Insights into the role and mechanism of the AAA+ adaptor ClpS

by

Jennifer Yuan Hou

Sc.B. Biochemistry
Brown University, 2002

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

DOCTOR OF PHILOSOPHY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2009

Signature of Author:_____
Department of Biology
May 22, 2009

Certified by:_____
Tania A. Baker
E. C. Whitehead Professor of Biology
Thesis Supervisor

Accepted by:_____
Stephen P. Bell
Professor of Biology
Co-Chair, Graduate Committee

Insights into the role and mechanism of the AAA+ adaptor ClpS

by

Jennifer Yuan Hou

Submitted to the Department of Biology on May 22, 2009 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy at the Massachusetts Institute of Technology

ABSTRACT

Protein degradation is a vital process in cells for quality control and participation in regulatory pathways. Intracellular ATP-dependent proteases are responsible for regulated degradation and are highly controlled in their function, especially with respect to substrate selectivity. Adaptor proteins that can associate with the proteases add an additional layer of control to substrate selection. Thus, understanding the mechanism and role of adaptor proteins is a critical component to understanding how proteases choose their substrates.

In this thesis, I examine the role of the intracellular protease ClpAP and its adaptor ClpS in *Escherichia coli*. ClpS binds to the N-terminal domain of ClpA and plays dual roles in ClpAP substrate selectivity: ClpS inhibits the degradation of some substrates such as ssrA-tagged proteins and enhances the degradation of other substrates such as N-end-rule proteins. We wished to elucidate how ClpS influences ClpAP substrate selection, and found that the stoichiometry of ClpS binding to ClpA is one level of regulation. Furthermore, we demonstrated that the N-terminal extension of ClpS is vital for the adaptor's role in delivering N-end-rule substrates to ClpAP for degradation, but this extension is not required for inhibition of ssrA-tagged proteins. Truncation studies of the ClpS N-terminal extension showed a dramatic length-dependence on N-end-rule protein delivery, and the chemical composition of this portion of ClpS also affected the ability to degrade N-degron-bearing substrates. Evidence suggests that ClpS allosterically affects the ClpA enzyme, causing a modulation in substrate specificity, and preliminary studies localized the point of contact by the ClpS N-terminal extension to the ClpA pore region. ClpS therefore represents a new type of adaptor protein that modulates substrate selection allosterically, rather than simply recruiting and tethering substrates to the protease.

To further understand the role of ClpS and ClpAP in the cell, we conducted a proteomic-based search for ClpS-dependent ClpAP substrates. A list of putative substrates was generated from these experiments, and we believe that ClpAP plays a key role in quality control, perhaps through the degradation of N-end-rule substrates. Combined with mechanistic studies, these physiological studies aid in the understanding of how ClpS influences substrate recognition by ClpAP.

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

ACKNOWLEDGEMENTS

There are so many people to thank who have helped me through my graduate school journey. I give great thanks to my advisor Tania Baker whose encouragement, patience, and extraordinary scientific intellect have tremendously shaped my project and my overall scientific education. She has fostered a supportive and engaging lab for which I will always be grateful. I thank my co-advisor Bob Sauer who has always been an unwavering source of support and beneficial advice. His passion for biochemistry is truly infectious. I am also grateful for the help and advice of my thesis committee: Thomas Schwartz, Phillip Sharp, and Daniel Finley.

Many thanks are given to all the members of the Baker Lab, past and present, who have created a wonderful environment in which to work. The Baker Lab has been an immense source of scientific expertise, support, and laughs. With whom else can I discuss protein mechanism and cupcakes in the same conversation?

Thanks are also extended to Biograd 2003. I feel so lucky to have met such a fantastic group of people. I have enjoyed the road we have shared together and expect to hear many great things from everyone in the future.

Lastly, I thank my dear Gary. He has been a constant source of support over the years, whether from close or from afar. His love, patