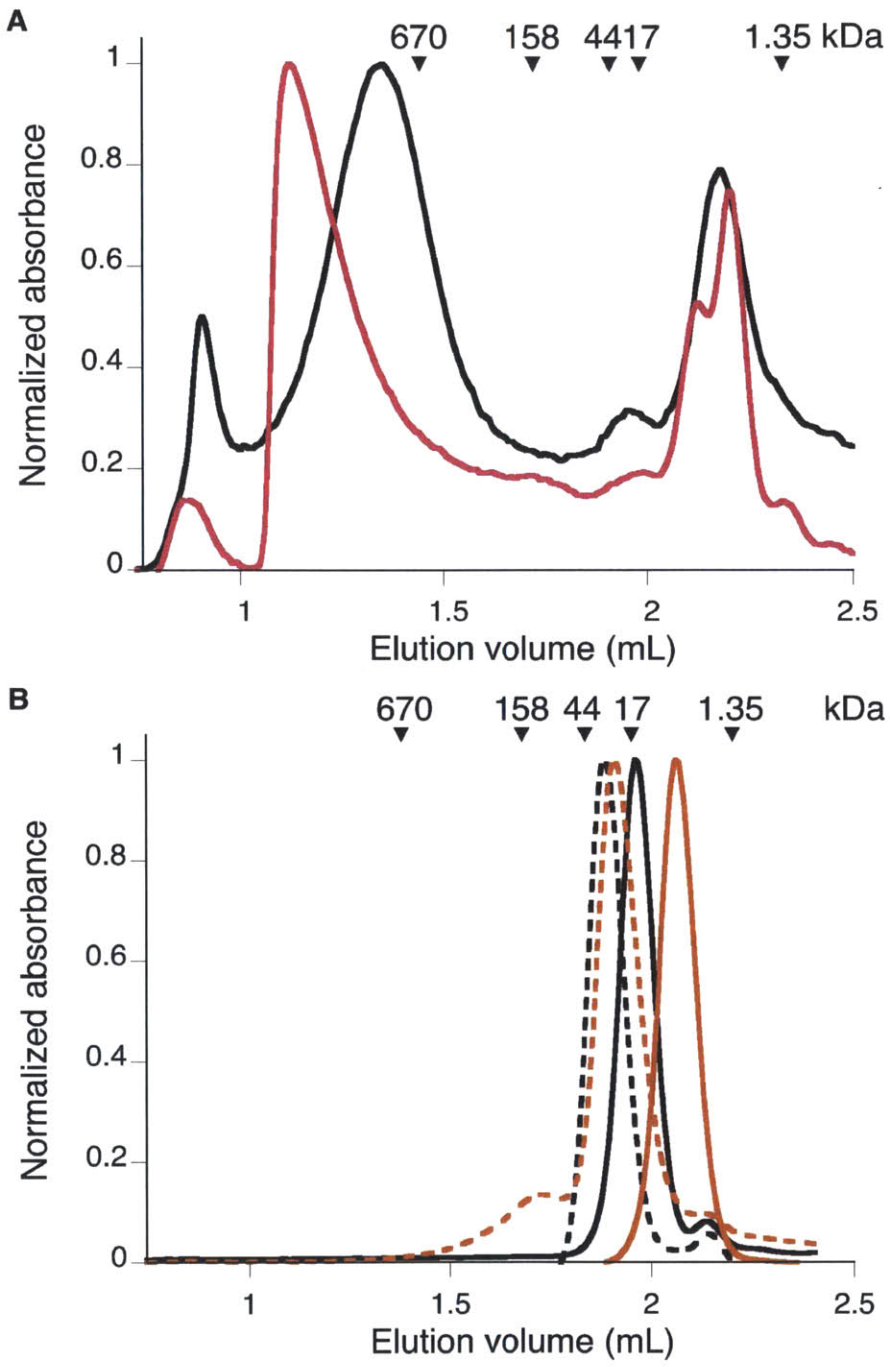

Figure 3.8 Substrates degraded poorly by Lon dodecamers form large assemblies

(A) Most IbpB (20 mM loading concentration, pink line) and FITC-casein (50 mM loading concentration, black line) eluted from a Superose 6 gel-filtration column (GE HealthCare, Uppsala Sweden) as large assemblies at positions before the largest molecular-weight standard (arrows). The column buffer was 50 mM HEPES-KOH (pH 8), 20% sucrose, 600 mM potassium glutamate, 0.1 mM TCEP. (B) Titin-I27-sul20 (black line), titin-I27^{CM}-sul20 (black dashed line), and β20-titin-I27^{CM} (orange

dashed line) eluted from a Superose 6 gel-filtration column at positions near those expected for native or denatured monomers. β20-titin-I27 (orange line) ran smaller than expected, possibly because of interactions with the column (on SDS-PAGE β20-titin-I27 ran at the expected size). For each titin variant, the loading concentration was 50 μM, and the column buffer was 50 mM HEPES-KOH (pH 7.6), 150 mM NaCl, 0.01 mM EDTA, 1 mM $MgCl_2$. Gel filtration was performed at room temperature, and molecular weight standards (Bio-Rad, Hercules, CA) were run under the same conditions as samples and eluted at positions marked by arrows.

We determined initial rates of Lon degradation of ^{35}S-labeled IbpB by assaying the production of acid-soluble peptides and normalized these rates by the total concentration of Lon subunits (Figure 3.9A). Importantly, IbpB was degraded more slowly at higher Lon concentrations, where more dodecamer was present. Figure 3.9B shows Michaelis-Menten plots for IbpB degradation using 1.5 and 6 μM Lon. K_M was similar at both Lon concentrations, but V_{max} was substantially lower at the higher Lon concentration. Fitting the Lon-concentration dependence of IbpB degradation gave a K_D of 1.8 ± 5 μM for the hexamer-dodecamer interaction, a maximum degradation rate of 0.057 ± 0.058 subunit^{-1} min^{-1} for the hexamer, and a rate of 0.00095 ± 0.021 subunit^{-1} min^{-1} for the dodecamer (Fig. 3.7B). Thus, the Lon dodecamer degrades IbpB ~60 times more slowly than the hexamer.

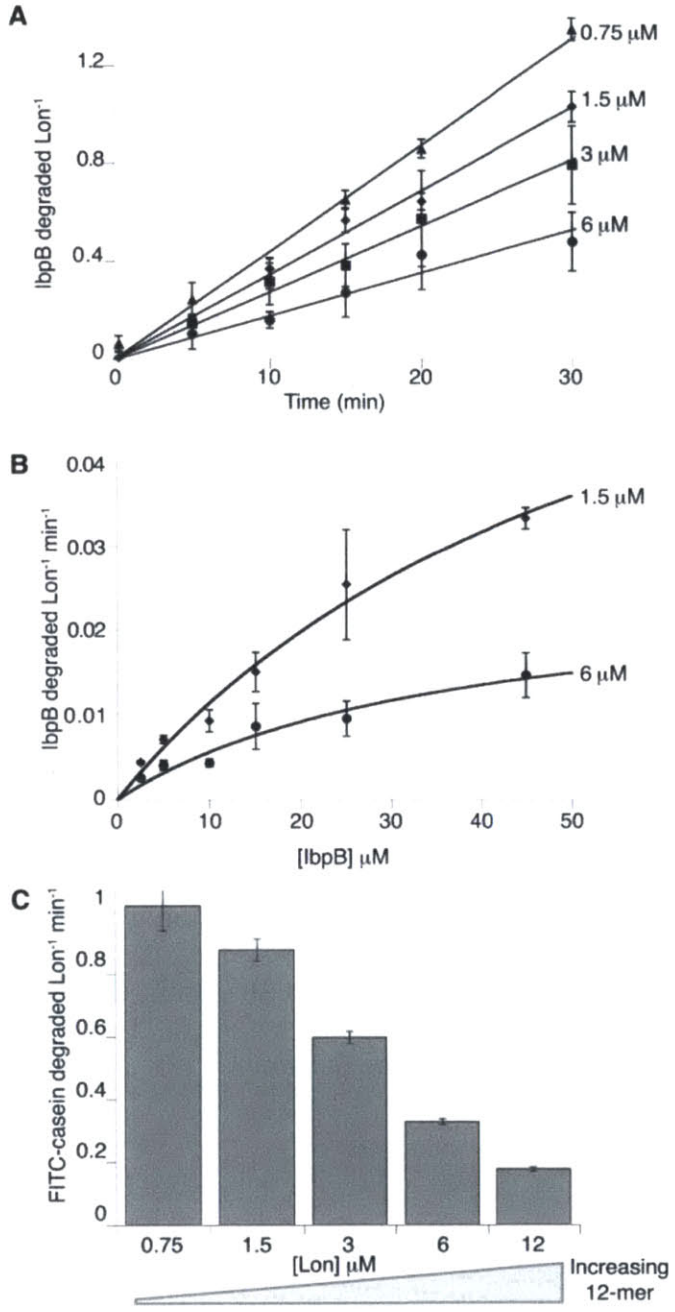

Figure 3.9 High concentrations of Lon degrade IbpB and FITC-casein less efficiently

(A) Normalized initial rates of IbpB degradation (45 μM) decreased as the Lon concentration increased. (B) Michealis-Menten plots of normalized rates of steady-state degradation of IbpB by 1.5 or 6 μM Lon. The lines are fits to the Michealis-Menten equation. (C) Normalized initial rates of degradation of FITC-β-casein (50 μM) also decreased with increasing Lon concentration. In all panels, data are plotted as averages ± 1 SEM (N = 3), and reactions were performed at 37° C and contained 4

mM ATP, 5 mM $MgCl_2$, 5 mM KCl, 2% DMSO, 360 mM potassium glutamate, 12% sucrose, and 50 mM HEPES-KOH (pH 8).

Next, we examined degradation of FITC-conjugated β-casein (monomer M_R ~25 kDa) by different concentrations of Lon. β-casein is unstructured but can form micelles and even as a monomer has a radius of gyration much larger than expected for a compact structure (Figure 3.8A and (Farrell et al. 2002; Lucey et al. 2000)). Normalized initial rates of FITC-casein degradation were determined by changes in fluorescence, and decreased substantially as the Lon concentration and dodecamer/hexamer ratio increased (Figure 3.9C). Fitting the Lon dependence of degradation gave a K_D of 1.3 ± 0.6 μM for the hexamer-dodecamer interaction, a maximum degradation rate of 1.8 ± 0.4 subunit^{-1} min^{-1} for the hexamer, and a rate of 0.00001 ± 0.038 subunit^{-1} min^{-1} for the dodecamer (Figure 3.7C). Thus, as with IbpB, the Lon dodecamer degrades this substrate far more slowly than does the Lon hexamer.

Finally, we used the appearance of acid-soluble peptides to assay Lon degradation of ^{35}S-labeled titin-I27 proteins (M_R ~12 kDa) with appended N- or C-terminal sul20 or β20 degrons (Gur and Sauer 2009; 2008). We also assayed degradation of some titin-I27 substrates following cysteine carboxymethylation, which unfolds the protein (Kenniston et al. 2003). Notably, Lon concentration had little effect on degron-tagged titin-I27 degradation, whether substrates were native or denatured or contained N-terminal or C-terminal degradation tags (Figure 3.10). Moreover, in the presence of carboxymethylated titin-I27-sul20, the rate of ATP hydrolysis by Lon showed a much smaller dependence on enzyme concentration

(Figure 3.10C), when compared with the rate of ATP hydrolysis in the absence of a protein substrate (Figure 3.6). We considered the possibility that binding of the sul20 degron to Lon might result in dodecamer dissociation. However, dodecamers were present in the EM experiments performed in the presence of sul20 peptide (Figure 3.4), and addition of this peptide did not detectably change the dodecamer/hexamer ratio in gel-filtration experiments (distributions were similar to Figure 3.2). In contrast to the case with the Ibps, the model degron substrates stimulate the rate of ATP-hydrolysis by Lon ~3 to 10-fold (Figure 3.10C). Nonetheless, the substrate-stimulated ATPase rate (normalized by Lon concentration) was lowest at the highest Lon concentrations (6 and 12 μM) (Figure 3.10C), indicating that both degron-enzyme interactions and the hexamer-dodecamer equilibrium influence ATP hydrolysis.

Taken together, our results support a model in which Lon hexamers and dodecamers are both active proteases. Importantly, however, the dodecamer only efficiently degraded the degron-tagged titin-I27 substrates, which behaved as much smaller species than the IbpB and β-casein substrates (Figure 3.8). As we discuss below, the portals created by dodecamer assembly may provide a "gating" mechanism that prevents larger substrates from entering the luminal chamber and being degraded.

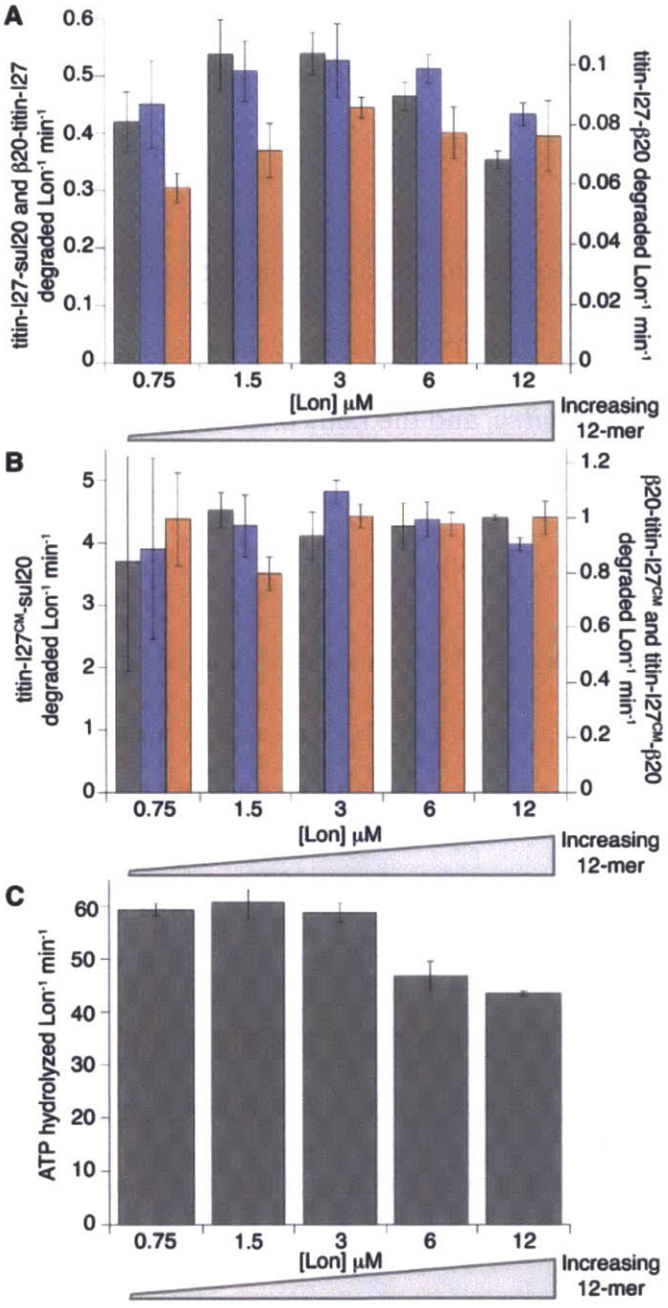

Figure 3.10 Lon hexamers and dodecamers degrade degron-tagged titin-I27 substrates with similar efficiencies

(A) Degradation of native substrates with C-terminal sul20 degrons (gray), C-terminal β20 degrons (blue), and N-terminal β20 degrons (orange). (B) Degradation of denatured titin-I27 substrates. Degron colors are the same as in panel A. (C) In the presence of carboxymethylated titin-I27-sul20 (40 μM), ATPase rates decreased only marginally at high Lon concentrations. In all panels, data are plotted as

averages ± 1 SEM (N = 3). Reactions were performed at 37° C and contained 25 mM Tris-HCl (pH 8.0), 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, and 2 mM ATP.

Discussion

Our results show that hexamers of *E. coli* Lon assemble into a dodecamer that displays different enzymatic properties. Hexamers and dodecamers are both populated at low μM concentrations *in vitro*, and the Lon concentration *in vivo* is ~2.5 μM. Thus, the hexamer-dodecamer equilibrium is likely to be a physiologically relevant factor in controlling Lon activity in cells.

Our EM structure of the Lon dodecamer reveals a face-to-face association of hexamers in which the N domains appear to be largely responsible for stabilizing the complex. This architecture positions the degradation chambers of each hexamer at the distal ends of the complex and has not been observed in other AAA+ proteases. As observed in the crystal structure of an archaeal LonB hexamer (Cha et al. 2010), the degradation chambers in the dodecamer appear to be sequestered from bulk solution. Thus, degradation still requires substrate unfolding/translocation by the Lon AAA+ ring, an apparently universal feature of AAA+ proteases (Sauer and Baker 2011).

The Lon N domains have been implicated in substrate recognition (Roudiak and Shrader 1998; Melnikov et al. 2008; Chir et al. 2009). Here we show an additional role for the N domain in dodecamer assembly. Thus, formation of the dodecamer may alter the substrate-recognition properties of the enzyme by creating or occluding substrate-binding sites. Furthermore, we find that when the N

domains interact with each other, and/or with the ATPase domain of the opposite hexamer in the dodecamer, the rate of ATP hydrolysis by Lon can be suppressed. A similar suppression of ATPase activity is observed when ClpX interacts ClpP (Joshi et al. 2004). Understanding the molecular basis of the suppression of ATPase activity upon dodecamer assembly, and the activation of ATP hydrolysis by some Lon substrates is likely to provide important insight into the allosteric mechanisms that are used by substrates to control Lon's enzymatic activities (Waxman and Goldberg 1986; Gur and Sauer 2009).

A notable feature of the dodecamer is the presence of six portals, each ~45 Å in diameter, spaced around the equator of the structure. The size of these portals should prevent entry of large substrates into the lumen of the dodecamer (Figure 3.11). Indeed, we found that IbpB and β-casein substrates, which have large radii of gyration, were degraded very slowly by Lon dodecamers when compared with hexamers. In fact, given the confidence of data fitting, we cannot exclude the possibility that the dodecamer cannot degrade IbpB or β-casein. By contrast, we found that smaller degron-tagged titin-I27 substrates in both folded and unfolded states were degraded at similar rates by Lon hexamers and dodecamers. These results support a model in which a major determinant of substrate degradation by the Lon dodecamer is the ability of a substrate to diffuse or be pulled through the portals to allow engagement by the degradation machinery. In principle, substrate size, charge, hydrophobicity, stability, or the length of the disordered degron tag could all contribute to determining whether specific substrates could enter the lumen and be degraded by Lon dodecamers.

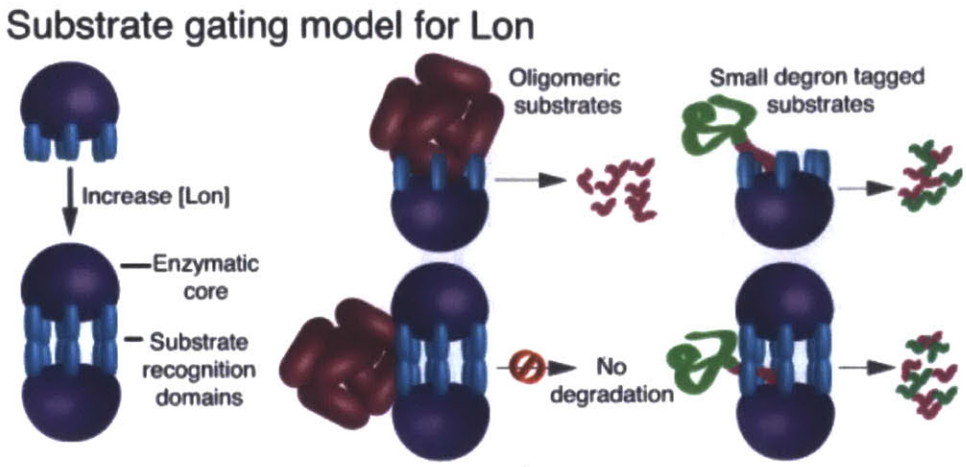

Figure 3.11 Substrate gating model for Lon

Potential roles of Lon hexamers and dodecamers in intracellular protein degradation.

At present, we can only speculate about possible biological functions for the Lon dodecamer. One possibility is that cellular stress results in very high concentrations of misfolded proteins that form large aggregates or bind the IbpA/IbpB chaperones, with degradation of such substrates by Lon hexamers swamping the enzyme's ability to degrade regulatory proteins and therefore inhibiting the recovery from stress. These regulatory proteins include SulA (19 kDa, C-terminal degron), which is rapidly degraded during recovery after DNA damage to allow cell division to resume (Gottesman et al. 1981; Mizusawa and Gottesman 1983). Similarly, SoxS (13 kDa, N-terminal degron), a transcriptional activator of the superoxide-response regulons (Shah and Wolf 2006), and UmuD (30 kDa dimer, N-terminal degron), a subunit of the lesion-bypass DNA polymerase (Gonzalez et al. 1998) need to be degraded by Lon specifically as cells recover from stress. In this scenario, Lon dodecamers could carry out regulatory degradation, while hexamers

triage and perform quality-control degradation of damaged/unfolded proteins. It is also possible, that dodecamers are more highly populated under stress conditions, reducing the rate of IbpA/IbpB degradation and increasing their concentration to suppress aggregation and promote refolding in cooperation with ClpB and DnaK (Mogk et al. 2003a; 2003b). Alternatively, equatorial interfaces in Lon dodecamers, which are absent in hexamers, might allow recognition and degradation of additional important substrates. ATP is abundant during exponential growth but scarce when nutrients are lacking. In principle, Lon dodecamer formation might serve to conserve ATP, as we find that basal hydrolysis by the dodecamer is much lower than by the hexamer. Thus, if Lon were predominantly dodecameric during stationary-phase growth, when substrates became available they could still be degraded to provide amino acids or for regulatory purposes but at a lower cost in terms of overall ATP consumption. Testing these models will require the design of Lon variants that are exclusively hexameric or dodecameric. We are pursuing higher resolution structures of the dodecamer to guide these efforts. *B. subtilis* and *Mycobacterium smegmatis* Lon have also been reported to form oligomers larger than hexamers (Roudiak et al. 1998; Duman and Löwe 2010; Rudyak et al. 2001). Thus, hexamer-dodecamer equilibria may be a relatively conserved feature of this important intracellular protease.

Acknowledgements

We thank S. Bissonnette and E. Gur for reagents, members of our labs for helpful

discussions, D. Pheasant at the MIT Biophysical Instrumentation Center for help with the AUC experiments, and the electron-microscopy facility at Harvard Medical School, which is supported by NIH grant PO1 GM62580. This work was supported by NIH grants GM-49224 and AI-16892, an NIH NRSA postdoctoral fellowship (F32GM094994) to E.F.V., and an NSF Graduate Research Fellowship to M.L.W. T.A.B. is an employee of the Howard Hughes Medical Institute.

References

Ali Azam T, Iwata A, Nishimura A, Ueda S, Ishihama A. 1999. Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J Bacteriol* **181**: 6361–6370.

Bender T, Lewrenz I, Franken S, Baitzel C, Voos W. 2011. Mitochondrial enzymes are protected from stress-induced aggregation by mitochondrial chaperones and the Pim1/LON protease. *Mol Biol Cell* **22**: 541–554.

Bernstein SH, Venkatesh S, Li M, Lee J, Lu B, Hilchey SP, Morse KM, Metcalfe HM, Skalska J, Andreeff M, et al. 2012. The mitochondrial ATP-dependent Lon protease: a novel target in lymphoma death mediated by the synthetic triterpenoid CDDO and its derivatives. *Blood*.

Bissonnette SA, Rivera-Rivera I, Sauer RT, Baker TA. 2010. The IbpA and IbpB small heat-shock proteins are substrates of the AAA+ Lon protease. *Mol Microbiol*.

Botos I, Melnikov EE, Cherry S, Tropea JE, Khalatova AG, Rasulova F, Dauter Z, Maurizi MR, Rotanova TV, Wlodawer A, et al. 2004. The catalytic domain of Escherichia coli Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J Biol Chem* **279**: 8140–8148.

Brown PH, Schuck P. 2006. Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys J* **90**: 4651–4661.

Cha S-S, An YJ, Lee C-R, Lee HS, Kim Y-G, Kim SJ, Kwon KK, De Donatis GM, Lee J-H, Maurizi MR, et al. 2010. Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *EMBO J* **29**: 3520–3530.

Chir J-L, Liao J-H, Lin Y-C, Wu S-H. 2009. The N-terminal sequence after residue 247 plays an important role in structure and function of Lon protease from

Brevibacillus thermoruber WR-249. *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS* **382**: 762–765.

Duman RE, Löwe J. 2010. Crystal structures of Bacillus subtilis Lon protease. *J Mol Biol* **401**: 653–670.

Farrell HM, Qi PX, Brown EM, Cooke PH, Tunick MH, Wickham ED, Unruh JJ. 2002. Molten globule structures in milk proteins: implications for potential new structure-function relationships. *J Dairy Sci* **85**: 459–471.

Fredriksson A, Ballesteros M, Dukan S, Nystrom T. 2005. Defense against protein carbonylation by DnaK/DnaJ and proteases of the heat shock regulon. *J Bacteriol* **187**: 4207–4213.

Goff SA, Casson LP, Goldberg AL. 1984. Heat shock regulatory gene htpR influences rates of protein degradation and expression of the lon gene in Escherichia coli. *Proc Natl Acad Sci USA* **81**: 6647–6651.

Goff SA, Goldberg AL. 1985. Production of abnormal proteins in E. coli stimulates transcription of lon and other heat shock genes. *Cell* **41**: 587–595.

Goldberg AL, Moerschell RP, Chung CH, Maurizi MR. 1994. ATP-dependent protease La (lon) from Escherichia coli. *Methods in Enzymology* **244**: 350–375.

Gonzalez M, Frank EG, Levine AS, Woodgate R. 1998. Lon-mediated proteolysis of the Escherichia coli UmuD mutagenesis protein: in vitro degradation and identification of residues required for proteolysis. *Genes Dev* **12**: 3889–3899.

Gottesman S, Halpern E, Trisler P. 1981. Role of sulA and sulB in filamentation by lon mutants of Escherichia coli K-12. *J Bacteriol* **148**: 265–273.

Gottesman S, Roche E, Zhou Y, Sauer RT. 1998. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev* **12**: 1338–1347.

Gur E, Sauer RT. 2009. Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine. *Proceedings of the National Academy of Sciences* **106**: 18503–18508.

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

Henderson R, Chen S, Chen JZ, Grigorieff N, Passmore LA, Ciccarelli L, Rubinstein JL, Crowther RA, Stewart PL, Rosenthal PB. 2011. Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *J Mol Biol* **413**: 1028–1046.

Jiao W, Qian M, Li P, Zhao L, Chang Z. 2005. The essential role of the flexible termini in the temperature-responsiveness of the oligomeric state and chaperone-like activity for the polydisperse small heat shock protein IbpB from Escherichia coli. *J Mol Biol* **347**: 871–884.

Joshi SA, Hersch GL, Baker TA, Sauer RT. 2004. Communication between ClpX and ClpP during substrate processing and degradation. *Nat Struct Mol Biol* **11**: 404–411.

Kenniston JA, Baker TA, Fernandez JM, Sauer RT. 2003. Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* **114**: 511–520.

Kowit JD, Goldberg AL. 1977. Intermediate steps in the degradation of a specific abnormal protein in Escherichia coli. *J Biol Chem* **252**: 8350–8357.

Lindsley JE. 2001. Use of a real-time, coupled assay to measure the ATPase activity of DNA topoisomerase II. *Methods Mol Biol* **95**: 57–64.

Lucey JA, Srinivasan M, Singh H, Munro PA. 2000. Characterization of commercial and experimental sodium caseinates by multiangle laser light scattering and size-exclusion chromatography. *J Agric Food Chem* **48**: 1610–1616.

Melnikov EE, Andrianova AG, Morozkin AD, Stepnov AA, Makhovskaya OV, Botos I, Gustchina A, Wlodawer A, Rotanova TV. 2008. Limited proteolysis of E. coli ATP-dependent protease Lon - a unified view of the subunit architecture and characterization of isolated enzyme fragments. *Acta Biochim Pol* **55**: 281–296.

Mizusawa S, Gottesman S. 1983. Protein degradation in Escherichia coli: the lon gene controls the stability of sulA protein. *Proc Natl Acad Sci USA* **80**: 358–362.

Mogk A, Deuerling E, Vorderwülbecke S, Vierling E, Bukau B. 2003a. Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Mol Microbiol* **50**: 585–595.

Mogk A, Schlieker C, Friedrich KL, Schönfeld H-J, Vierling E, Bukau B. 2003b. Refolding of substrates bound to small Hsps relies on a disaggregation reaction mediated most efficiently by ClpB/DnaK. *J Biol Chem* **278**: 31033–31042.

Ngo JK, Davies KJA. 2009. Mitochondrial Lon protease is a human stress protein. *Free Radic Biol Med* **46**: 1042–1048.

Ngo JK, Pomatto LCD, Bota DA, Koop AL, Davies KJA. 2011. Impairment of lon-induced protection against the accumulation of oxidized proteins in senescent wi-38 fibroblasts. *J Gerontol A Biol Sci Med Sci* **66**: 1178–1185.

Nørby JG. 1988. Coupled assay of Na+,K+-ATPase activity. *Methods in Enzymology*

156: 116–119.

Park S-C, Jia B, Yang J-K, Van DL, Shao YG, Han SW, Jeon Y-J, Chung CH, Cheong G-W. 2006. Oligomeric structure of the ATP-dependent protease La (Lon) of Escherichia coli. *Mol Cells* **21**: 129–134.

Phillips TA, VanBogelen RA, Neidhardt FC. 1984. lon gene product of Escherichia coli is a heat-shock protein. *J Bacteriol* **159**: 283–287.

Rosen R, Biran D, Gur E, Becher D, Hecker M, Ron EZ. 2002. Protein aggregation in Escherichia coli: role of proteases. *FEMS Microbiology Letters* **207**: 9–12.

Roudiak SG, Seth A, Knipfer N, Shrader TE. 1998. The lon protease from Mycobacterium smegmatis: molecular cloning, sequence analysis, functional expression, and enzymatic characterization. *Biochemistry* **37**: 377–386.

Roudiak SG, Shrader TE. 1998. Functional role of the N-terminal region of the Lon protease from Mycobacterium smegmatis. *Biochemistry* **37**: 11255–11263.

Rudyak SG, Brenowitz M, Shrader TE. 2001. Mg2+-linked oligomerization modulates the catalytic activity of the Lon (La) protease from Mycobacterium smegmatis. *Biochemistry* **40**: 9317–9323.

Sauer RT, Baker TA. 2011. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **80**: 587–612.

Sauer RT, Bolon DN, Burton BM, Burton RE, Flynn JM, Grant RA, Hersch GL, Joshi SA, Kenniston JA, Levchenko I, et al. 2004. Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell* **119**: 9–18.

Shah IM, Wolf RE. 2006. Sequence requirements for Lon-dependent degradation of the Escherichia coli transcription activator SoxS: identification of the SoxS residues critical to proteolysis and specific inhibition of in vitro degradation by a peptide comprised of the N-terminal 21 amino acid residues. *J Mol Biol* **357**: 718–731.

Shearstone JR, Baneyx F. 1999. Biochemical characterization of the small heat shock protein IbpB from Escherichia coli. *J Biol Chem* **274**: 9937–9945.

Shineberg B, Zipser D. 1973. The ion gene and degradation of beta-galactosidase nonsense fragments. *J Bacteriol* **116**: 1469–1471.

Starkova NN, Koroleva EP, Rumsh LD, Ginodman LM, Rotanova TV. 1998. Mutations in the proteolytic domain of Escherichia coli protease Lon impair the ATPase activity of the enzyme. *FEBS Lett* **422**: 218–220.

van Dijl JM, Kutejová E, Suda K, Perecko D, Schatz G, Suzuki CK. 1998. The ATPase

and protease domains of yeast mitochondrial Lon: roles in proteolysis and respiration-dependent growth. *Proc Natl Acad Sci USA* **95**: 10584–10589.

Van Melderen L, Aertsen A. 2009. Regulation and quality control by Lon-dependent proteolysis. *Res Microbiol.*

Van Melderen L, Gottesman S. 1999. Substrate sequestration by a proteolytically inactive Lon mutant. *Proc Natl Acad Sci USA* **96**: 6064–6071.

Venkatesh S, Lee J, Singh K, Lee I, Suzuki CK. 2011. Multitasking in the mitochondrion by the ATP-dependent Lon protease. *Biochim Biophys Acta.*

Waxman L, Goldberg AL. 1986. Selectivity of intracellular proteolysis: protein substrates activate the ATP-dependent protease (La). *Science* **232**: 500–503.

Chapter 4

Roles of the N domain of the AAA+ Lon protease in substrate recognition, allosteric regulation, and chaperone activity

Abstract

A wide spectrum of intracellular substrates are targeted to the AAA+ Lon protease of *E. coli* by specific degrons. These sequences bind Lon and regulate rates of ATP hydrolysis and proteolysis. Little is known about how any degron binds Lon. Here, we show that a degron comprising the 20 C-terminal residues of the cell-division inhibitor SulA binds to the N domain of Lon and identify N-domain mutations that weaken this binding and alter degradation and allosteric activation by sul20-tagged substrates but not other substrate classes. In addition, residues in the Lon axial pore are required for efficient degradation and robust stimulation of ATP hydrolysis by sul20-tagged substrates, as well as substrates bearing different degrons. Experiments *in vivo* suggest that simple binding to the N domain of Lon is sufficient to inactivate SulA, whereas Lon-mediated relief of proteotoxic stress requires substrate translocation and remodeling but does not require degradation. These results support a model in which Lon can function as a protease or as a remodeling chaperone, with degron binding playing an important role in partitioning substrates between these alternative fates.

Introduction

An *Escherichia coli* cell contains more than 4000 different proteins, with copy numbers ranging from a few to greater than 50,000 (Kitagawa et al. 2005; Moran et al. 2010; Bakshi et al. 2012). Under conditions that result in protein misfolding, about half of cytosolic protein degradation in *E. coli* is carried out by the AAA+ Lon protease (Chung and Goldberg 1981). Lon appears to recognize damaged proteins by binding to hydrophobic residues exposed as a consequence of unfolding or misfolding. Lon also degrades native proteins, including the SulA cell-division inhibitor (Higashitani et al. 1997; Gur and Sauer 2008). Lon-family proteases are present in most bacteria, in archaea, and in endosymbiotic organelles (Van Melderen and Aertsen 2009). Lon is necessary for rapid progression through the cell cycle or for full pathogenicity in some bacteria (Wright et al. 1996; Robertson et al. 2000; Ingmer and Brøndsted 2009; Breidenstein et al. 2012b; 2012a; Gora et al. 2013). Overexpression of mitochondrial Lon increases fungal lifespan, whereas knockdown of mitochondrial Lon is toxic to lymphoma cells, which are presumably subject to increased oxidative and proteotoxic stress (Luce and Osiewacz 2009; Luo et al. 2009; Bernstein et al. 2012).

Like other AAA+ proteases, Lon sequesters its proteolytic active sites within a barrel-like chamber, uses a hexameric ring and ATP hydrolysis to unfold and translocate proteins through a narrow axial pore into this chamber, and recognizes substrates predominantly by binding to degrons or peptide tags (Baker and Sauer 2006; Cha et al. 2010). Unlike many AAA+ proteases, however, the AAA+ ATPase module and protease domain of Lon are part of a single polypeptide, and degron

binding regulates Lon activity in addition to serving a recognition function. For example, when otherwise identical proteins are tagged with either the sul20 or β20 degron, which correspond respectively to the C-terminal 20 residues of SulA and a β-galactosidase sequence buried in the native protein, the maximal rate of *E. coli* Lon degradation can differ almost 10-fold (Higashitani et al. 1997; Ishii and Amano 2001; Gur and Sauer 2008; 2009). These results suggest that degron binding can shift Lon into conformations with higher or lower protease activity. A Lon conformation with no protease activity has been proposed to unfold misfolded substrates by ATP-dependent translocation, allowing them a chance to refold properly after release (Gur and Sauer 2009).

In addition to its AAA+ module and peptidase domain, *E. coli* Lon contains a family specific N domain, which is necessary but not sufficient for hexamerization (Lee et al. 2004b; Melnikov et al. 2008). Crystal structures are known for parts of the N domain, but none are in a hexameric or dodecameric oligomeric state. (Duman and Löwe 2010; Li et al. 2010; 2005). The N domain is thought to be involved in substrate binding (Ebel et al. 1999; Roudiak and Shrader 1998; Rudyak and Shrader 2000; Lee et al. 2004b; Adam et al. 2012). Consistently, N-domain mutations or truncations result in defects in Lon activity *in vitro* (Cheng et al. 2012), but these results could be explained by hexamer destabilization, and substrate binding to the N domain has not been directly demonstrated.

Here, we show that the sul20 degron binds to the Lon N domain, identify N-domain mutations that define the binding site, and demonstrate that sul20 binding to this site plays a role in allosteric activation of Lon protease activity. Residues in

the axial pore of the Lon hexamer are also required for degradation of sul20- or β-tagged substrates and for robust stimulation of ATP hydrolysis by these substrates. Experiments *in vivo* suggest that simple binding to Lon is sufficient to inactivate SulA, whereas Lon-mediated relief of proteotoxic stress requires substrate translocation and remodeling but not degradation. In combination, these results suggest that Lon can function as a protease or as a remodeling chaperone, with degron binding to the N domain playing an important role in how substrates partition between these alternative fates.

Materials and Methods

Protein cloning, expression, and purification

Variants of *E. coli* Lon were cloned into pBAD33. For heat-shock assays, the chloramphenicol resistance marker of pBAD33 was replaced with an ampicillin resistance marker cloned from pSH21. *E. coli* $ClpX^{\Delta N}$ and chimeras were cloned into HTUA vector and contained an N-terminal His_6 tag followed by a TEV protease site. $Chimera^{+307}$ contained Lon residues 1-307, a two residue scar (EL, resulting from cloning into a Sac I restriction site), and $ClpX^{\Delta N}$ (residues 62-424 of wild-type ClpX). $Chimera^{+211}$ contained Lon residues 1-211, a GSSG linker, the EL dipeptide, and $ClpX^{\Delta N}$. In addition, $chimera^{+211}$ contained the C39S Lon mutation and C169S ClpX mutation to remove exposed cysteines and the ClpX T66C and P388C mutations to form inter-subunit disulfide bonds to stabilize hexamer formation (Glynn et al. 2012). ClpP was cloned into a pET22b vector with a His_6 tag on the C-terminus. $Titin^{I27}$ variants were cloned into a pSH21 vector with an N-terminal His_6 tag. b20-

cp6-GFP and cp6-GFP-sul20 were cloned into a pCOLADuet1 vector with an N-terminal His_6 tag followed by a PreScission protease site. Mutations were generated either by QuickChange PCR (Stratagene) or by standard PCR techniques.

Lon was over-expressed with slight modifications from a method described previously (Wohlever et al. 2013). Briefly, cells were grown at 37 °C until OD_{600} = 1.0, induced with 0.2% arabinose at 37 °C for 3.5 h, harvested, and resuspended in LBA [100 mM potassium phosphate (pH 6.5), 1 mM DTT, 1 mM EDTA, and 10% glycerol] to a final volume of 20 mL. Cells were incubated with lysozyme before sonication, and the crude cell lysate was cleared by high-speed centrifugation. The cleared lysate was incubated on ice for 20 min with 2 μL of benzonase (250 U/mL, Sigma) and then bound to P11 phosphocellulose resin (Whatman) equilibrated in LBA buffer. This resin was washed twice with LBA and once with LBA plus 100 mM potassium phosphate (pH 6.5). Lon was eluted from the P11 resin using LBA buffer plus 300 mM potassium phosphate (pH 6.5). The eluant was filtered to remove phosphocellulose, polyethyleneimine (PEI) was added to a final concentration of 0.12% to precipitate nucleic acids, additional phosphocellulose was added to remove excess PEI, and the mixture was filtered, concentrated, and chromatographed on an S200 column (GE Healthcare) equilibrated in 50 mM HEPES (pH 7.5), 2 M NaCl, AND 1 mM DTT. Peak fractions from this column were pooled, buffer exchanged into storage buffer (50 mM HEPES (pH 7.5), 150 mM NaCl, 10 μM EDTA, 1 mM DTT, and 10% glycerol) and frozen at -80° C.

ClpP, cp6-GFP-sul20, b20-cp6-GFP, and $titin^{I27}$ variants were expressed, purified, and carboxymethylated (if applicable) as described (Glynn et al. 2012; Gur

and Sauer 2009; Wohlever et al. 2013). For ^{35}S-labeling, cells were grown in a rich defined medium lacking methionine (TekNova) until OD_{600} = 0.6 , and ^{35}S-methionine (Perkin-Elmer) was added after 20 min of induction with 1 mM IPTG. ^{35}S-labeled proteins were purified through the Ni-NTA step and then mixed at a 1:19 ratio with purified unlabeled substrate.

Cells expressing $ClpX^{\Delta N}$ and chimeras were grown until OD_{600} = 1.0, induced with 1 mM IPTG for 3.5 h at room temperature, harvested, resuspended in lysis buffer [25 mM HEPES (pH 7.5), 400 mM NaCl, 100 mM KCl, 20 mM imidazole, 10% glycerol, and 10 mM 2-mercaptoethanol] to a total volume of 20 mL, and lysed by incubation with lysozyme and sonication. Following lysis, 2 μL of benzonase (250 U/mL, Sigma) and PMSF (final concentration 1 mM) were added, the lysate was cleared by high-speed centrifugation, and the supernatant was bound to Ni-NTA resin equilibrated in lysis buffer. The resin was washed with 30 mL of lysis buffer and eluted with lysis buffer plus 250 mM imidazole. For $ClpX^{\Delta N}$ and chimera^{+307}, the eluant was chromatographed on S300 column (GE Healthcare) equilibrated in 50 mM Tris (pH 8.0), 300 mM KCl, 1 mM DTT, and 10% glycerol. Appropriate fractions were pooled, concentrated, and frozen at -80° C. After elution of chimera^{+211} from the Ni-NTA resin, the protein was buffer exchanged into low-salt buffer [25 mM HEPES (pH 7.5), 50 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 10% glycerol], incubated with 1 μM TEV protease for 90 min at room temperature to remove the His_6 tag, chromatographed on an S300 column as described above, and treated with copper phenanthroline to catalyze disulfide-bond formation between subunits as described (Glynn et al. 2012). The disulfide-bonded chimera^{+211} was purified on a Superose 6

column, concentrated, and frozen at -80° C as described above, except in a buffer lacking DTT.

Peptides

Peptides were synthesized in house, purified by reverse-phase HPLC, and the expected masses were verified by mass spectrometry. The F-β20-Q peptide (sequence Z-QLRSLNGEWRFAWFPAPEAV-nY-A, where Z is a para-aminobenzoic acid fluorophore and nY is a nitrotyrosine quencher) was dissolved in dimethylsulfoxide and concentration was determined by absorbance (ε_{381} = 2200 M^{-1} cm^{-1}). The sul20 peptide (sequence ASSHATRQLSGLKIHSNLYH) was dissolved in 50 mM HEPES (pH 7.5) and concentration was measured by absorbance (ε_{280} = 1490). The sul20 peptide with an N-terminal fluorescein was dissolved in 25 mM Tris (pH 8.0) and concentration was determined by absorbance (ε_{495} = 83,397 M^{-1} cm^{-1}).

Biochemical assays

Unless noted, biochemical assays were performed in 25 mM Tris (pH 8.0), 100 mM KCl, 10 mM $MgCl_2$, at 37 °C using enzyme concentrations calculated for hexamer equivalents. Kinetic and anisotropy assays were performed in a SpectraMax M5 plate reader using 384-well clear plates (Corning) for absorbance assays and 96-well flat bottom ½-area plates (Corning) for fluorescence assays. ATPase assays contained supplemental 5 mM DTT, 2 mM ATP, lactate dehydrogenase (10 U/mL), and an ATP regeneration system [rabbit muscle

pyruvate kinase (Sigma, 10 U/mL), 20 mM phosphoenolpyruvate (Sigma)]. The rate of ATP hydrolysis was measured by monitoring changes in absorbance at 340 nm, and reactions were initiated by the addition of $MgCl_2$ that had been pre-warmed to 37 °C. Degradation assays contained supplemental 1 mM DTT, an ATP regeneration system, and 2 mM ATP, which was used to initiate the reaction. Fluorescent substrates were incubated in plate reader until the fluorescence was constant prior to initiation of degradation. For degradation assays monitored by SDS-PAGE, 10 μL aliquots were taken at specified time points and mixed with 3.3 μL of 4X loading buffer [8% SDS, 250 mM Tris (pH 6.8), 40% glycerol, 160 mM DTT, and bromophenol blue]. The rate of degradation of ^{35}S-labeled titin^{I27}-sul20 variants was determined by measuring the amount of soluble radioactive products following precipitation with ice-cold trichloroacetic acid (Gottesman et al. 1998). The binding of fluorescent sul20 peptide to the proteolytically inactive Lon^{S679A} variant was measured in the presence of 1 mM ATPγS to prevent translocation; fluorescence anisotropy values were corrected for G-factor and scattering and fitted to a hyperbolic equation to determine a K_D value.

Cross-linking and mass spectrometry

Reactions were performed in the dark until the photo-activation step. The sul20 peptide (1 mM in 50 mM HEPES (pH 7.5), 150 mM NaCl, 10 μM EDTA, 10% glycerol) was incubated with 1 mM Sulfo-SBED (Pierce) for 30 min at room temperature, precipitated material was removed by centrifugation, and unreacted crosslinker was removed by dialysis using a 2 kDa MWCO membrane. The

crosslinker-modified sul20 peptide (200 μM) was incubated with 10 μM Lon^{S679A} hexamer, 1 mM ATPγS, and 1 mM $MgCl_2$ at room temperature for 5 min. Crosslinking was initiated by UV irradiation (365 nm) with a handheld lamp at a distance of 2 cm for 15 min. To reduce the disulfide bond linking the sul20 peptide to the crosslinker and Lon, 100 mM 2-mercaptoethanol was added and the reaction was incubated at room temperature for 1 h. Free sul20 peptide was removed by two consecutive microbio spin columns (BioRad). Labeling of Lon was verified by western blotting with an anti-biotin antibody. The modified Lon protein was digested with sequencing grade trypsin (Roche) using a 1:100 enzyme:substrate ratio at 37 °C for 14 h, and cleavage was quenched with 1 mM TLCK. Biotinylated peptides were enriched by passage over a Monomeric Avidin Resin (Pierce) and were eluted from this column with 100 mM glycine buffer (pH 2.8). Samples were analyzed by nanospray LC-MS using a QSTAR Elite quadrupole-time-of-flight mass spectrometer. Samples were loaded onto a reverse phase protein trap, which was desalted on-line and eluted isocratically. Deconvolution of the electrospray data to generate molecular-weight spectra was performed with the BioAnalyst software included with the QSTAR Elite data system.

Biological assays

For *in vivo* degradation of SulA, W3110 *lon*::Kan^R cells were transformed with pBAD33 vectors containing wild-type Lon, Lon^{33-35}, or the empty pBAD33 vector and grown in LB until a final OD_{600} = 0.9 – 1.3. Cultures were diluted into fresh LB to give a final OD_{600} = 0.25. 10 μL of 10x serial dilutions were spotted on LB plates

containing 25 μg/mL Kanamycin and 10 μg/mL chloramphenicol. Plates were exposed to 254 nm UV light from a handheld lamp at a distance of 5 cm for 10 seconds, and then incubated in the dark at 37° C overnight.

For proteotoxic stress assays, strain BB7357 (*clpXP-lon*::Cm^R, $P_{A1\text{-}lacO\text{-}1}$ *dnaK, J lacI*^{q}) (Tomoyasu et al. 2001) were transformed with ampicillin resistant pBAD33 vectors containing wild-type Lon, $Lon^{33\text{-}35}$, Lon^{E240K}, Lon^{S679A}, or the empty vector control. Cultures were grown in LB with 1 mM IPTG at 30° C until late-log phase, diluted to a final OD_{600} = 0.1, and serially diluted 5-fold in LB. Cultures were then spotted onto LB plates with 25 μM IPTG, 25 μg/mL Kanamycin, 10 μg/mL Chloramphenicol, and 100 μg/mL Ampicillin and grown at 30° C or 42° C.

For western blots, wild-type W3110 cells or W3110 *lon*::Kan^R cells with pBAD33 plasmids with Lon variants were grown at 37° C in LB until OD_{600} = 0.3, at which point 200 μL aliquots were taken, pelleted, decanted, and stored at -20° C. Frozen pellets were resuspended in a final volume of 40 uL of 4x SDS PAGE Loading Buffer, boiled at 100° C for 10 min. and then 6 uL were loaded onto a 4% - 20% polyacrylamide gradient gel. The gels were transferred onto filter paper using a semi-wet transfer apparatus (Bio-Rad), probed with anti-Lon polyclonal antibody (produced by Covance Research Products) at a 1:2,000 dilution for 1.5 h at room temperature, incubated with ECL anti-rabbit IgG-HRP conjugate (Amersham Biosciences) for 1 h, and developed with ECL Plus Western Blotting Detection System (Amersham Biosciences).

Results

The Lon N domain binds the sul20 degron

We sought to test if the sul20 degron binds to a site in the N domain of *E. coli* Lon. However, N-domain fragments do not form stable hexamers (Lee et al. 2004b; Li et al. 2010), raising potential problems if substrate binding requires hexamerization or if interactions with hydrophobic surfaces normally buried in subunit-subunit interfaces create spurious non-specific binding. To circumvent these problems, we fused the Lon N domain to *E. coli* ClpX$^{\Delta N}$, a AAA+ enzyme that forms stable ring hexamers (Figure 4.1A). Chimera^{+307} contained the entire Lon N domain (residues 1-307) fused to ClpX$^{\Delta N}$, whereas chimera^{+211} contained the first 211 N-domain residues (Figure 4.1B). Both Lon-N-ClpX$^{\Delta N}$ chimeras were highly soluble, hydrolyzed ATP, associated with the ClpP peptidase (the normal proteolytic partner of ClpX), and degraded an ssrA-tagged ClpXP substrate (Figure 4.2).

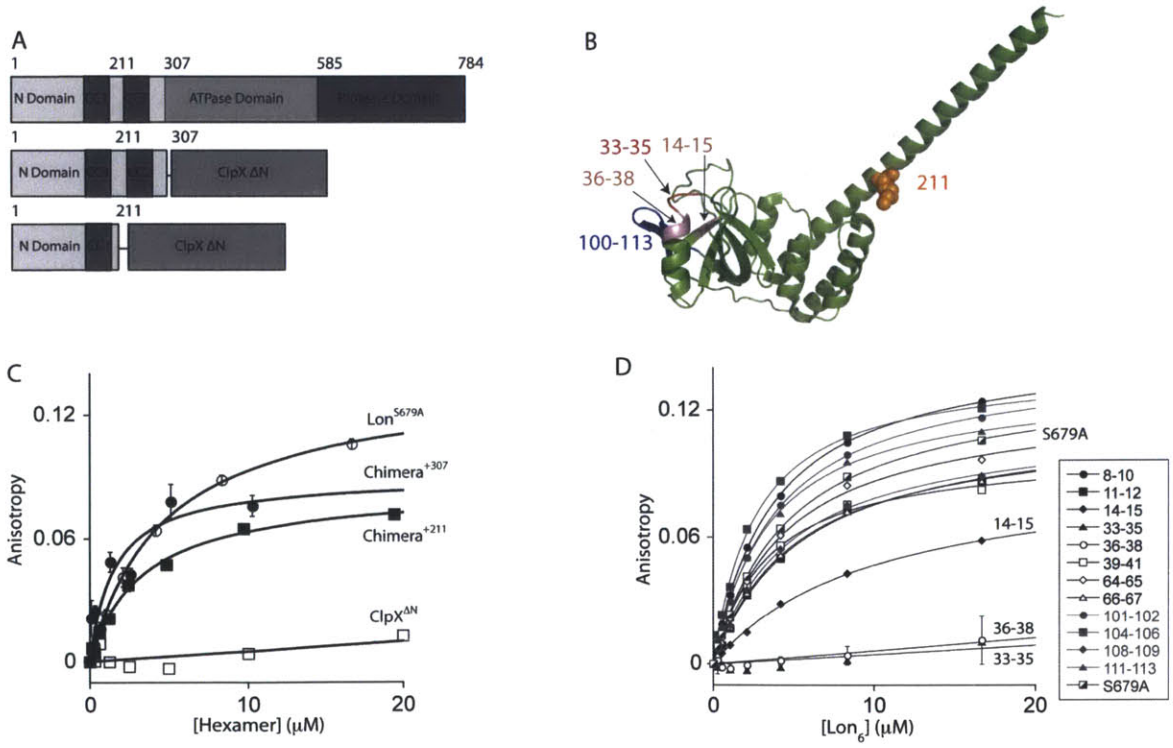

Figure 4.1 Identification of a sul20-binding site in the Lon N domain
(A) Domain organization of *E. coli* Lon, chimera^{+211} and chimera^{+307}. The two coiled-coil regions in the N domain are labeled as CC1 and CC2. Chimera^{+307} contains a two residue cloning-scar between the N domain and ClpX$^{\Delta N}$ whereas a short GSSG linker and a two residue cloning scar connects residues 1-211 of Lon to ClpX$^{\Delta N}$ in chimera^{+211}. (B) Location of key residues in the N domain (3LJC.pdb). Residues 14-15 and 33-38 form a binding pocket for the sul20 degron in close proximity to residues 100-113, which was modified by cross-linking. (C) Binding of fluorescein-labeled sul20 peptide (200 nM, excitation 494 nm, emission 521 nm) to Lon^{S679A}, chimera^{+211}, chimera^{+307}, and ClpX$^{\Delta N}$ was assayed by changes in anisotropy. Error bars are standard error of the mean, n ≥ 2. All values were baseline corrected by subtracting the anisotropy at 0 μM hexamer. Solid lines are fits to the equation Anisotropy = (A_{Max} * [hexamer]) / (K_D + [hexamer]). (D) Binding of fluorescein-labeled sul20 peptide to variants of Lon with different alanine-scan mutations. Binding conditions and baseline correction were the same as in Figure 4.1C. For clarity, error bars are only shown for the three mutants with defects in sul20 binding.

As assayed by changes in fluorescence anisotropy, chimera^{+307}, chimera^{+211}, and wild-type Lon bound to a fluorescently labeled sul20 peptide with K_D's of ~2-5 μM, whereas ClpX$^{\Delta N}$ alone did not bind this peptide (Figure 4.1C). The binding of chimera^{+211} to the fluorescent peptide was inhibited by increasing concentrations of an unlabeled sul20 peptide but was not inhibited by an unrelated peptide (Figure 4.2D). Thus, the sul20 peptide binds specifically to a site contained within the first 211 residues of the N domain of *E. coli* Lon.

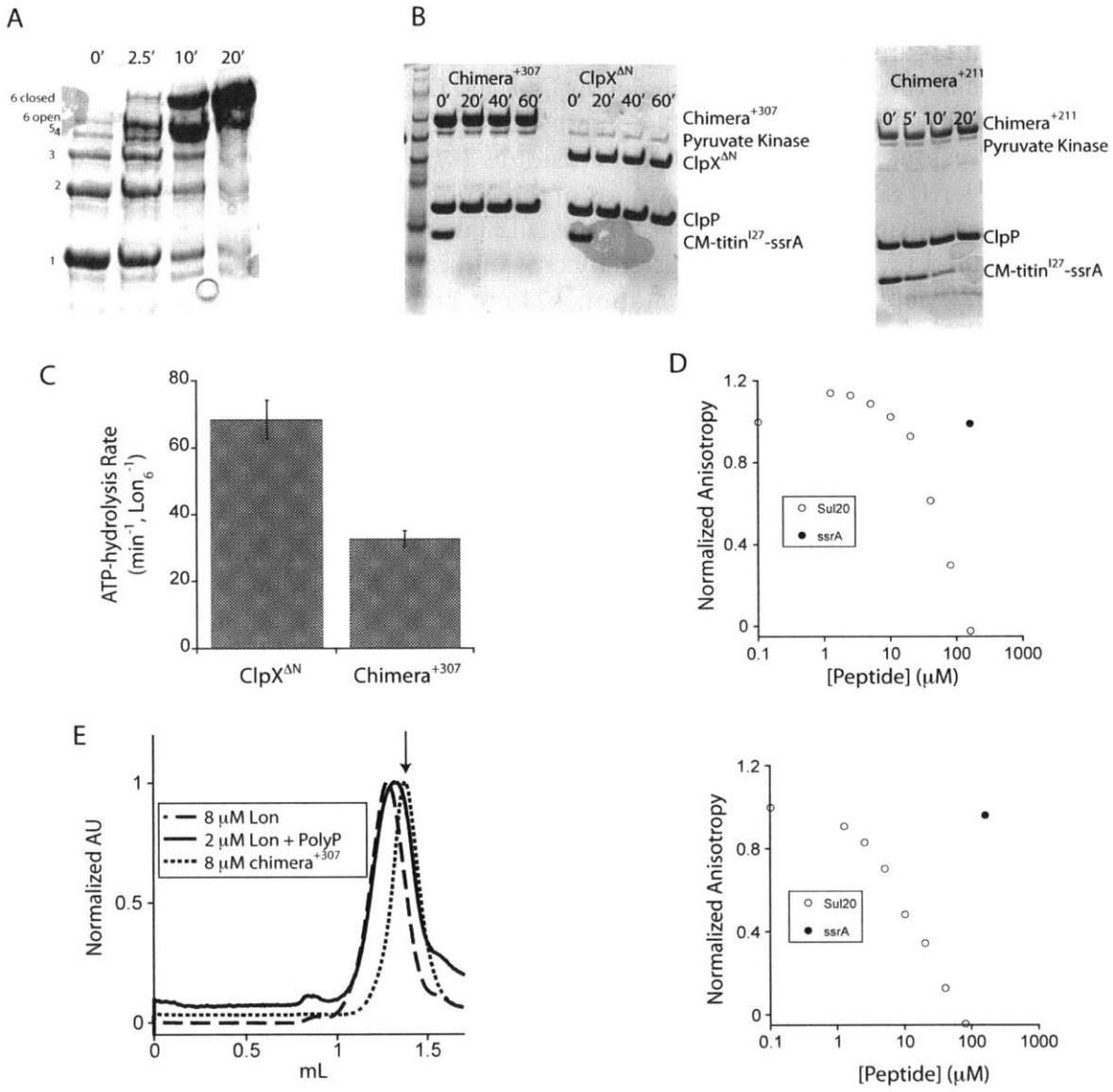

Figure 4.2 Characterization of chimera proteins

(A) Oxidation of disulfide bonds with copper phenanthroline leads to formation of covalent linear (open) or covalently closed hexamers of chimera^{+211}, as assayed by non-reducing SDS-PAGE. (B) Left panel: gel degradation assay of CM-titin^{I27}-ssrA (10 μM) by chimera^{+307} (1 μM) or ClpX$^{\Delta N}$ (1 μM) with ClpP (1.5 μM) at 30° C. Right panel: gel degradation assay of CM-titin^{I27}-ssrA (10 μM) by chimera^{+211} (0.5 μM) with ClpP (1 μM) at 30° C. Note that this is a reducing gel, so even though chimera^{+211} is oxidized during the assay, it runs as a monomer. Assays were performed in PD Buffer [25 mM HEPES (pH 7.5), 100 mM KCl, 5 mM $MgCl_2$, 10% glycerol at 30° C. Reactions were initiated by the addition of substrate, and the 0 min time point was taken immediately after addition of substrate. (C) Basal ATPase activity of ClpX$^{\Delta N}$ and Chimera^{+307} at 37° C. Values are averages ± SD (n = 3). (D) Unlabeled sul20 peptide but not ssrA peptide competes for binding of fluorescein-

labeled sul20 peptide (200 nM) to Lon^{S679A} (6 μM; top panel) or chimera^{+211} (6 μM; bottom panel) as assayed by fluorescence anisotropy (excitation 494 nm, emission 521 nm). Data were normalized so that anisotropy at 0 μM competitor peptide was equal to 1 and anisotropy with 0 μM hexamer (Lon or chimera^{+211}) and 160 μM competitor peptide was equal to 0. Assays were performed with 1 mM ATPγS at 37° C. (E) Chromatography of Lon^{S679A} (8 μM) or chimera^{+307} (8 μM) on a Superose-6 gel-filtration column. In the presence of 100 μM polyphosphate (Sigma, n = 45), Lon chromatographs as a hexamer. Chimera^{+307} chromatographed at a position smaller than expected for a dodecamer (calculated MW of dodecamer = 900 kDa), under conditions where Lon runs predominantly as a dodecamer, indicating that the N domain is not sufficient for stable dodecamer formation. Samples were dialyzed against FPLC Buffer [50 mM HEPES (pH 7.5), 150 mM NaCl, 10 μM EDTA, 10% glycerol] for 30 min and 1 mM $MgCl_2$ and 1 mM ATPγS were added immediately before loading onto the column (25 μL injection). For this experiment, chimera^{+307} contained disulfide stabilized hexamers, identical to those used for chimera^{+211}. Data were normalized so that maximum absorbance was equal to 1. Arrow indicates the position for the 670 kDa standard (BioRad).

Mapping the sul20 binding site

We attached a UV-activatable crosslinker that contained a biotin and cleavable disulfide to the a- and/or e-amino groups of a sul20 "bait" peptide. Following incubation of the bait peptide with Lon, we activated crosslinking by UV irradiation, reduced the disulfide to remove the sul20 portion of the crosslinked moiety, cleaved with trypsin, and enriched for biotinylated peptides. Mass spectrometry identified a single biotinylated Lon peptide (residues 100-113) (Figure 4.3), suggesting that the sul20 binding site was within 14 Å (the linker length between the "bait" peptide and the crosslinker) of this peptide in the structure of N domain. Next, we performed alanine-scanning mutagenesis of solvent-exposed residues within 14 Å of residues 100-113 in the crystal structure of a N-domain fragment (3LJC.pdb). We mutated blocks of two or three residues, purified the mutant proteins, and assayed for defects in sul20-peptide binding by

fluorescence anisotropy. Mutating residues 14-15, 33-35, or 36-38 to alanines resulted in significant loss of binding (Figure 4.1D). All of these residues were close in the 3LJC crystal structure, suggesting that all of the mutations affect the same binding site (Figure 4.1B). We focused further studies on the R33A/E34A/K35A mutant (henceforth called Lon$^{33-35}$), which had the largest and most reproducible defect in sul20 binding.

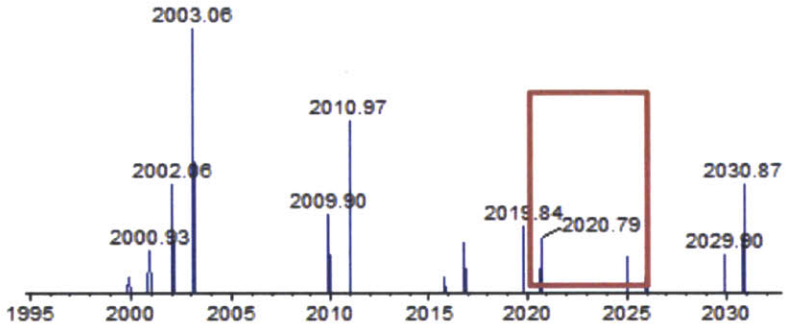

Figure 4.3 Mass spectrometry of biotin-labeled Lon peptide
Mass spectrometry identified one Lon peptide with a mass (2020.79 Da) very close to the mass expected for residues 100-113 plus the biotin label (2022.93 Da).

Degradation by Lon$^{33-35}$ in vitro

To test the importance of residues 33-35 for Lon degradation *in vitro*, we assayed proteolysis of variants of titin^{I27}-sul20 and titin^{I27}-β20 that had been unfolded by carboxymethylation (CM) of cysteines normally buried in the hydrophobic core (Kenniston et al. 2003). As assayed by SDS-PAGE, wild-type Lon degraded both substrates, whereas Lon$^{33-35}$ degraded CM-titin^{I27}-β20 at a rate similar to the wild-type enzyme but degraded CM-titin^{I27}-sul20 slowly (Figure 4.4). We found that Lon$^{33-35}$ also degraded β20-cp6-GFP, a stable native substrate (Wohlever et al. 2013), and FITC-casein, another unfolded substrate, with K_M and V_{max} values similar to those of wild-type Lon (Figure 4.4B & C). Thus, the

R33A/E34A/K35A mutations do not affect degradation of β20-tagged substrates or FITC casein but appear to cause a selective defect in degradation of a sul20-tagged substrate.

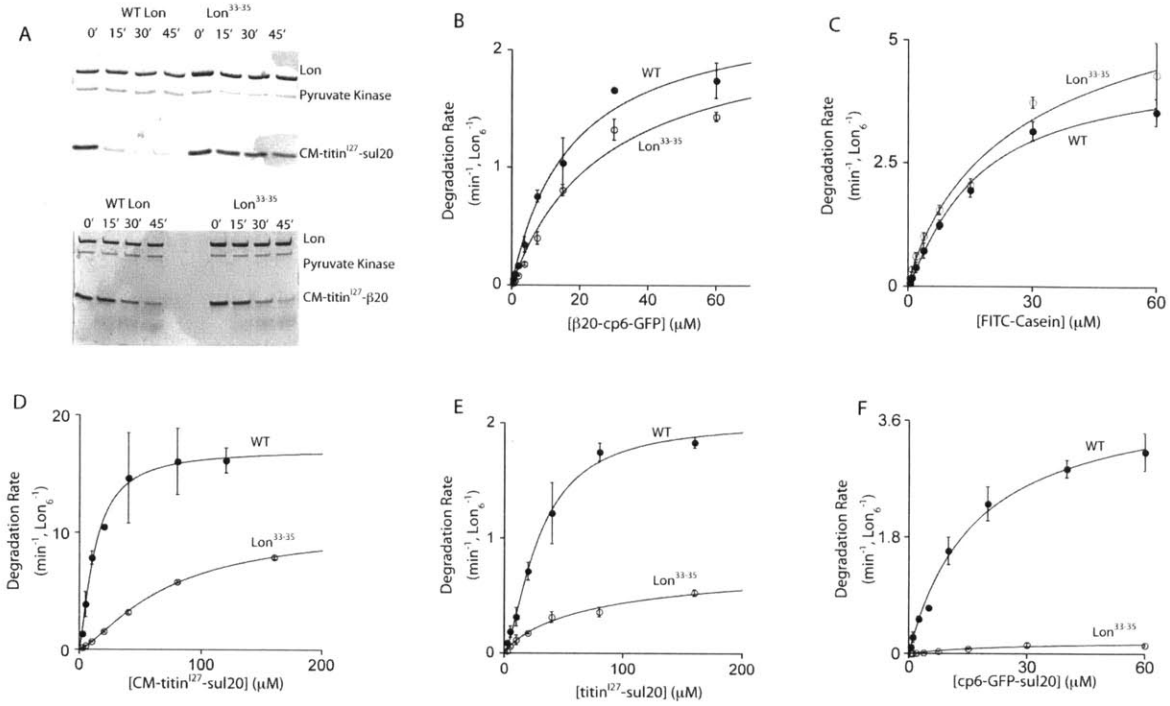

Figure 4.4 Degradation of model substrates by Lon and Lon$^{33-35}$
(A) SDS PAGE assay of the degradation of CM-titin^{I27}-sul20 (10 μM, top panel) or CM-titin^{I27}-β20 (10 μM, bottom panel) by wild-type Lon (0.3 μM) or Lon$^{33-35}$ (0.3 μM). (B) Substrate dependence of degradation of β20-cp6-GFP assayed by changes in fluorescence (excitation 467 nm; emission 511 nm). Values are means ± SEM (n ≥ 2). Wild-type Lon data was taken from (Wohlever et al. 2013). (C) Substrate dependence of degradation of FITC-Casein Type II (Sigma) assayed by changes in fluorescence (excitation 490 nm; emission 525 nm). Values are means ± SEM (n = 3). (D) Substrate dependence of degradation of ^{35}S-labeled CM-titin^{I27}-sul20 assayed by acid-soluble radioactivity. Values are means ± SEM (n = 3). (E) Substrate dependence of degradation of ^{35}S-labeled titin^{I27}-sul20 assayed by acid-soluble radioactivity. Values are means ± SEM (n = 3). (F) Substrate dependence of degradation of cp6-GFP-sul20 assayed by changes in fluorescence (excitation 467 nm; emission 511 nm). Values are means ± SEM (n ≥ 2). Wild-type Lon data was taken from (Wohlever et al. 2013). In panels B-F, solid lines are fits to the Hill form of the Michaelis-Menton equation $V = V_{Max} / (1+ (K_M /[S])^n)$. All assays were performed at 37° C using 0.3 μM Lon or Lon$^{33-35}$.

		Proteolysis			ATPase			
Substrate	Lon Variant	V_{max} (min^{-1}, Lon_6^{-1})	Apparent Km (μM)	Hill Constant	V_{max} (min^{-1}, Lon_6^{-1})	Apparent Km (μM)	Hill Constant	ATP per substrate
FITC-casein	WT	4.4 ± 0.4	16 ± 4	1.2 ± 0.1	nd	nd	nd	nd
	33-35	7 ± 2	31 ± 23	0.9 ± 2	nd	nd	nd	nd
cp6-GFP-sul20	WT*	3.6 ± 0.3	13 ± 2	1.2 ± 0.1	217 ± 16	9.3 ± 1.3	1.2 ± 0.1	60
	33-35	0.15 ± 0.01	13 ± 2	2 ± 0.4	84 ± 13	3 ± 0.8	1.5 ± 0.6	560
β20-cp6-GFP	WT*	2.3 ± 0.3	21 ± 6	1.2 ± 0.2	143 ± 5	2.6 ± 0.2	1.7 ± 0.1	62
	33-35	1.6 ± 0.1	14 ± 1	1.7 ± 0.2	nd	nd	nd	nd
CM-titin^{I27}-β20	WT	5.5 ± 0.1	18 ± 1	1.4 ± 0.1	174 ± 7	1.6 ± 0.2	0.9 ± 0.1	32
	33-35	nd	nd	nd	250 ± 50	5 ± 3	0.7 ± 0.2	nd
	Y398A	nd	nd	nd	25 ± 2	1.4 ± 0.4	1 ± 0.3	nd
CM-titin^{I27}-sul20	WT	17.1 ± 0.7	12 ± 1	1.4 ± 0.2	200 ± 20	0.6 ± 0.1	1.3 ± 0.4	12
	33-35	10.5 ± 0.4	70 ± 5	1.4 ± 0.1	270 ± 70	13 ± 10	0.7 ± 0.2	36
	Y398A	nd	nd	nd	23 ± 1	2.3 ± 0.5	3 ± 1.8	nd
titin^{I27}-sul20	WT	2 ± 0.1	29 ± 3	1.5 ± 0.2	118 ± 5	1 ± 0.1	1.1 ± 0.1	59
	33-35	0.8 ± 0.3	80 ± 70	0.9 ± 0.2	79 ± 8	3.5 ± 0.9	1.2 ± 0.3	99

Table 4.1 Steady-state kinetic parameters

The error is that of non-linear-least-squares fitting. ATP per substrate was calculated by dividing V_{max} for ATP hydrolysis by V_{max} for proteolysis. * Data taken from Wohlever et al. 2013. nd = not determined

To assess how the 33-35 mutations alter degradation of sul20-tagged proteins, we determined steady-state kinetics for Lon and Lon$^{33-35}$ degradation (Figure 4.4 D-F; Table 1) of unfolded CM-titin^{I27}-sul20 and two native substrates (titin^{I27}-sul20 and cp6-GFP-sul20). In each case, the 33-35 mutations resulted in increases in K_M (ranging from ~1.5 to ~5 fold), decreases in V_{max} (ranging from ~2 to ~20 fold), and decreases in the second-order degradation rate constant (V_{max}/K_M; ranging from ~6 to ~30 fold). For both the native and unfolded titin^{I27}-sul20 proteins, the 33-35 mutations decreased V_{max} by ~2 fold, suggesting a general defect in degradation, perhaps translocation or proteolytic cleavage, rather than in substrate unfolding. For degradation of cp6-GFP-sul20, however, the 33-35 mutations decreased V_{max} ~20 fold, indicating a strong defect in unfolding of this native substrate. Although the detailed effects of the 33-35 mutations on degradation of sul20-tagged proteins vary for different substrates, binding of the

sul20 degron to the wild-type N domain of Lon is clearly required for efficient proteolysis of these substrates.

An allosteric role for the sul20-binding site in the N domain

The effects of the 33-35 mutations on K_M can be rationalized if binding of the sul20 degron to the wild-type N domain allows direct handoff to the translocation machinery in the axial pore or passively increases the local substrate concentration near the pore. However, the effects of the 33-35 mutations on V_{max} require allosteric linkage between binding of sul20 degrons to the wild-type N domain and the enzymatic machinery required for translocation, unfolding, and/or degradation. To test for possible effects that depend on the rate of ATP hydrolysis, we assayed ATPase activity by Lon or Lon$^{33\text{-}35}$ as a function of the concentration of sul20-tagged substrates. Saturating concentrations of cp6-GFP-sul20 stimulated ATP hydrolysis to 3-fold lower levels for Lon$^{33\text{-}35}$ than for wild-type Lon (Figure 4.5A). This decrease almost certainly contributes to the poor degradation of cp6-GFP-sul20 by Lon$^{33\text{-}35}$ compared to Lon. By contrast, saturating concentrations of CM-titin^{I27}-sul20 resulted in similar rates of ATP hydrolysis for Lon$^{33\text{-}35}$ and wild-type Lon (Figure 4.5B). Thus, slow ATP hydrolysis is not responsible for the ~2-fold reduction in V_{max} for CM-titin^{I27}-sul20 degradation by Lon$^{33\text{-}35}$ compared to Lon.

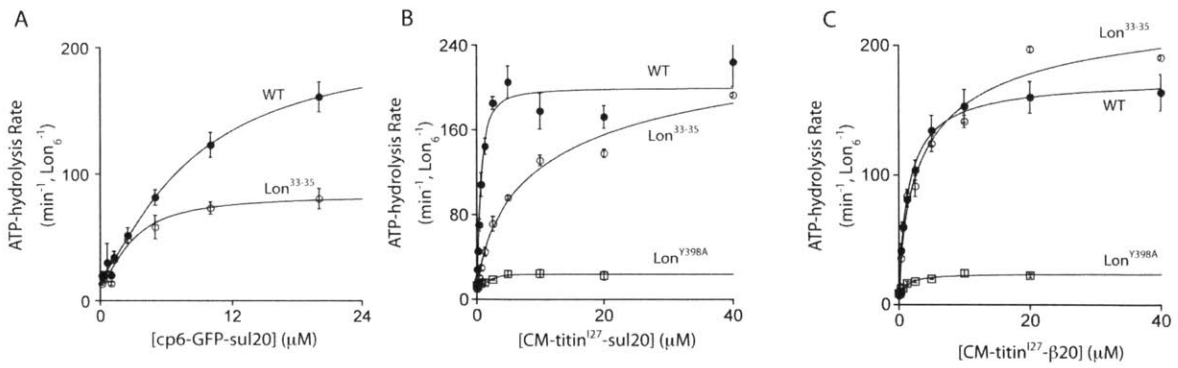

Figure 4.5 Substrate dependence of ATPase stimulation
Stimulation of ATP hydrolysis by Lon or variants (0.15 μM) by (A) cp6-GFP-sul20, (B) CM-titin^{I27}-sul20, and (C) CM-titin^{I27}-β20. Values are means ± SEM (n ≥ 3). Solid lines are fits to the Hill Equation rate = basal + amp/(1 + (K_M/[S])^n) where V_{Max} = basal + amp.

Similar concentrations of cp6-GFP-sul20 half-maximally stimulated ATP hydrolysis and degradation by wild-type Lon. By contrast, for Lon$^{33-35}$, ~10-fold higher concentrations of cp6-GFP-sul20 were required for half-maximal stimulation of degradation compared to ATP hydrolysis (Figure 4.6A). This offset, which suggests that many ATP-hydrolysis events lead to outcomes other than degradation, could also contribute to the very low V_{max} for Lon$^{33-35}$ degradation of cp6-GFP-sul20. CM-titin ^{I27}-sul20 half stimulated ATP hydrolysis at ~10-fold lower concentration than degradation both for wild-type Lon and Lon$^{33-35}$ (Figure 4.6B), but there was a marked increase in the number of ATP hydrolysis events required for substrate degradation for Lon$^{33-35}$ compared to wild-type Lon (Figure 4.6C), again suggesting a disruption in coordination of ATP hydrolysis and degradation. Thus, in addition to effects on substrate binding, the 33-35 mutations can affect V_{max} for substrate-stimulated ATP hydrolysis and/or the coordination between ATP hydrolysis and degradation, but in a manner that depends upon the detailed properties of the sul20-tagged substrate.

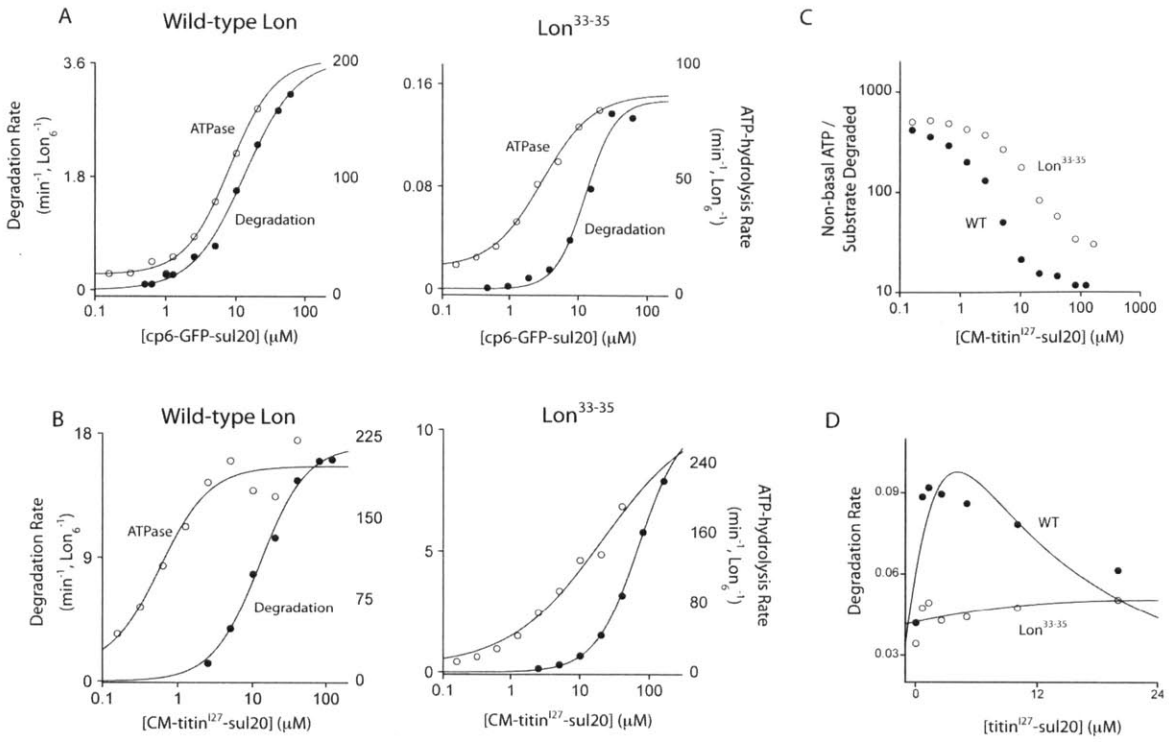

Figure 4.6 The 33-35 mutation alters coupling of ATPase and protease activity

For panels (A) and (B), ATPase and protease data were taken from Figures 4.4 and 4.5 and solid lines are fits as described in each of those figure legends. (A) For wild-type Lon (left panel), the $K_{1/2}$ values for degradation of and ATPase stimulation by cp6-GFP-sul20 are similar, whereas for Lon^{33-35} (right panel) the $K_{1/2}$ for ATPase stimulation is ~10-fold tighter than the $K_{1/2}$ for degradation. (B) For wild-type Lon (left Panel) and Lon^{33-35} (right Panel) the $K_{1/2}$ for ATPase stimulation by CM-$titin^{I27}$-sul20 is ~10-fold tighter than the $K_{1/2}$ for degradation of this substrate. (C) Lon^{33-35} requires more ATP hydrolysis events to degrade CM-$titin^{I27}$-sul20 than wild-type Lon. Y-axis plots non-basal ATPase rate (ATPase rate – basal rate) divided by the rate of substrate degradation. A higher value reflects a lower efficiency of substrate degradation, as measured by ATP molecules hydrolyzed per substrate degraded. (D) Degradation of F-β20-Q peptide (2 μM) by wild-type Lon (0.15 μM) or Lon^{33-35} (0.15 μM) was assayed by changes in fluorescence (excitation 320 nm; emission 422 nm) in the presence of increasing concentrations of $titin^{I27}$-sul20. Data are averages (n=5). Solid lines are fits to the equation rate = $C^*\alpha^*(1 + \alpha + \beta)/(L + (1 + \alpha + \beta)^2)$. Where C is a scaling factor, $\alpha = [F\text{-}\beta20\text{-}Q]/K_M$, $\beta = [titin^{I27}\text{-}sul20]/K_{0.5}$, and L is a conformational equilibrium constant (Segel 1993).

To test for an effect of the 33-35 mutations on activation of substrate degradation *in trans*, we monitored cleavage of a F-β20-Q peptide (cleavage

separates a fluorophore and quencher and increases fluorescence) by Lon^{33-35} and wild-type Lon in the presence of increasing concentrations of $titin^{I27}$-sul20 (Figure 4.6D). As expected from previous studies (Gur and Sauer 2009), low concentrations of $titin^{I27}$-sul20 activated cleavage of F-β20-Q by wild-type Lon, whereas higher concentrations resulted in decreased cleavage. By contrast, $titin^{I27}$-sul20 activated Lon^{33-35} cleavage of F-β20-Q to a much smaller extent, strongly supporting a role for sul20 binding to the N domain in allosteric activation.

Effects of an axial pore-loop mutation

Although the 33-35 mutations weaken or eliminate binding of the sul20 degron to the N domain, sul20-tagged substrates can still stimulate ATP hydrolysis and be degraded by Lon^{33-35} (Figure 4.4 & 4.5). Based on previous results (Wohlever et al. 2013), we suspected that these activities might be mediated by sul20 binding in the axial pore. To test this model, we mutated a highly conserved tyrosine in the axial pore to alanine (Y398A) (Figure 4.7A). In other AAA+ proteases, this tyrosine is required for translocation, and the Tyr→Ala mutation abrogates proteolytic activity and alters ATPase rates (Park et al. 2005; Martin et al. 2008). As expected, Lon^{Y398A} had no detectable proteolytic activity against an unfolded substrate (Figure 4.7B), confirming an important role for Tyr^{398} in substrate translocation and degradation by Lon. When CM-$titin^{I27}$-sul20 or CM-$titin^{I27}$-β20 were titrated against Lon^{Y398A}, maximal stimulation of ATPase activity was reduced more than 10-fold compared to wild-type Lon (Figure 4.5B & C). This decrease in maximal ATPase stimulation was not caused by altered binding of the sul20 degron to the N domain, as a sul20

peptide bound Lon^{Y398A} with the same affinity as wild-type Lon (Figure 4.7C). These results suggest that the loops that line the axial pore of wild-type Lon play important roles in substrate stimulation of ATP hydrolysis regardless of degron identity.

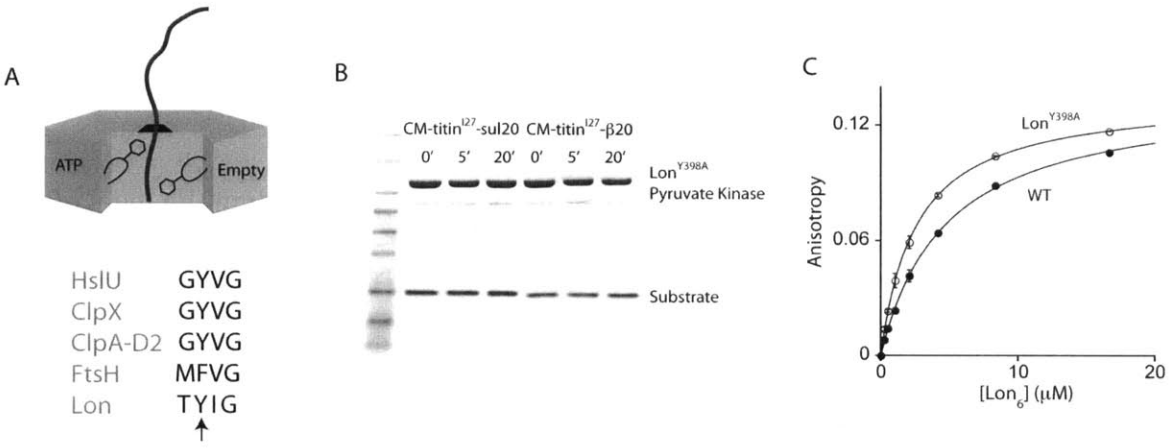

Figure 4.7 The role of the axial-pore loop in Lon activity
(A) Cartoon showing sequence conservation of axial-pore loops in AAA+ proteases from *E. coli*. The arrow points to Y398 in Lon. (B) SDS-PAGE assay shows that Lon^{Y398A} (0.6 μM) did not degrade the unfolded substrates CM-titin^{I27}-sul20 (10 μM) or CM-titin^{I27}-β20 (10 μM). (C) Binding of fluorescein-labeled sul20 peptide (200 nM) to Lon^{S679A} or Lon^{Y398A} assayed by fluorescence anisotropy (excitation 494 nm, emission 521 nm). Data were baseline corrected to have no anisotropy at 0 μM Lon. Values are averages ± SEM(n≥2).

Cellular phenotypes

To characterize activity *in vivo*, we cloned different Lon variants into low-copy plasmids and assayed different phenotypes in *E. coli* strains lacking the chromosomal *lon* gene. Inactivation of SulA by Lon is required for resumption of robust growth following repair of UV-induced DNA damage (Gottesman et al. 1981). However, Van Melderen and Gottesman (1999) found that overexpression of proteolytically inactive Lon^{S679A} rescues growth, indicating that SulA degradation is

not necessary for its inactivation. In our assays, strains expressing wild-type Lon or Lon^{E240K} grew similarly after UV irradiation, whereas cells with an empty vector or cells expressing $Lon^{33\text{-}35}$ grew very poorly (Figure 4.8A). Thus, the 33-35 mutations in the N domain of Lon are sufficient to prevent SulA inactivation *in vivo*. Lon variants bearing a single mutation in the proteolytic active site (S679A) or with this mutation and a mutation in the axial-pore tyrosine (Y389A) grew well following UV irradiation (Figure 4.8A). Western blots revealed that the different Lon variants were expressed at similar levels, which were only slightly higher than the level of Lon expressed from its normal chromosomal locus (Figure 4.8B). Taken together, these results suggest that Lon can inhibit SulA by simple binding, with no need for ATP-dependent remodeling or translocation into the proteolytic chamber.

We also tested the ability of different Lon variants to support growth of cells subjected to proteotoxic stress by high temperature, the absence of the ClpXP protease, and low levels of the DnaJ and DnaK chaperones (Tomoyasu et al. 2001). In this background at 42° C, $Lon^{33\text{-}35}$ and Lon^{S679A} supported growth less well than wild-type Lon but better than cells with the empty vector (Figure 4.8C). The partial rescue by proteolytically inactive Lon^{S679A} indicates that activities other than degradation contribute to control of proteotoxic stress. The $Lon^{Y398A/S679A}$ double mutant showed no rescue, suggesting that translocation of substrates through the axial pore is required to suppress proteotoxic stress. When the 33-35 mutations in the N domain were combined with the S679A active-site mutation, the resulting variant supported growth as well as wild-type Lon (Figure 4.8C). This surprising result suggests that the 33-35 and S679A mutations act in a synergistic fashion,

possibly by independently stabilizing a Lon conformation that reduces proteotoxic stress in a translocation-dependent remodeling reaction.

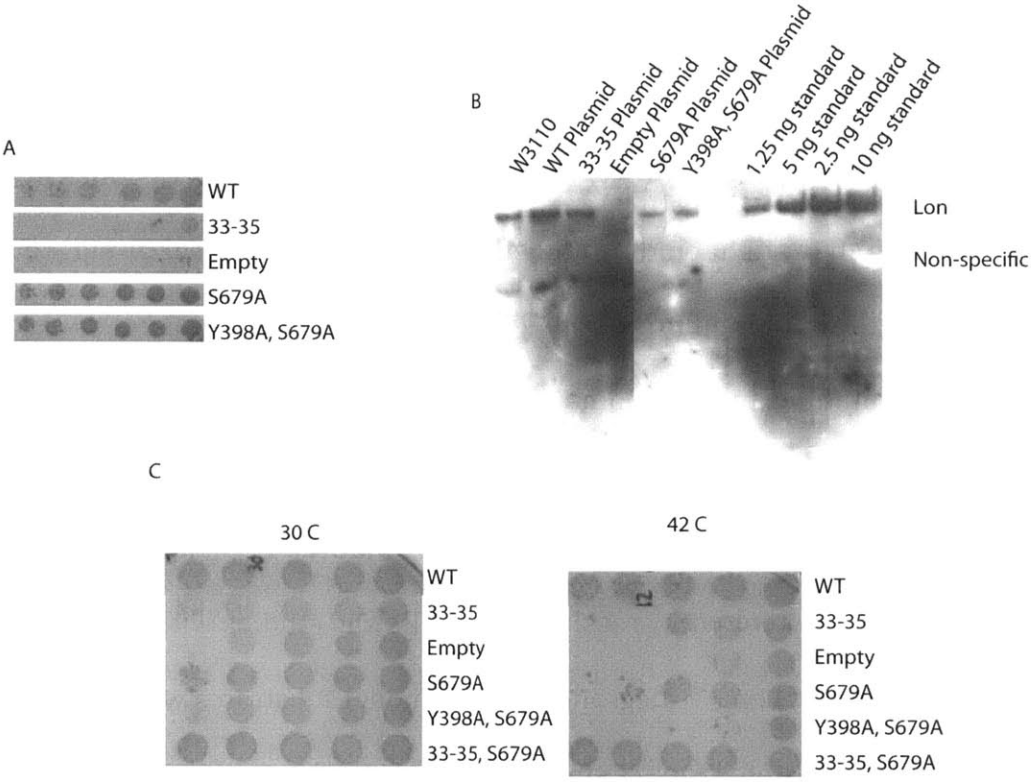

Figure 4.8 Activity of Lon variants *in vivo*

(A) Inactivation of SulA *in vivo*. Following UV-induced activation of the SOS-response, W3110 *lon*::Kan^{R} cells with pBAD33 plasmids encoding wild-type Lon, Lon^{S679A}, or Lon$^{Y398A/S679A}$ were able to rescue growth, whereas a plasmid encoding Lon$^{33-35}$ or the empty pBAD33 vector could not. (B) Western blot assay of intracellular levels of chromosomal Lon and Lon variants expressed from pBAD33. (C) Rescue of cells from proteotoxic stress caused by deletion of *clpXP-lon*, reduced expression of DnaK/J, and growth at 42° C. Plasmids expressing Lon$^{33-35}$ and Lon^{S679A} provided partial rescue of growth under conditions of proteotoxic stress (right panel), whereas plasmids expressing wild-type Lon and Lon$^{33-35/S679A}$ provided full rescue of growth. The empty pBAD33 plasmid and the plasmid expressing the Lon$^{Y398A/S679A}$ double mutant had the lowest level of rescue from proteotoxic stress.

Discussion

The biochemical and mutational experiments described here provide evidence that the N domain of *E. coli* Lon participates in substrate binding as well as in regulating the rate of substrate proteolysis. A set of spatially adjacent residues in the N domain is required for tight binding and efficient degradation of sul20-tagged substrates, including the SulA inhibitor of cell division. For example, the R33A/E34A/K35A mutations in the N domain weaken binding to a sul20 peptide, increase K_M and reduce V_{max} for degradation of sul20-tagged substrates, and prevent Lon relief of SulA inhibition of cell division following UV-induced DNA damage. The location of this site is consistent with crosslinking results and studies with a Lon/ClpX chimera. Notably, Lon^{33-35} degrades several model substrates with non-sul20 degrons with steady-state kinetics similar to those of wild-type Lon. Thus, the site defined by the R33A/E34A/K35A mutations is only required for the binding and/or degradation of a subset of Lon substrates, including sul20-tagged substrates.

The R33A/E34A/K35A mutations also prevent efficient *trans* activation of cleavage of a β20 peptide by a sul20-tagged protein, and alter stimulation of ATP hydrolysis by sul20-tagged substrates. These results support a model in which binding of the sul20 degron to the site defined by the R33A/E34A/K35A mutations causes allosteric changes in Lon conformation that stimulate ATP hydrolysis and proteolysis (Figure 4.9). Although SulA is restricted to γ proteobacteria, Lon residues 33-35 are highly conserved across α, β, and γ proteobacteria as well as in

Gram-positive bacteria (Figure 4.10), suggesting that this site also serves to bind substrates with sul20-related degrons.

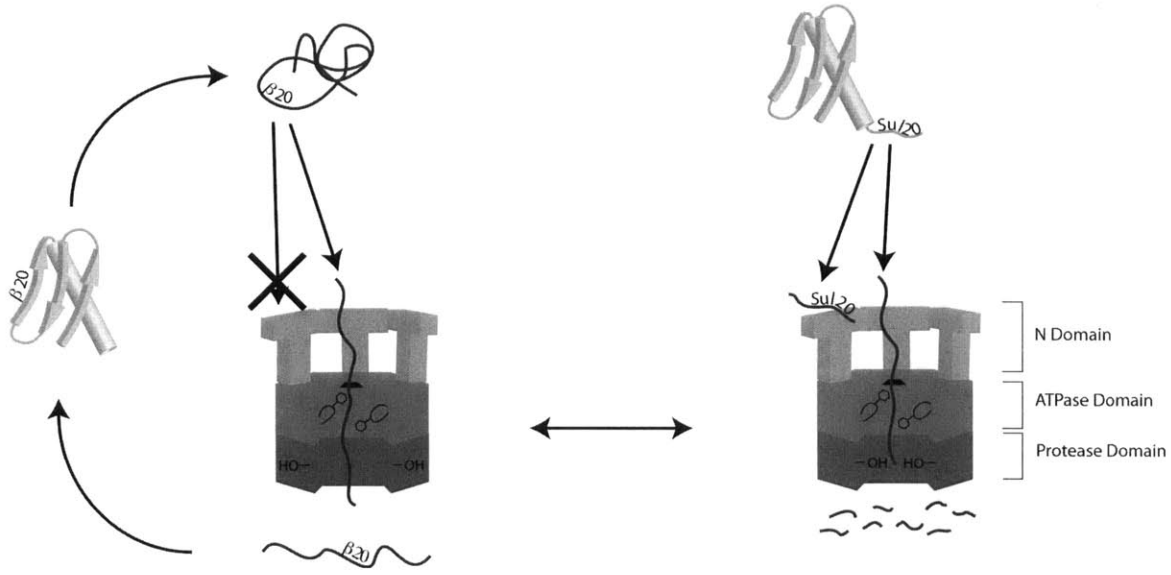

Figure 4.9 Model for degron-mediated regulation of Lon chaperone activity

Lon equilibrates between a chaperone state (left) and a protease state (right). Binding of degrons, such as sul20, to the N-domain binding site defined by the 33-35 mutations drives the equilibrium to the right, resulting in higher levels of degradation. Misfolded substrates that do not interact with this site, such as β20-tagged proteins, do not drive the equilibrium toward the right, resulting in lower levels of degradation and higher rates of unfolding and release, potentially allowing these substrates to refold properly.

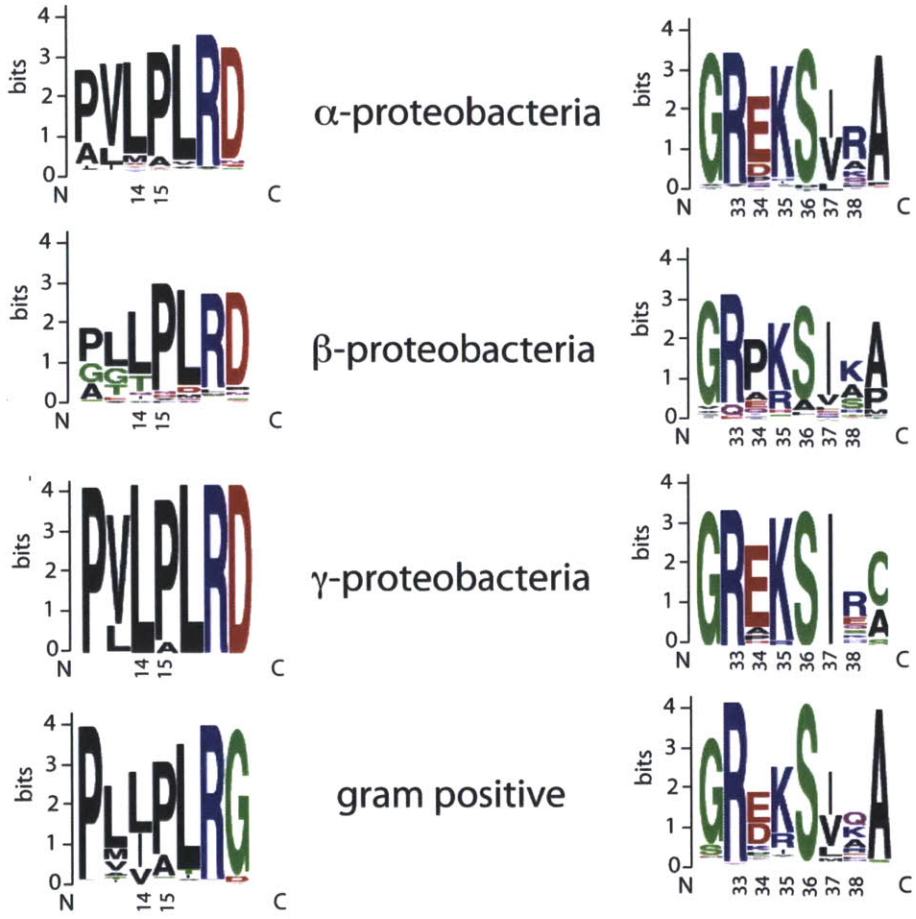

Figure 4.10 Conservation of residues forming the N-domain binding site for sul20

Conservation of Lon residues 14-15 and 33-38. *E. coli* Lon was blasted (NCBI BlastP) against each bacterial subdivision and the top 500 non-redundant sequences were aligned using ClustalW. Results of the sequence alignment were loaded into Weblogo (weblogo.berkeley.edu). Numbering refers to *E. coli* Lon.

It is likely that substrates interact with multiple sites on Lon. Indeed, the sul20 degron of some cp6-GFP-sul20 substrates are proteolytically clipped by Lon without global degradation, implying that the sul20 tag is the first part of these molecules to pass through the axial pore and enter the degradation chamber (Wohlever et al. 2013). However, the sul20 degron also binds to the N domain of Lon. The sul20 degron may initially bind to the N domain and subsequently be

transferred to the axial pore in a hand-off reaction. Alternatively, the sul20 degrons of some substrate molecules may bind to the N domain and allosterically activate proteolysis of other substrates whose degrons are engaged by the pore. Indeed, the latter model is supported by transactivation experiments reported here and previously (Gur and Sauer 2009). We also find that the Y398A mutation, which truncates a highly conserved aromatic side chain in the axial-pore loop, prevents degradation of unfolded substrates bearing the sul20 or β20 degrons and dramatically reduces the maximal level of stimulation of ATP hydrolysis by these substrates. Nevertheless, Lon^{Y398A} binds a sul20 peptide with wild-type affinity, indicating that this mutation does not impair binding of this degron to the N domain. These results are consistent with independent binding of sul20 degrons on different substrate molecules to the N domain and to the axial pore. Studies with other AAA+ proteases show that mutations corresponding to Y398A prevent or greatly slow the rate of substrate translocation (Martin et al. 2008). Thus, both allosteric activation via degron binding to the N domain and interaction of a translocating segment of polypeptide with the axial-pore loops appear to be required for normal coordination of substrate binding, ATP hydrolysis, and translocation by Lon.

SulA can be inactivated by Lon without degradation. Indeed, Van Melderen and Gottesman (1999) found that overproduction of proteolytically inactive Lon^{S679A} allowed resumption of cell growth after UV irradiation, whereas overproduction of a Lon variant with a mutation in the Walker-A motif did not. Because the Walker-A motif is required for ATP binding and hydrolysis, they suggested that Lon inactivates SulA by unfolding and translocating it into the chamber. Our results

confirm that Lon^{S679A} inactivates SulA *in vivo*. However, we find that Lon^{Y398A}, which is defective in translocation and unfolding, also inactivates SulA *in vivo*. Thus, simple binding of the sul20 degron of SulA to the N domain of Lon appears to be sufficient to prevent inhibition of cell division. SulA binding to Lon may be ATP dependent, explaining why the Walker-A mutant fails to rescue cell growth.

There are hints in the literature that Lon may have chaperone-like activity (Rep et al. 1996; Lee et al. 2004a; Coleman et al. 2009). For some client proteins, binding to the N domain of Lon may suffice to prevent aggregation or help direct folding. However, based on the finding that certain degrons stabilize a Lon state with high ATPase activity but little or no protease activity, Gur and Sauer (2009) proposed that Lon could exist in a conformation in which misfolded substrates were unfolded and released, giving them a chance to fold properly. In support of this model, we find that the protease-defective Lon^{S679A} variant partially suppresses proteotoxic stress *in vivo*, whereas the protease-defective and translocation-defective Lon$^{Y398A/S679A}$ variant does not. Surprisingly, a variant containing both the 33-35 and S679A mutations is as active as wild-type Lon in suppressing proteotoxic stress, suggesting that it may have enhanced chaperone function and that binding of client proteins to the N-domain site altered by the 33-35 mutations is not required for this activity.

The family-specific N domains of the ClpX and ClpA AAA+ unfoldases serve as binding platforms for some substrates and adaptor proteins but can be deleted without compromising hexamer formation or robust degradation of certain substrates by ClpXP and ClpAP (Sauer and Baker 2011). By contrast, although the N

domain of Lon also participates in substrate binding, it is far more highly integrated into overall enzyme architecture and function, including hexamer formation and allosteric control of ATP hydrolysis and protease activity. Understanding in structural terms how the N domain accomplishes these tasks is an important future challenge and goal.

Acknowledgements

We thank D. Barthelme, V. Baytshtok, S. Kim, E. Vieux, J. Chen, and members of the Sauer and Baker labs for reagents and helpful discussions. This work was supported by National Institutes of Health grant AI-16982, and an National science Foundation Graduate Research Fellowship to M.L.W.

References

Adam C, Picard M, Déquard-Chablat M, Sellem CH, Hermann-Le Denmat S, Contamine V. 2012. Biological roles of the Podospora anserina mitochondrial Lon protease and the importance of its N-domain. *PLoS ONE* **7**: e38138.

Baker TA, Sauer RT. 2006. ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends Biochem Sci* **31**: 647–653.

Bakshi S, Siryaporn A, Goulian M, Weisshaar JC. 2012. Superresolution imaging of ribosomes and RNA polymerase in live Escherichia coli cells. *Mol Microbiol* **85**: 21–38.

Bernstein SH, Venkatesh S, Li M, Lee J, Lu B, Hilchey SP, Morse KM, Metcalfe HM, Skalska J, Andreeff M, et al. 2012. The mitochondrial ATP-dependent Lon protease: a novel target in lymphoma death mediated by the synthetic triterpenoid CDDO and its derivatives. *Blood*.

Breidenstein EBM, Bains M, Hancock REW. 2012a. Involvement of the Lon protease in the SOS response triggered by ciprofloxacin in Pseudomonas aeruginosa PAO1. *Antimicrobial Agents and Chemotherapy*.

Breidenstein EBM, Janot L, Strehmel J, Fernandez L, Taylor PK, Kukavica-Ibrulj I,

Gellatly SL, Levesque RC, Overhage J, Hancock REW. 2012b. The Lon Protease Is Essential for Full Virulence in Pseudomonas aeruginosa. *PLoS ONE* **7**: e49123.

Cha S-S, An YJ, Lee C-R, Lee HS, Kim Y-G, Kim SJ, Kwon KK, De Donatis GM, Lee J-H, Maurizi MR, et al. 2010. Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *EMBO J* **29**: 3520–3530.

Cheng I, Mikita N, Fishovitz J, Frase H, Wintrode P, Lee I. 2012. Identification of a Region in the N-Terminus of Escherichia coli Lon That Affects ATPase, Substrate Translocation and Proteolytic Activity. *J Mol Biol.*

Chung CH, Goldberg AL. 1981. The product of the lon (capR) gene in Escherichia coli is the ATP-dependent protease, protease La. *Proc Natl Acad Sci USA* **78**: 4931–4935.

Coleman JL, Katona LI, Kuhlow C, Toledo A, Okan NA, Tokarz R, Benach JL. 2009. Evidence that two ATP-dependent (Lon) proteases in Borrelia burgdorferi serve different functions. *PLoS Pathog* **5**: e1000676.

Duman RE, Löwe J. 2010. Crystal structures of Bacillus subtilis Lon protease. *J Mol Biol* **401**: 653–670.

Ebel W, Skinner MM, Dierksen KP, Scott JM, Trempy JE. 1999. A conserved domain in Escherichia coli Lon protease is involved in substrate discriminator activity. *J Bacteriol* **181**: 2236–2243.

Glynn SE, Nager AR, Baker TA, Sauer RT. 2012. Dynamic and static components power unfolding in topologically closed rings of a AAA+ proteolytic machine. *Nat Struct Mol Biol.*

Gora KG, Cantin A, Wohlever M, Joshi KK, Perchuk BS, Chien P, Laub MT. 2013. Regulated proteolysis of a transcription factor complex is critical to cell cycle progression in Caulobacter crescentus. *Mol Microbiol.*

Gottesman S, Halpern E, Trisler P. 1981. Role of sulA and sulB in filamentation by lon mutants of Escherichia coli K-12. *J Bacteriol* **148**: 265–273.

Gottesman S, Roche E, Zhou Y, Sauer RT. 1998. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev* **12**: 1338–1347.

Gur E, Sauer RT. 2009. Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine. *Proceedings of the National Academy of Sciences* **106**: 18503–18508.

Gur E, Sauer RT. 2008. Recognition of misfolded proteins by Lon, a AAA(+) protease.

Genes Dev **22**: 2267–2277.

Higashitani A, Ishii Y, Kato Y, Koriuchi K. 1997. Functional dissection of a cell-division inhibitor, SulA, of Escherichia coli and its negative regulation by Lon. *Mol Gen Genet* **254**: 351–357.

Ingmer H, Brøndsted L. 2009. Proteases in bacterial pathogenesis. *Res Microbiol* **160**: 704–710.

Ishii Y, Amano F. 2001. Regulation of SulA cleavage by Lon protease by the C-terminal amino acid of SulA, histidine. *Biochem J* **358**: 473–480.

Kenniston JA, Baker TA, Fernandez JM, Sauer RT. 2003. Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* **114**: 511–520.

Kitagawa M, Ara T, Arifuzzaman M, Ioka-Nakamichi T, Inamoto E, Toyonaga H, Mori H. 2005. Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. *DNA Res* **12**: 291–299.

Lee AY-L, Chen M-Y, Wu S-H. 2004a. Identification of a gene encoding Lon protease from Brevibacillus thermoruber WR-249 and biochemical characterization of its thermostable recombinant enzyme. *Eur J Biochem* **271**: 834–844.

Lee AY-L, Hsu C-H, Wu S-H. 2004b. Functional domains of Brevibacillus thermoruber lon protease for oligomerization and DNA binding: role of N-terminal and sensor and substrate discrimination domains. *J Biol Chem* **279**: 34903–34912.

Li M, Gustchina A, Rasulova FS, Melnikov EE, Maurizi MR, Rotanova TV, Dauter Z, Wlodawer A. 2010. Structure of the N-terminal fragment of Escherichia coli Lon protease. *Acta Crystallogr D Biol Crystallogr* **66**: 865–873.

Li M, Rasulova F, Melnikov EE, Rotanova TV, Gustchina A, Maurizi MR, Wlodawer A. 2005. Crystal structure of the N-terminal domain of E. coli Lon protease. *Protein Sci* **14**: 2895–2900.

Luce K, Osiewacz HD. 2009. Increasing organismal healthspan by enhancing mitochondrial protein quality control. *Nat Cell Biol* **11**: 852–858.

Luo J, Solimini NL, Elledge SJ. 2009. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell* **136**: 823–837.

Martin A, Baker TA, Sauer RT. 2008. Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nat Struct Mol Biol* **15**: 1147–1151.

Melnikov EE, Andrianova AG, Morozkin AD, Stepnov AA, Makhovskaya OV, Botos I, Gustchina A, Wlodawer A, Rotanova TV. 2008. Limited proteolysis of E. coli ATP-dependent protease Lon - a unified view of the subunit architecture and characterization of isolated enzyme fragments. *Acta Biochim Pol* **55**: 281–296.

Moran U, Phillips R, Milo R. 2010. SnapShot: key numbers in biology. *Cell* **141**: 1262–1262.e1.

Park E, Rho YM, Koh O-J, Ahn SW, Seong IS, Song J-J, Bang O, Seol JH, Wang J, Eom SH, 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.

Rep M, van Dijl JM, Suda K, Schatz G, Grivell LA, Suzuki CK. 1996. Promotion of mitochondrial membrane complex assembly by a proteolytically inactive yeast Lon. *Science* **274**: 103–106.

Robertson GT, Kovach ME, Allen CA, Ficht TA, Roop RM. 2000. The Brucella abortus Lon functions as a generalized stress response protease and is required for wild-type virulence in BALB/c mice. *Mol Microbiol* **35**: 577–588.

Roudiak SG, Shrader TE. 1998. Functional role of the N-terminal region of the Lon protease from Mycobacterium smegmatis. *Biochemistry* **37**: 11255–11263.

Rudyak SG, Shrader TE. 2000. Polypeptide stimulators of the Ms-Lon protease. *Protein Sci* **9**: 1810–1817.

Sauer RT, Baker TA. 2011. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **80**: 587–612.

Segel IH. 1993. *Enzyme kinetics*. Wiley-Interscience.

Tomoyasu T, Mogk A, Langen H, Goloubinoff P, Bukau B. 2001. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the Escherichia coli cytosol. *Mol Microbiol* **40**: 397–413.

Van Melderen L, Aertsen A. 2009. Regulation and quality control by Lon-dependent proteolysis. *Res Microbiol*.

Wohlever ML, Nager AR, Baker TA, Sauer RT. 2013. Engineering fluorescent protein substrates for the AAA+ Lon protease. *Protein Eng Des Sel*.

Wright R, Stephens C, Zweiger G, Shapiro L, Alley MR. 1996. Caulobacter Lon protease has a critical role in cell-cycle control of DNA methylation. *Genes Dev* **10**: 1532–1542.

Chapter 5

The E240K mutation in the Lon N domain stabilizes dodecamers and alters degradation of model substrates

Abstract

E. coli Lon, a AAA+ protease, recognizes and degrades many different substrates, including regulatory proteins such as RcsA and SulA. In Chapter 4, a binding site was identified in the Lon N domain that interacts with a subset of substrates, including SulA, but how Lon recognizes the rest of its substrate repertoire is unknown. The E240K mutation in Lon was isolated more than a decade ago and was proposed to disrupt the binding site for RcsA. However, characterization of the E240K mutant *in vitro* suggests that the effects of this mutation are more complex than simple disruption of a binding site. For example, Lon^{E240K} exists almost exclusively as a dodecamer, whereas wild-type Lon equilibrates between hexamers and dodecamers. Lon^{E240K} appears to possess chaperone function similar wild-type Lon *in vivo* but displays degradation defects *in vitro* that do not correlate simply with substrate stability, degron identity, or dodecamer formation. Because residue 240 undergoes nucleotide-dependent conformational changes, this region may be important for coupling substrate binding in the N domain with allosteric activation of Lon protease and ATPase activity.

Introduction

AAA+ proteases play important biological roles in bacteria, eukaryotes, and archaea (Hanson and Whiteheart 2005; Baker and Sauer 2006; Sauer and Baker 2011). Inhibition of the AAA+ Lon protease reduces virulence in several pathogenic bacteria and is highly toxic to lymphoma cells, whereas overexpression of Lon improves lifespan of some fungi but kills *Escherichia coli* (Goldberg et al. 1994; Robertson et al. 2000; Ingmer and Brøndsted 2009; Luce and Osiewacz 2009; Breidenstein et al. 2012; Bernstein et al. 2012). Like most AAA+ proteases, Lon recognizes substrates by binding to specific amino-acid sequences called degrons or degradation tags (Baker and Sauer 2006; Sauer and Baker 2011). In *E. coli*, Lon degrades many native regulatory proteins, including the RcsA transcription factor and the SulA inhibitor of cell division, and also degrades the majority of misfolded proteins, including β-galactosidase (Gottesman and Zipser 1978; Chung and Goldberg 1981; Torres-Cabassa and Gottesman 1987). The RcsA degron is unknown, whereas the degrons for SulA (called sul20) and β-galactosidase (called β20) have been identified (Higashitani et al. 1997; Gur and Sauer 2008). The binding of the sul20 degron appears to stabilize a Lon conformation with high protease activity, whereas the binding of the β20 degron appears to stabilize a conformation that may function as a chaperone, possibly by unfolding misfolded substrates and allowing them to refold properly (Gur and Sauer 2009).

E. coli Lon is active as a homohexamer. Each subunit of 784 amino acids contains an N domain (~300 residues), a AAA+ module (~275 residues) consisting of large and small domains, and a peptidase domain (~200 residues) (Rotanova et al.

2006; Botos et al. 2004). The active sites for peptide-bond cleavage are sequestered within a chamber formed by the peptidase domains (Cha et al. 2010). Access to this chamber is regulated by the hexameric AAA+ ring of Lon, which couples ATP hydrolysis to conformational changes that unfold and translocate substrates through a narrow axial pore and into the chamber. As shown in Chapter 4, the N domain binds the sul20 degron and also coordinates the catalytic activities of the AAA+ domains and peptidase domains. The N domain is also required for stable hexamer formation (Lee et al. 2004; Melnikov et al. 2008). As discussed in Chapter 3, wild-type hexamers and dodecamers are in equilibrium at physiological Lon concentrations (~2-4 μM subunit equivalents), with the dodecamer degrading certain substrates (e.g., sul20-tagged CM-titin^{I27} or titin^{I27}) as well as the hexamer but degrading other substrates (e.g., FITC-casein) at substantially reduced rates. Although crystal structures are known for most parts of Lon, structures of the full-length hexamer or dodecamer have not been solved and it is not known how the N domain stabilizes hexamers or dodecamers or regulates the activities of the Lon catalytic domains (Li et al. 2005; 2010; Duman and Löwe 2010).

The E240K mutation in the N domain was isolated in a genetic screen for Lon variants that could degrade SulA but not RcsA (Ebel et al. 1999), suggesting that the region around residue 240, which forms a coiled-coil, may serve as an RcsA binding site. However, experiments using hydrogen-deuterium exchange and limited proteolysis also show that the region flanking residue 240 undergoes nucleotide-dependent changes in conformation (Vasilyeva et al. 2002; Cheng et al. 2012), and thus the effects of the E240K mutant could be indirect. Here, we characterize the

biochemical properties of the Lon^{E240K} mutant. Notably, this mutation stabilizes the Lon dodecamer, results in severe defects in the proteolysis of a subset of model sul20-tagged and β20-tagged substrates, but suppresses proteotoxic stress *in vivo* as well as wild-type Lon.

Materials and Methods

Variants of *E. coli* Lon were cloned into pBAD33. For proteotoxic-stress assays, the chloramphenicol resistance marker of pBAD33 was replaced with an ampicillin resistance marker cloned from pSH21. $Titin^{I27}$ variants were cloned into a pSH21 vector with an N-terminal His_6 tag. β20-cp6-GFP and cp6-GFP-sul20 were cloned into a pCOLADuet1 vector with an N-terminal His_6 tag followed by a PreScission protease site. Mutations were generated either by QuickChange PCR (Stratagene) or by standard PCR techniques.

Lon variants, cp6-GFP-sul20, $titin^{I27}$-sul20 20-cp6-GFP, and $titin^{I27}$ variants were expressed, purified, and carboxymethylated (if applicable) as described (Gur and Sauer 2009; Wohlever et al. 2013) (Chapter 4). Assays for degradation, ATP hydrolysis, and binding of the sul20 peptide were performed as described in Chapter 4.

Prior to ultracentrifugation, Lon^{E240K} was dialyzed overnight against 50 mM HEPES (pH 7.5), 150 mM NaCl, 10 μM EDTA, and 100 μM TCEP. Immediately before loading samples into dual-sector charcoal-filled epon centerpieces, 1 mM $MgCl_2$ and 100 μM ATPγS were added. Sedimentation-velocity analysis was performed at 16,000 rpm and 20 °C in a Beckman OptimaXL-1 analytical ultracentrifuge

(Biophysical Instrumentation Facility, MIT) using an An60-Ti rotor. SEDFIT (Brown and Schuck 2006) was used to calculate the continuous distribution of sedimentation coefficients from 0 to 60 *S* at a resolution of 200 scans per concentration with a confidence level of 0.95. Calculations were performed using a Lon partial specific volume of 0.7431 mL/g (SEDNTERP; J. Philo; http://www.jphilo.mailway.com), a density of 1.00831 g/mL, and a viscosity of 0.010475 poise (Chapter 3).

For the mucoidy assay, W3110 *lon*::*Kan*^{R} cells were transformed with pBAD33 vectors containing Lon variants. Liquid cultures were grown in LB broth until early-log phase, diluted 1000x in LB broth, and then spread onto minimal media plates containing 0.4% glycerol (Davis and Mingioli 1950), 25 μg/mL kanamycin, and 10 μg/mL chloramphenicol. Cells were grown at 30 °C for 48 h. Western blots and the proteotoxic stress assay were performed as described in Chapter 4.

Results

Lon^{E240K} forms a stable dodecamer

We expressed and purified Lon^{E240K} and performed sedimentation-velocity ultracentrifugation to characterize its oligomeric state. Strikingly, the Lon^{E240K} protein sedimented almost exclusively at ~21*S*, the dodecamer value, at concentrations in hexamer equivalents ranging from 0.5 to 3 μM (Figure 5.1A). The wild-type dodecamer degrades FITC-casein poorly compared to the hexamer, and thus the enzyme-normalized rate of degradation decreases at higher Lon

concentrations where the fraction of dodecamer increases (Chapter 3). By contrast, for Lon^{E240K}, the enzyme-normalized degradation rate of FITC-casein remained relatively constant over a 20-fold range of Lon^{E240K} concentration (Figure 5.1B). Under conditions where wild-type Lon is largely hexameric, it degraded FITC-casein with a V_{max} of 4.4 min^{-1} enz^{-1} and a K_M of 16 μM (Figure 5.1C). Under conditions where Lon^{E240K} is largely dodecameric, it degraded FITC-casein with roughly similar values of V_{max} and K_M (Figure 5.1C; Table 5.1). Although the enzymatic properties of the Lon^{E240K} dodecamer are clearly different than the wild-type dodecamer, the E240K mutation does not appear to interfere with recognition of the FITC-casein substrate. We also found that Lon^{E240K} bound to a fluorescent sul20 peptide with wild-type affinity (Figure 5.1D), indicating that the binding site for this degron is not occluded in the dodecamer.

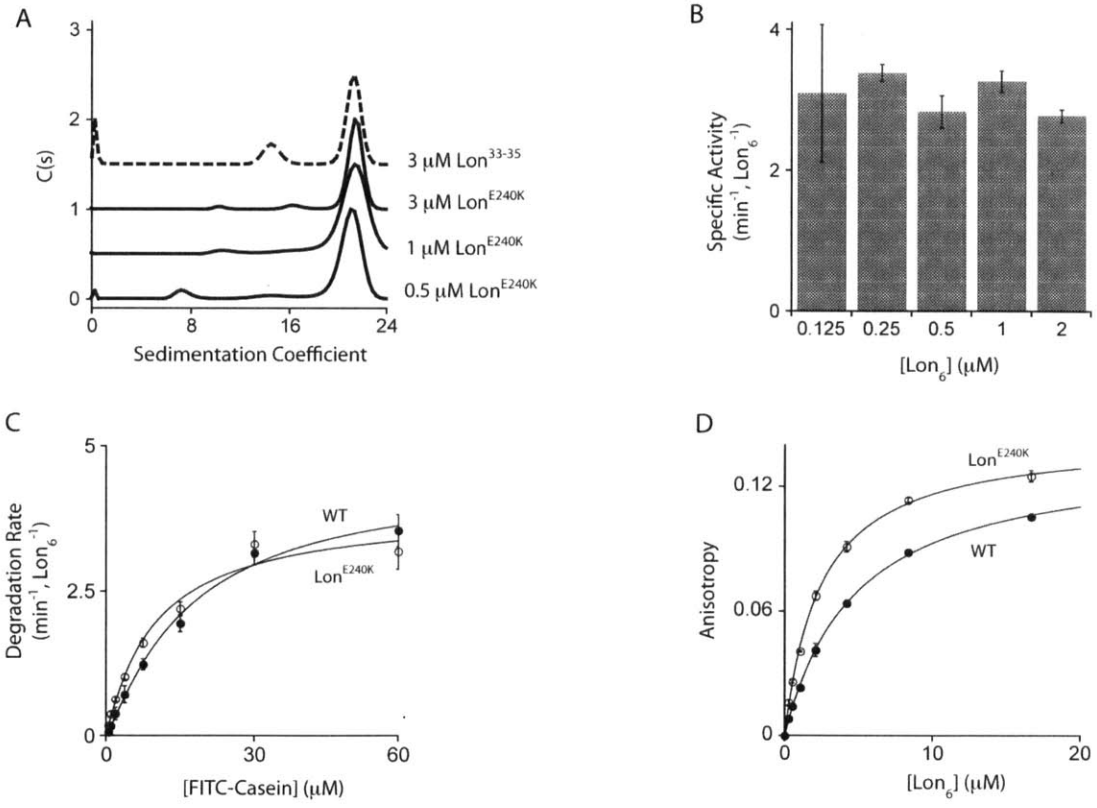

Figure 5.1 Lon^{E240K} forms a dodecamer that degrades FITC-casein well
(A) Sedimentation velocity analytical ultracentrifugation of Lon^{E240K}. Traces were normalized to have a maximum C(s) equal to 1 and were offset on the y-axis for clarity. For comparison, Lon^{33-35}, which forms both hexamers and dodecamers, is also shown. All concentrations listed are in hexamer equivalents. Experiments were performed at 20° C and 16,000 rpm in 50 mM HEPES (pH 7.5), 150 mM NaCl, 0.01 mM EDTA, 0.1 mM TCEP, 1 mM $MgCl_2$, and 0.1 mM ATPγS. (B) Specific activity of FITC-Casein (Sigma, 50 μM) degradation was assayed a different concentrations of Lon^{E240K} by changes in fluorescence (excitation 490 nm; emission 525 nm). Values are means ± SEM (n ≥ 3). (C) Substrate dependence of degradation of FITC-Casein by Lon or Lon^{E240K} (0.3 μM each). Values are means ± SEM (n = 3). The line is fit to the Hill form of the Michaelis-Menten equation $V = V_{Max} / (1+ (K_M/[S])^n)$. (D) Binding of fluorescein-labeled sul20 peptide (200 nM) by Lon or Lon^{E240K} was assayed by changes in fluorescence anisotropy (excitation 494 nm; emission 521 nm). Values are means ± SEM (n ≥ 2). All values are baseline corrected to have no anisotropy at 0 μM hexamer. Lines are fits to the equation $Anisotropy = (A_{Max} * [hexamer]) / (K_M + [hexamer])$.

Lon^{E240K} has selective defects in degrading model substrates

We purified RcsA (207 residues) for biochemical experiments, but it ran in the excluded volume of a gel-filtration column, suggesting that it forms soluble aggregates. Unfortunately, this protein was not degraded by wild-type Lon (not shown). It is possible that aggregation prevents recognition by Lon, or that another cellular protein is required for degradation of RcsA by Lon.

To determine if the E240K mutation affects degradation of other model substrates, we tested degradation of CM-titin^{I27}-sul20, CM-titin^{I27}-β20, cp6-GFP-sul20, and β20-cp6-GFP (Figures 5.2A-D). The unfolded CM-titin^{I27} substrates were degraded with V_{max} values roughly half of the wild-type values and with K_M's similar to or higher than the wild-type K_M (Figures 5.2A & B; Table 5.1). Thus, Lon^{E240K} displays minor defects in degrading unfolded CM-titin^{I27} substrates. By contrast, the native cp6-GFP substrates were degraded by Lon^{E240K} with V_{max} values reduced by

10-fold or more relative to wild-type Lon and with K_M's about twice the wild-type values (Figures 5.2C & D; Table 5.1). Clearly, Lon^{E240K} can degrade unfolded substrates, like CM-titin^{I27} and FITC casein, almost as well as wild-type Lon but displays severe defects in degrading the native GFP substrates. These defects are independent of the degron tag. The very low rate of degradation of cp6-GFP-sul20 by Lon^{E240K} is not a property shared by the wild-type dodecamer, as enzyme-normalized rates of cp6-GFP-sul20 degradation by wild-type Lon were relatively constant over concentrations where the dodecamer/hexamer ratio increased substantially (Figure 5.2E).

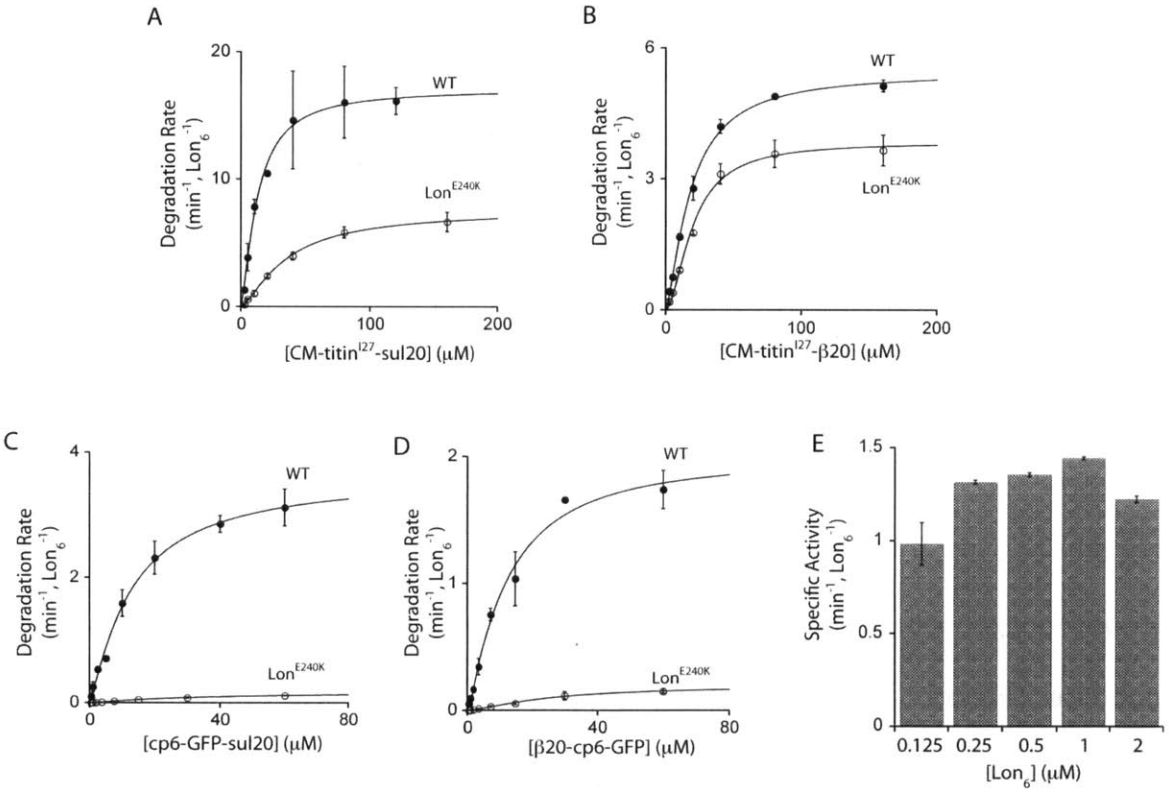

Figure 5.2 Substrate dependence of proteolysis

Substrate dependence of degradation of (A) CM-titin^{I27}-sul20, (B) CM-titin^{I27}-β20, (C) cp6-GFP-sul20, and (D) β20-cp6-GFP by wild-type Lon and Lon^{E240K} (0.3 μM each). Values are means ± SEM (n ≥ 2) and lines are fit to the Hill form of the Michaelis-Menten equation, $V = V_{Max} / (1+ (K_M/[S])^n)$. (E) Wild-type Lon degraded

cp6-GFP-sul20 (20 μM) with a relatively constant specific activity at concentrations spanning the hexamer-dodecamer equilibrium constant as assayed by changes in fluorescence (excitation 467 nm; emission 525 nm). Values are means ± SEM (n = 3).

		Proteolysis			ATPase			
Substrate	Lon Variant	V_{max} (min^{-1}, Lon_6^{-1})	Apparent Km (μM)	Hill Constant	V_{max} (min^{-1}, Lon_6^{-1})	Apparent Km (μM)	Hill Constant	ATP per substrate
FITC-casein	WT	4.4 ± 0.4	16 ± 4	1.2 ± 0.1	nd	nd	nd	nd
	E240K	3.9 ± 0.5	10 ± 3	1 ± 0.2	nd	nd	nd	nd
cp6-GFP-sul20	WT*	3.6 ± 0.3	13 ± 2	1.2 ± 0.1	217 ± 16	9.3 ± 1.3	1.2 ± 0.1	60
	E240K	0.17 ± 0.01	29 ± 3	1.2 ± 0.1	15 ± 1	10 ± 2	3.3 ± 1.7	88
β20-cp6-GFP	WT*	2.3 ± 0.3	21 ± 6	1.2 ± 0.2	143 ± 5	2.6 ± 0.2	1.7 ± 0.1	62
	E240K	0.19 ± 0.02	25 ± 4	1.5 ± 0.2	39 ± 4	7 ± 3	nd	205
CM-titin^{I27}-β20	WT	5.5 ± 0.1	18 ± 1	1.4 ± 0.1	174 ± 7	1.6 ± 0.2	0.9 ± 0.1	32
	E240K	3.8 ± 0.2	20 ± 2	1.8 ± 0.2	130 ± 10	9 ± 2	2.5 ± 0.9	34
CM-titin^{I27}-sul20	WT	17.1 ± 0.7	12 ± 1	1.4 ± 0.2	200 ± 20	0.6 ± 0.1	1.3 ± 0.4	12
	E240K	7.5 ± 0.3	35 ± 3	1.4 ± 0.1	240 ± 20	16 ± 2	1.3 ± 0.1	32
V13P-titin^{I27}-sul20	WT	6.7 ± 0.6	38 ± 6	1.6 ± 0.2	nd	nd	nd	nd
	E240K	1.5 ± 0.1	34 ± 4	1.8 ± 0.2	nd	nd	nd	nd
Y9P-titin^{I27}-sul20	WT	3.0 ± 0.1	29 ± 2	1.2 ± 0.1	nd	nd	nd	nd
	E240K	1.0 ± 0.1	53 ± 8	1.3 ± 0.1	nd	nd	nd	nd
V15P-titin^{I27}-sul20	WT	5.5 ± 0.4	55 ± 8	1.3 ± 0.1	nd	nd	nd	nd
	E240K	0.65 ± 0.06	31 ± 5	2.3 ± 0.6	nd	nd	nd	nd
titin^{I27}-sul20	WT	2 ± 0.1	29 ± 3	1.5 ± 0.2	118 ± 5	1 ± 0.1	1.1 ± 0.1	59
	E240K	1.2 ± 0.1	38 ± 3	1.4 ± 0.1	nd	nd	nd	nd

Table 5.1 Steady-state kinetic parameters for Lon^{E240K}

Errors are for non-linear-least-squares fitting. ATP per substrate was calculated by dividing V_{max} for ATP hydrolysis by V_{max} for proteolysis. * Data taken from Wohlever et al. 2013. nd = not determined

Does the E240K mutation cause a global defect in Lon's ability to unfold proteins? To test this possibility, we assayed Lon and Lon^{E240K} degradation of a set of sul20-tagged variants of native titin^{I27} with different thermodynamic and kinetic stabilities (Figure 5.3). Lon^{E240K} degraded titin^{I27}-sul20, the most stable variant, ~2-fold more slowly than wild-type Lon, a value similar to its defect in degradation of unfolded CM-titin^{I27}-sul20. Surprisingly, however, Lon^{E240K} showed larger defects relative to wild-type Lon in degrading the Y9P, V13P, and V15P titin^{I27}-sul20 variants. Because the native structures of these substrates are less stable than the

structure of titin^{I27}-sul20 (Kenniston et al. 2003), Lon^{E240K} appears to have greater difficultly in unfolding certain substrates in a way that does not correlate simply with their thermodynamic or kinetic stabilities.

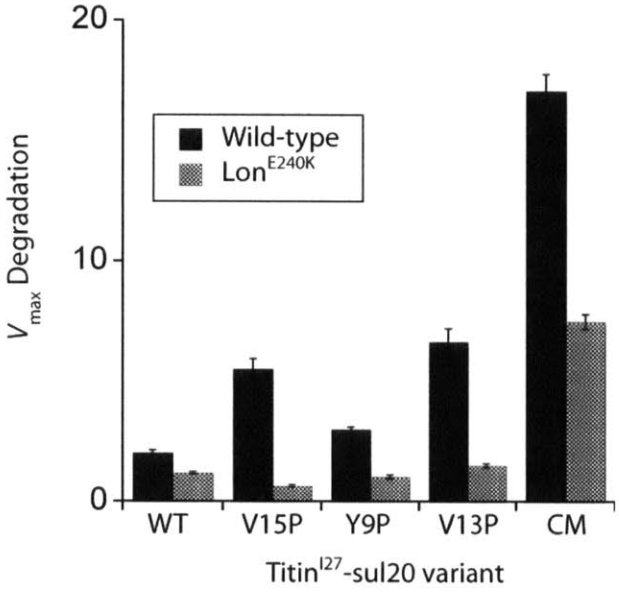

Figure 5.3 Degradation of titin^{I27}-sul20 stability variants
V_{max} values for degradation of titin^{I27}-sul20 stability variants by wild-type Lon and Lon^{E240K}. The substrates are arranged in decreasing order of stability, with the most stable substrate (wild-type titin^{I27}-sul20) on the left. Rates of degradation, based on formation of acid-soluble fragments of ^{35}S-labeled substrate, were averaged for each concentration of substrate for three independent experiments and fit to the Hill form of the Michaelis-Menten equation, $V = V_{max} / (1+(K_M/[S])^n)$. Errors are the uncertainty of V_{max} from non-linear-least-squares fitting.

Substrate-stimulated ATP hydrolysis by Lon^{E240K}

We measured stimulation of Lon^{E240K} ATP hydrolysis by CM-titin^{I27}-sul20 or CM-titin^{I27}-β20 and observed near wild-type levels of ATP hydrolysis at saturating substrate, although saturation required substantially higher substrate concentrations in comparison to wild-type Lon (Figures 5.4A & B; Table 5.1). By contrast, the maximal level of cp6-GFP-sul20 or β20-GFP-cp6 stimulation of ATP hydrolysis by Lon^{E240K} was reduced markedly in comparison with wild-type Lon

(Figures 5.4C & D; Table 5.1). This very low level of ATPase stimulation by native GFP substrate could be the reason for their very slow degradation, or the slow rate of degradation could be responsible for the low level of ATPase stimulation. In combination, our results show that the E240K mutation impairs degradation and substrate stimulation of ATPase activity in a manner that depends on the detailed properties of the substrate.

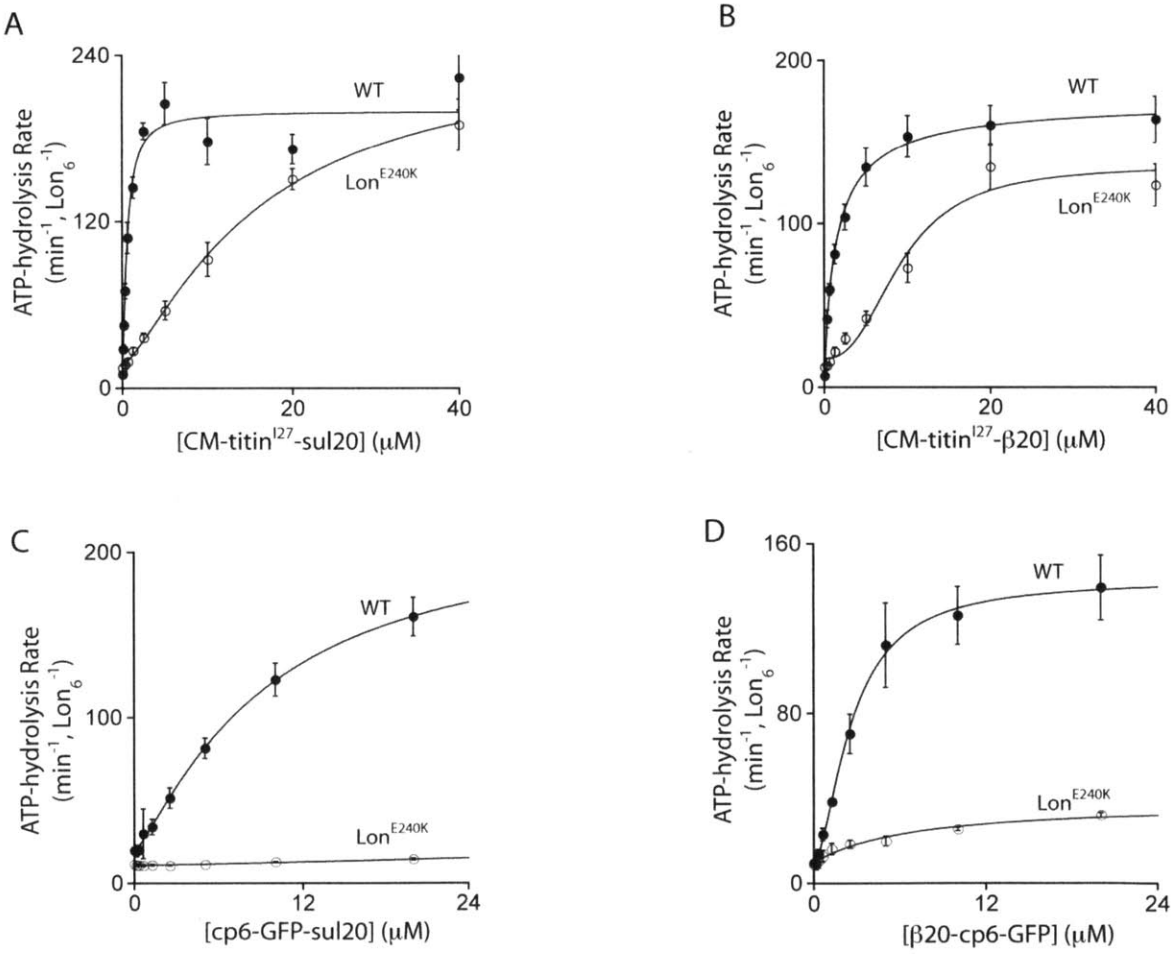

Figure 5.4 Substrate dependence of ATPase activity

ATPase stimulation of Lon or Lon^{E240K} (0.15 μM each) by (A) CM-$titin^{I27}$-sul20, (B) CM-$titin^{I27}$-β20, (C) cp6-GFP-sul20, and (D) β20-cp6-GFP. Values shown are means ± SEM ($n \geq 3$). Solid lines are fits to the Hill equation, rate = basal + amp/(1 + $(K_M/[S])^n$). V_{max} = basal + amp.

Activity of Lon^{E240K} in vivo

We cloned Lon, Lon^{E240K}, and proteolytically inactive Lon^{S679A} into low-copy plasmids, transformed *E. coli* strains harboring a deletion of the chromosomal *lon* gene, and tested for a phenotype that depends upon RcsA-activated transcription of enzymes that synthesize capsular polysaccharide (Gottesman et al. 1985; Van Melderen and Gottesman 1999). Cells without Lon secrete excess polysaccharide and are mucoid. As expected, strains expressing wild-type Lon formed non-mucoid colonies, whereas cells expressing Lon^{E240K}, or harboring an empty vector were mucoid (Figure 5.5A). Cells expressing Lon^{S679A} were also mucoid, suggesting that active RcsA degradation is needed to maintain low RcsA levels. Van Melderen and Gottesman (1999) found that expression of Lon^{S679A} from a high-copy plasmid was sufficient to keep RcsA levels low, suggesting that simple binding to RcsA can prevent transcriptional activation. It is likely that our assay requires RcsA degradation because the intracellular levels of Lon^{S679A}, expressed from a low-copy plasmid, are not high enough to bind a sufficient quantity of RcsA.

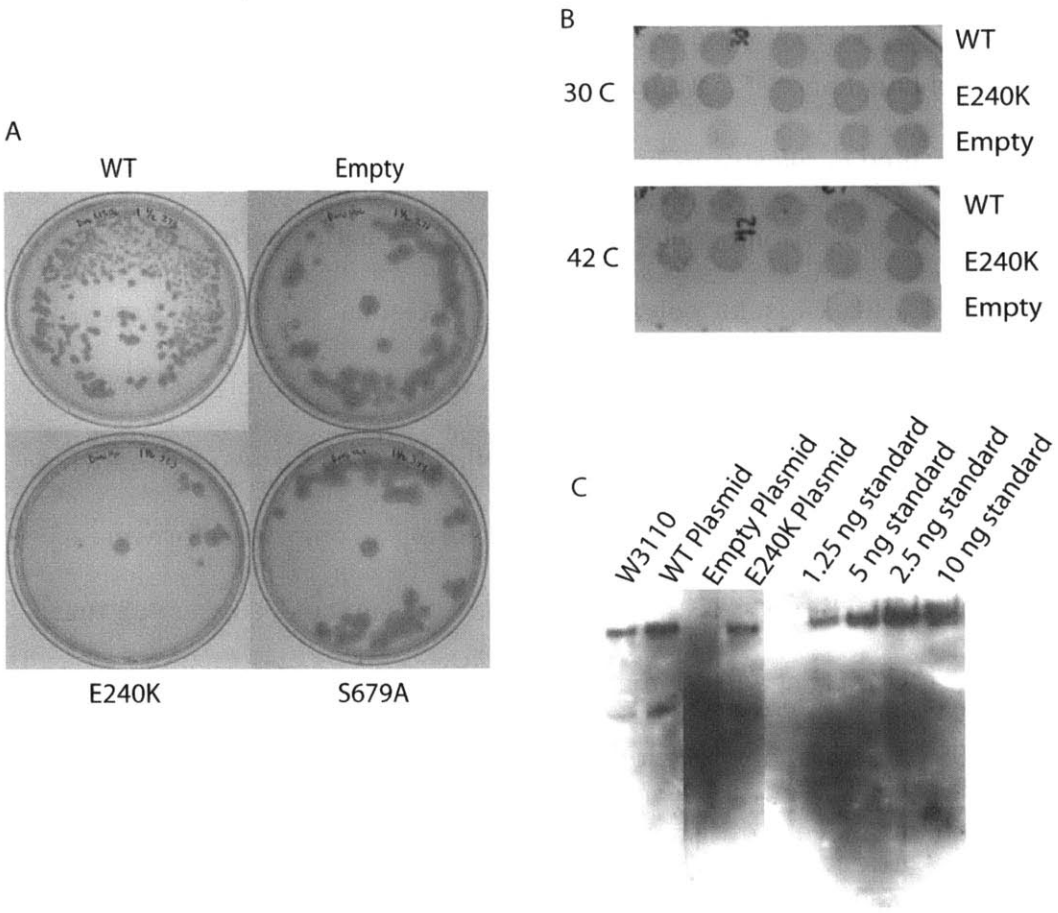

Figure 5.5 Activity of Lon^{E240K} *in vivo*
(A) Mucoidy assay to monitor the degradation of RcsA. W3110 *lon*::Kan^{R} cells were transformed with pBAD33 vectors expressing wild-type Lon, Lon^{E240K}, Lon^{S679A}, or with the empty vector. Failure to degrade RcsA leads to overproduction of capsular polysaccharides and formation of large colonies. (B) Rescue of cells from proteotoxic stress caused by deletion of *clpXP-lon*, reduced expression of DnaK/J, and growth at 42° C. Plasmids with wild-type Lon or Lon^{E240K} rescued growth equally well. (C) Western blot assay of expression levels of chromosomal Lon and Lon variants on pBAD33 plasmid.

We next tested if Lon^{E240K} could support the growth of cells subjected to proteotoxic stress by a lack of the ClpXP protease, low levels of the DnaK and DnaJ chaperones, and growth at 42 °C (Tomoyasu et al. 2001). Importantly, Lon^{E240K} supported growth as well as wild-type Lon, whereas cells with the empty vector died (Figure 5.5B). Western blotting showed that wild-type Lon and Lon^{E240K} were

expressed at similar levels, slightly higher than the level of Lon expressed from its normal chromosomal locus (Figure 5.5C).

Discussion

Lon^{E240K} was originally isolated in a genetic screen as a variant that failed to degrade RcsA but did degrade SulA (Ebel et al. 1999). Although this phenotype is consistent with an RcsA-binding defect, our results suggest that the failure of Lon^{E240K} to degrade RcsA could be substantially more complicated. For example, we find that Lon^{E240K} degrades native cp6-GFP-sul20 and β20-cp6-GFP very slowly under conditions of substrate saturation but degrades the unfolded substrates FITC-casein, CM-titin^{I27}-β20, and CM-titin^{I27}-sul20 with kinetic parameters similar to or only slightly slower than wild-type Lon. The E240K mutation did not cause a general defect in substrate unfolding as the V_{max} for degradation of native titin^{I27}-sul20 and unfolded CM- titin^{I27}-sul20 were both ~50% of the wild-type values.

The ability of substrates to stimulate ATP hydrolysis by Lon^{E240K} also differs in comparison with wild-type Lon. For example, Lon^{E240K} has wild-type levels of ATP hydrolysis when stimulated by CM-titin^{I27}-sul20 and CM-titin^{I27}-β20, but ~10-fold higher substrate concentrations are required for maximal stimulation. However, the maximal levels of cp6-GFP-sul20 or β20-GFP-cp6 stimulation of ATP hydrolysis by Lon^{E240K} are much lower than those for wild-type Lon These results suggest that the detailed properties of different substrate play important roles in determining how they interact with and control proteolysis by wild-type and mutant Lon enzymes. Such properties could include the geometric relationship between the position of

the primary degron and other regions of the substrate that might interact with different Lon sites, including the axial pore. The E240K mutation falls within a coiled-coil region of the N domain (residues 232-250) that lies between the binding site for the sul20 degron and the ATPase domain (Li et al. 2010) (Chapter 4) and undergoes nucleotide-dependent changes in conformation (Cheng et al. 2012). Thus, the E240K mutation may alter allosteric communication between substrate-binding sites in the N domain and the sites in the AAA+ module of Lon that bind and hydrolyze ATP to power translocation and unfolding.

Unexpectedly, we found that Lon^{E240K} exists almost exclusively as a dodecamer at concentrations where wild-type Lon forms hexamers and dodecamers. However, the wild-type Lon dodecamer degrades FITC-casein poorly, whereas Lon^{E240K} degrades this substrate well. These results suggest that Lon dodecamers, like hexamers, can exist in multiple conformational states, with the distinct properties of Lon^{E240K} arising from an alteration in this conformational ensemble. Structural studies of the Lon^{E240K} dodecamer could provide insight into regulation of Lon activity. Despite the catalytic defects observed for some substrates *in vitro*, Lon^{E240K} suppresses proteotoxic stress as well as wild-type Lon. As shown in Chapter 4, substrate translocation but not degradation is required for this phenotype, supporting a role for Lon as a chaperone. We assume that Lon^{E240K} also forms a stable dodecamer *in vivo*. Thus, dodecamers can apparently function both as a protease and as a chaperone.

Acknowledgements

We thank D. Barthelme, V. Baytshtok, S. Kim, E. Vieux, J. Chen, and members of the Sauer and Baker labs for reagents and helpful discussions, and D. Pheasant at the MIT Biophysical Instrumentation Center for help with the AUC experiments. This work was supported by National Institutes of Health grant AI-16982, and an National science Foundation Graduate Research Fellowship to M.L.W.

References

Baker TA, Sauer RT. 2006. ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends Biochem Sci* **31**: 647–653.

Bernstein SH, Venkatesh S, Li M, Lee J, Lu B, Hilchey SP, Morse KM, Metcalfe HM, Skalska J, Andreeff M, et al. 2012. The mitochondrial ATP-dependent Lon protease: a novel target in lymphoma death mediated by the synthetic triterpenoid CDDO and its derivatives. *Blood*.

Botos I, Melnikov EE, Cherry S, Tropea JE, Khalatova AG, Rasulova F, Dauter Z, Maurizi MR, Rotanova TV, Wlodawer A, et al. 2004. The catalytic domain of Escherichia coli Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J Biol Chem* **279**: 8140–8148.

Breidenstein EBM, Janot L, Strehmel J, Fernandez L, Taylor PK, Kukavica-Ibrulj I, Gellatly SL, Levesque RC, Overhage J, Hancock REW. 2012. The Lon Protease Is Essential for Full Virulence in Pseudomonas aeruginosa. *PLoS ONE* **7**: e49123.

Brown PH, Schuck P. 2006. Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys J* **90**: 4651–4661.

Cha S-S, An YJ, Lee C-R, Lee HS, Kim Y-G, Kim SJ, Kwon KK, De Donatis GM, Lee J-H, Maurizi MR, et al. 2010. Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *EMBO J* **29**: 3520–3530.

Cheng I, Mikita N, Fishovitz J, Frase H, Wintrode P, Lee I. 2012. Identification of a Region in the N-Terminus of Escherichia coli Lon That Affects ATPase, Substrate Translocation and Proteolytic Activity. *J Mol Biol*.

Chung CH, Goldberg AL. 1981. The product of the lon (capR) gene in Escherichia coli is the ATP-dependent protease, protease La. *Proc Natl Acad Sci USA* **78**: 4931–4935.

DAVIS BD, MINGIOLI ES. 1950. Mutants of Escherichia coli requiring methionine or vitamin B12. *J Bacteriol* **60**: 17–28.

Duman RE, Löwe J. 2010. Crystal structures of Bacillus subtilis Lon protease. *J Mol Biol* **401**: 653–670.

Ebel W, Skinner MM, Dierksen KP, Scott JM, Trempy JE. 1999. A conserved domain in Escherichia coli Lon protease is involved in substrate discriminator activity. *J Bacteriol* **181**: 2236–2243.

Goldberg AL, Moerschell RP, Chung CH, Maurizi MR. 1994. ATP-dependent protease La (lon) from Escherichia coli. *Methods in Enzymology* **244**: 350–375.

Gottesman S, Trisler P, Torres-Cabassa A. 1985. Regulation of capsular polysaccharide synthesis in Escherichia coli K-12: characterization of three regulatory genes. *J Bacteriol* **162**: 1111–1119.

Gottesman S, Zipser D. 1978. Deg phenotype of Escherichia coli lon mutants. *J Bacteriol* **133**: 844–851.

Gur E, Sauer RT. 2009. Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine. *Proceedings of the National Academy of Sciences* **106**: 18503–18508.

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

Hanson PI, Whiteheart SW. 2005. AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol* **6**: 519–529.

Higashitani A, Ishii Y, Kato Y, Koriuchi K. 1997. Functional dissection of a cell-division inhibitor, SulA, of Escherichia coli and its negative regulation by Lon. *Mol Gen Genet* **254**: 351–357.

Ingmer H, Brøndsted L. 2009. Proteases in bacterial pathogenesis. *Res Microbiol* **160**: 704–710.

Kenniston JA, Baker TA, Fernandez JM, Sauer RT. 2003. Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* **114**: 511–520.

Lee AY-L, Hsu C-H, Wu S-H. 2004. Functional domains of Brevibacillus thermoruber lon protease for oligomerization and DNA binding: role of N-terminal and sensor

and substrate discrimination domains. *J Biol Chem* **279**: 34903–34912.

Li M, Gustchina A, Rasulova FS, Melnikov EE, Maurizi MR, Rotanova TV, Dauter Z, Wlodawer A. 2010. Structure of the N-terminal fragment of Escherichia coli Lon protease. *Acta Crystallogr D Biol Crystallogr* **66**: 865–873.

Li M, Rasulova F, Melnikov EE, Rotanova TV, Gustchina A, Maurizi MR, Wlodawer A. 2005. Crystal structure of the N-terminal domain of E. coli Lon protease. *Protein Sci* **14**: 2895–2900.

Luce K, Osiewacz HD. 2009. Increasing organismal healthspan by enhancing mitochondrial protein quality control. *Nat Cell Biol* **11**: 852–858.

Melnikov EE, Andrianova AG, Morozkin AD, Stepnov AA, Makhovskaya OV, Botos I, Gustchina A, Wlodawer A, Rotanova TV. 2008. Limited proteolysis of E. coli ATP-dependent protease Lon - a unified view of the subunit architecture and characterization of isolated enzyme fragments. *Acta Biochim Pol* **55**: 281–296.

Robertson GT, Kovach ME, Allen CA, Ficht TA, Roop RM. 2000. The Brucella abortus Lon functions as a generalized stress response protease and is required for wild-type virulence in BALB/c mice. *Mol Microbiol* **35**: 577–588.

Rotanova TV, Botos I, Melnikov EE, Rasulova F, Gustchina A, Maurizi MR, Wlodawer A. 2006. Slicing a protease: structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains. *Protein Sci* **15**: 1815–1828.

Sauer RT, Baker TA. 2011. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **80**: 587–612.

Tomoyasu T, Mogk A, Langen H, Goloubinoff P, Bukau B. 2001. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the Escherichia coli cytosol. *Mol Microbiol* **40**: 397–413.

Torres-Cabassa AS, Gottesman S. 1987. Capsule synthesis in Escherichia coli K-12 is regulated by proteolysis. *J Bacteriol* **169**: 981–989.

Van Melderen L, Gottesman S. 1999. Substrate sequestration by a proteolytically inactive Lon mutant. *Proc Natl Acad Sci USA* **96**: 6064–6071.

Vasilyeva OV, Kolygo KB, Leonova YF, Potapenko NA, Ovchinnikova TV. 2002. Domain structure and ATP-induced conformational changes in Escherichia coli protease Lon revealed by limited proteolysis and autolysis. *FEBS Lett* **526**: 66–70.

Wohlever ML, Nager AR, Baker TA, Sauer RT. 2013. Engineering fluorescent protein substrates for the AAA+ Lon protease. *Protein Eng Des Sel.*

Chapter 6

Perspectives and Future Directions

Introduction

In this thesis, I have developed folded fluorescent substrates for Lon, described collaborative experiments that show that Lon equilibrates between a hexamer and a dodecamer with altered substrate profiles for each oligomer, identified a substrate-binding site within the N domain that allosterically regulates Lon activity, and provided evidence that *E. coli* Lon may function as a chaperone *in vivo*. Lon activity is probably regulated in multiple ways. In Chapter 3, my co-authors and I proposed that Lon alters its substrate profile by changing oligomeric states. In Chapters 4 & 5, I proposed that substrate binding to the N domain shifts Lon between states with differential protease and chaperone activities. Here, I describe experiments and/or tools that could be used to test these models and further our understanding of the role of Lon in proteostasis. For many of these experiments I have gathered materials and/or shown proof of principle.

Study of chaperone mechanism *in vitro*

One of my most interesting results is the observation that proteolytically inactive Lon^{S679A} rescues Δ*lon* cells from heat shock and proteotoxic stress, supporting a model in which Lon can function as a chaperone as well as a protease. The next question is how does Lon act as a chaperone? As outlined in Chapter 1, chaperones can function as holdases, unfoldases, or refolding enzymes. Because Lon contains a AAA+ unfoldase domain and a functional pore loop appears necessary for chaperone function, an obvious place to begin is by examining the ability of Lon to serve as an unfolding chaperone, similar to ClpB.

Because the protease and ATPase domains are contained within a single polypeptide, assaying Lon-mediated substrate unfolding that is independent of degradation requires use of proteolytically inactive variants, having a misfolded substrate that can be shown to unfold and refold faster than it is degraded, or having a native substrate that can be shown to unfold at a faster rate than it is degraded. In terms of the second possibility, a truncated version of *Photinus pyralis* luciferase that has undergone several freeze-thaw cycles is misfolded and inactive but does not aggregate (Sharma et al. 2010). When unfolded, either by a chemical denaturant or by DnaK/J, and allowed to refold in solution, this luciferase variant regains bioluminescence. Because this assay measures an increase in bioluminescence from background, it is much more sensitive than assays that measure loss of a signal. Moreover, any unfolding event that is paired with degradation will not give a signal, allowing for easy separation of these two catalytic activities. I have obtained and cloned this luciferase substrate with either a sul20 or β20 degron, but I have not yet characterized these substrates. Based on our model, Lon assumes a chaperone-like conformation upon allosteric activation by the β20 degron. However, the hydrophobic β20 degron may alter the folding properties of the luciferase substrate, so an orthogonal approach should also be used.

In terms of the third possibility, one could use the well-characterized substrate titin^{I27}-β20 and measure the rate of modification of the two buried cysteine residues by a thiol-reactive fluorophore, such as fluorescein-maleimide. SDS-PAGE could be used to separate proteolytic fragments with modified cysteines from fluorescent, full-length titin, thus distinguishing between chaperone and

protease activity. If full-length titin^{I27}-β20 were modified by maleimide faster in the presence than absence of Lon, then this result would provide evidence for Lon-mediated unfolding that is uncoupled from degradation.

Identification of β20 binding site

The identification of the sul20-binding site in the Lon N domain provided a powerful tool for examining regulation of Lon activity. Because the β20 degron leads to different allosteric regulation, identification of the β20 binding sites would allow for further exploration of degron-mediated regulation. An obvious starting point for this endeavor would be to repeat the cross-linking experiments described in Chapter 4. However, this approach would likely be quite challenging given the hydrophobic nature of the β20 degron. Indeed, attaching the sulfo-SBED cross-linker to any protein dramatically reduces solubility.

An orthogonal approach would be to use a genetic screen to select for mutants that are capable of degrading sul20-tagged substrates, but not β20-tagged substrates. DHFRII-β20 and trimethoprim can be used to select for Lon variants that are unable to degrade β20-tagged substrates. Trimethoprim is an inhibitor of the essential enzyme dihydrofolate reductase (DHFR), but the variant DHFRII is resistant to trimethoprim. Growing cells in the presence of trimethoprim makes DHFRII essential. Attaching the β20 degron to DHFRII will select for Lon variants that are unable to recognize and degrade the β20-tagged substrate.

A mutant of phenylalanine-tRNA synthetase (PheS) with the sul20 degron could be used to counter-select against Lon mutants that have general proteolytic

defects not specific to binding of the β20 degron (truncations, protease dead mutants, etc.). The reduced substrate specificity of PheS allows misincorporation of the amino acid analog *p*-chlorophenylalanine, which, when added to the selection media, leads to cell death. The only way for cells to survive is for PheS-sul20 to be degraded. To verify the efficacy of this screen, it can be set up in reverse (DHFR-sul20, PheS-β20) with Lon^{33-35} as a control. All components necessary for this screen have been cloned, but they have not yet been tested or optimized.

Although unlikely, it is possible that there is no defined allosteric binding site for the β20 degron, and instead this degron only binds to the pore. It will be exceedingly difficult to rigorously show that there is no β20-binding site outside of the pore loops, but this possibility should be considered. Alternatively, the β20-binding site may not be essential for degradation of the β20-tagged substrates, which would also cause this genetic selection to fail. Despite these challenges, identification of the β20-binding site should be a high priority, as it will allow for detailed mechanistic and physiological studies of Lon.

Structural studies and development of tools for mechanistic investigations

A major hindrance to mechanistic studies on Lon is the lack of a high-resolution structure of the full-length enzyme. There are many possible reasons why Lon has failed to crystallize, but one likely culprit is oligomeric heterogeneity. The Lon^{E240K} mutant appears to exist exclusively as a dodecamer, rather than equilibrating between a hexamer and a dodecamer like wild-type Lon. This reduced oligomeric heterogeneity may allow for crystallization of the full-length enzyme. An

alternative approach is to use cryo-electron microscopy (cryo-EM) to obtain a high-resolution structure. Indeed, initial cryo-EM results (Figure 5.1) suggest that this is a promising approach. Surprisingly, cryo-EM results with wild-type Lon suggest that dodecameric Lon exists in multiple conformations, including a "lock-washer" conformation in addition to a closed hexameric ring (Ellen Vieux and James Chen, personal communication). The significance of multiple conformations of the dodecamer is an interesting area of further investigation, which may be aided by the Lon^{E240K} mutant.

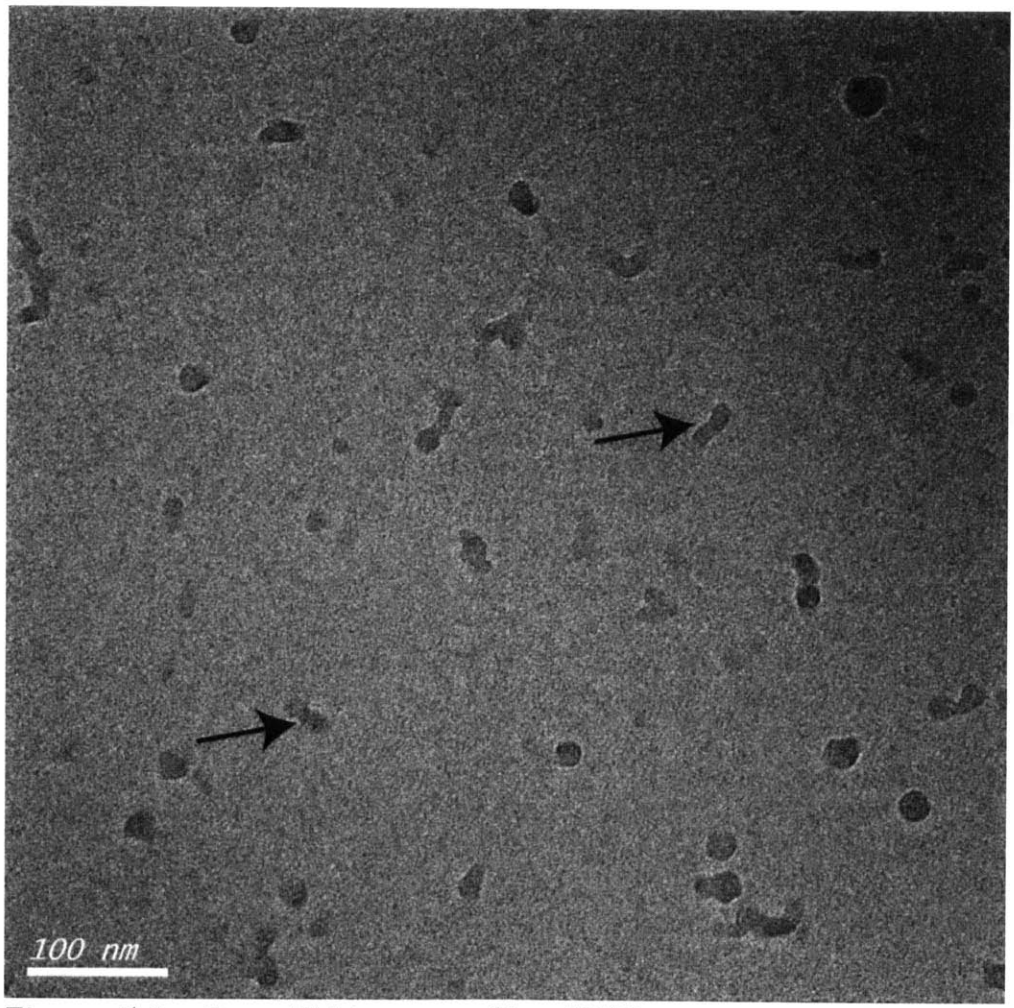

Figure 5.1 Cryo-EM images of Lon^{E240K}

Initial cryo-EM images for Lon^{E240K} (0.2 mg/mL) with ATPγS (200 μM) and sul20 peptide (4 μM). Arrows distinguish apparent Lon dodecamers from ice contaminants. Experiment done in collaboration with James Z. Chen.

Obtaining a high-resolution structure of full-length Lon will open the door for the development of new tools for dissecting structure-function relationships for the Lon hexamer and dodecamer. Mutants could be designed that prevent dodecamer formation, thereby allowing more rigorous testing of the substrate-gating model proposed in Chapter 3. Currently, the only way to only way to obtain hexameric Lon is to add polyphosphate (Figure 5.2), which has pleiotropic effects on Lon activity.

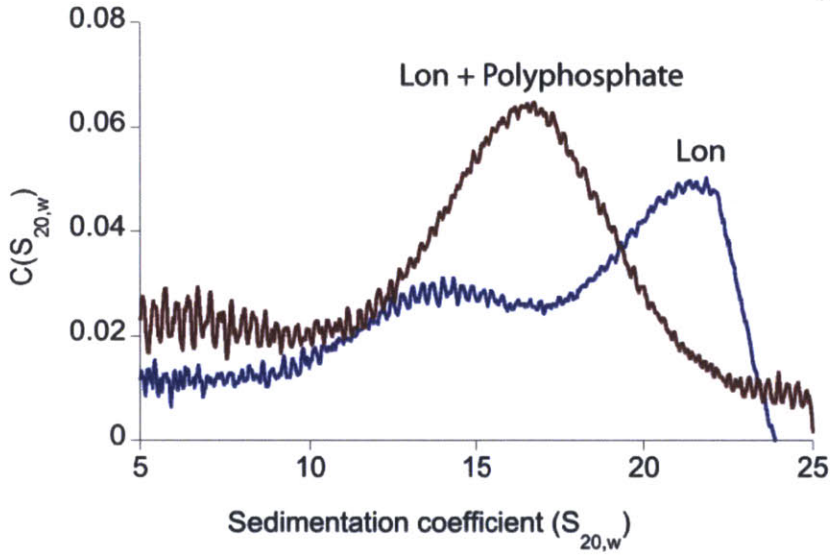

Figure 5.2 Polyphosphate changes Lon oligomerization
Sedimentation Velocity Analytical Ultracentrifugation with 250 nM Lon^{S679A} (hexamer equivalents) with or without 100 uM polyphosphate (n = 35, Sigma). The peak at 21 S corresponds to a dodecamer while the peak around 15 S corresponds to a hexamer. 100 μM Polyphosphate appears to shift Lon exclusively into a hexameric conformation. Experiment was performed at 20° C at 17,000 RPM in 50 mM Hepes (pH 7.5), 150 mM NaCl, 0.01 mM EDTA, 1 mM MgCl2, and 0.1 mM ATPγS. Experiment was performed and analyzed as described in chapter 3. Experiment done in collaboration with Ellen F. Vieux.

Another tool that could be developed based on a high-resolution structure would be to introduce inter-subunit disulfide bonds to covalently link adjacent

subunits. If multiple cysteine pairs that form disulfides are possible, then this would allow for the insertion of mutants into specific subunits in the Lon hexamer/dodecamer (Figure 5.3A). Alternatively, inter-subunit disulfide bonds could be designed to trap Lon in the lock-washer conformation, allowing the function of this conformation to be more rigorously investigated.

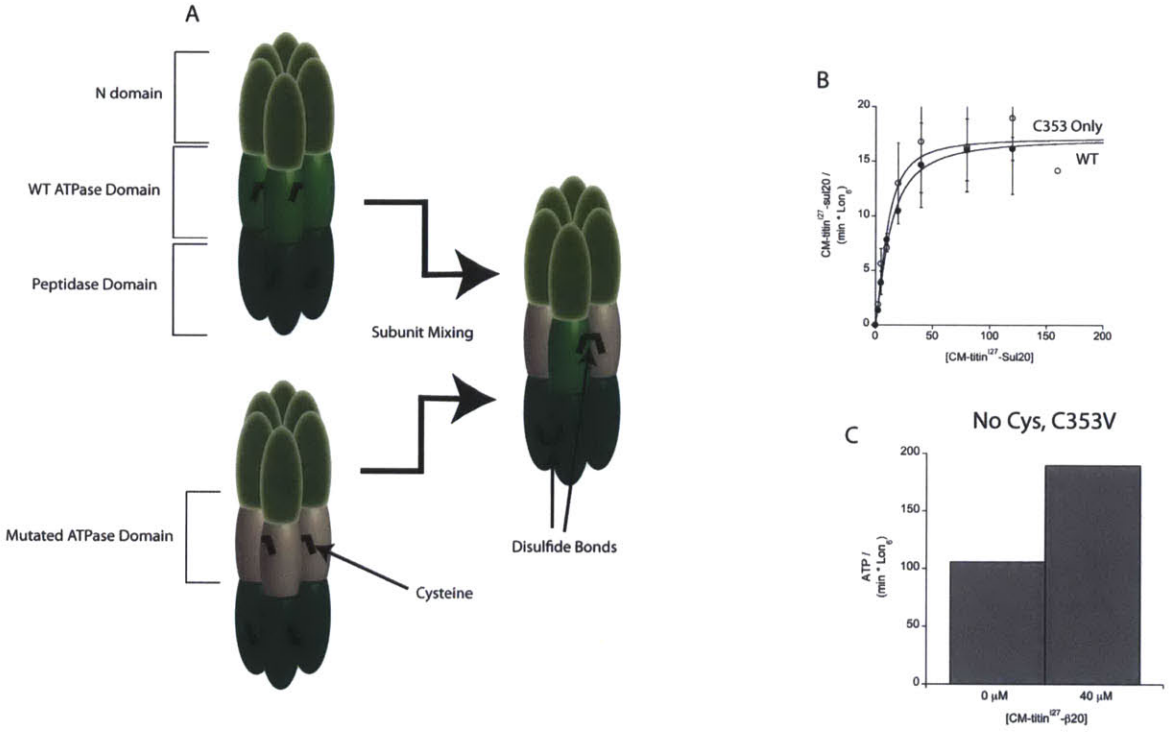

Figure 5.3 Cysteines in Lon

(A) Diagram showing how two different sets of designed disulfides can be mixed together to generate a covalently linked Lon hexamer with a mutation in the ATPase domains of alternating subunits. (B) Lon with all native cysteines except C353 mutated to serine (C353 only, 0.3 μM) behaves indistinguishably from wild-type Lon (0.3 μM) in degradation of CM-titin^{I27}-sul20. Values shown are mean ± SEM (n≥2). Solid lines are fits to the Hill form of the Michaelis-Menton equation $V = V_{Max} / (1+ (K_M/[S])^n)$. Assays were performed as described in Chapter 4. (C) Lon with all native cysteines mutated to serine except and C353 mutated to valine (no Cys, C353V) retains ATPase activity, which can be stimulated by CM-titin^{I27}-β20. Assays were performed as described in Chapter 4.

I have developed a cysteine free Lon variant that retains protease and ATPase activity. Lon contains six native cysteine residues, five of which (C39, C552,

C617, C685, C691) can be mutated to serine with no effect on activity (Figure 5.3B). C353 is predominantly buried, close to the ATP binding site, and is highly conserved. Because it is buried, C353 is not highly reactive. However, if experiments require no cross-reactivity, C353 can be mutated to valine and still retain activity (Figure 5.3C)

Proteomic approaches

As outlined in the Introduction, the Lon protease has many physiological substrates, but a complete list of substrates is still lacking. Proteomic approaches have been used to identify a comprehensive list of substrates for other proteases, such as ClpXP (Flynn et al. 2003). In these experiments, the protease-active site mutant is used to create a substrate trap; where substrates are translocated into the degradation chamber but not degraded. Pull downs and subsequent mass spectrometry allow for identification of substrates. Elizabeth Oakes, a former postdoc in Tania Baker's lab, attempted a similar approach with Lon (personal communication). Over 550 potential substrates were identified, but these data were never published due to some challenges with reproducibility.

The development of fluorescent Lon substrates that work *in vivo* (Chapter 2) allows for an orthogonal approach to substrate identification. I have done initial proof of principle experiments that demonstrate that co-expression of a known Lon substrate with cp6-GFP-sul20 protects the fluorescent substrate from degradation, presumably via competition for binding to Lon (Figure 5.4A). Importantly, co-expression of a model substrate that is not a Lon substrate does not protect cp6-GFP-sul20 from degradation (Figure 5.4B). Furthermore, I have shown that this

effect can be observed with fluorescent activated cell sorting (FACS), providing a ready means to separate fluorescent and non-fluorescent cells. If a library with most or all *E. coli* open-reading frames was transformed into cells containing cp6-GFP-sul20, then FACS and deep sequencing could be used to enrich for and identify Lon substrates on a proteomic level (Figure 5.4C).

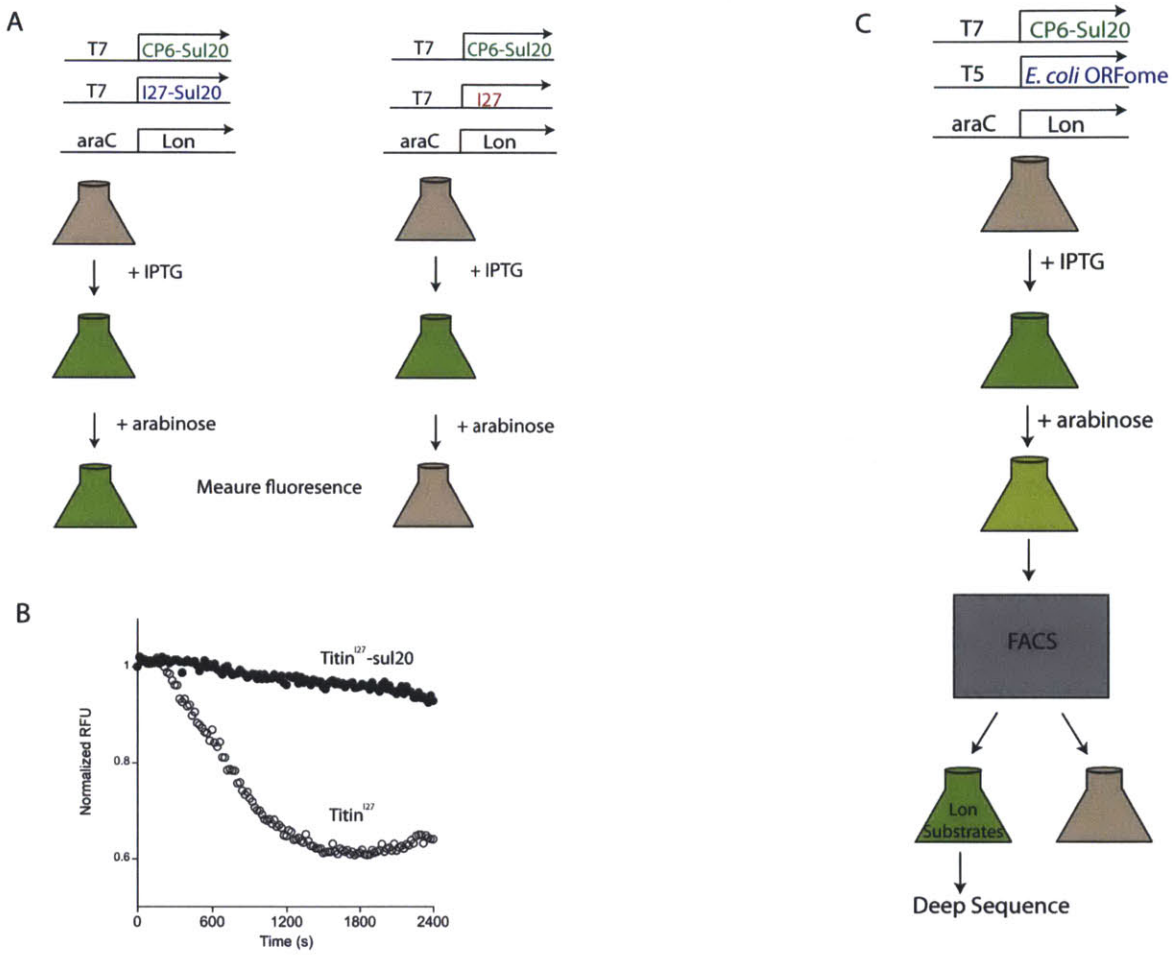

Figure 5.4 *In vivo* competition for cp6-GFP-sul20 degradation
(A) Diagram showing experimental set up for competition assays *in vivo*. Δlon cells contain three separate plasmids. cp6-GFP-sul20 and titin^{I27} variants are under control of the T7 promoter, whereas Lon is under control of an arabinose-inducible promoter. cp6-GFP-sul20 and titin^{I27} variants are induced with IPTG for 2 h prior to induction of Lon expression with arabinose. (B) Co-expression of a known Lon substrate, titin^{I27}-sul20, competes with cp6-GFP-sul20 for Lon-mediated degradation, causing the cells to remain fluorescent. Co-expression of a titin ^{I27}, which is not degraded by Lon, leads to a loss of fluorescence. T = 0 time point

corresponds to induction of Lon expression with arabinose. Overall experimental setup is similar to that described in chapter 2. Fluorescence was measured in a fluorescent plate reader with excitation at 467 nm and emission at 511 nm. Cells used for this experiment are ER2566, which are *lon* deficient. (C) Diagram showing how competition for Lon degradation *in vivo* can be used to identify Lon substrates. Similar experimental setup to figure 5.4A, except instead of expression of known Lon substrates, the *E. coli* ORFome is expressed from the second IPTG-inducible plasmid.

Rather than attempting to identify all possible Lon substrates, another approach would be to identify substrates that interact with the N domain sul20 binding site (including residues 33-35) as identified in Chapter 4. The *in vivo* experiments in Chapter 4 suggest that substrates other than SulA interact with this site in the N domain. Comparative enrichment experiments using the wild-type or $Lon^{33\text{-}35}$ trap variants could provide a list of substrates that interact with this binding site. Alternatively, if the competition assay *in vivo* is optimized with β20-cp6-GFP as the fluorescent substrate, instead of cp6-GFP-sul20, then a comparative analysis with wild-type and $Lon^{33\text{-}35}$ could also provide a list of substrates that interact with this site. The information gleaned from these experiments could provide insights into how Lon prioritizes unfolding and/or degradation of its many substrates. Provided that structural or other studies lead to the generation of Lon variants that are exclusively dodecameric, like Lon^{E240K}, these strategies could also be used to investigate the substrate profiles of the hexamer and dodecamer and provide a more rigorous test of the substrate-gating model presented in Chapter 3.

References

Flynn JM, Neher SB, Kim YI, Sauer RT, Baker TA. 2003. Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Molecular Cell* **11**: 671–683.

Sharma SK, De los Rios P, Christen P, Lustig A, Goloubinoff P. 2010. The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. *Nat Chem Biol* **6**: 914–920.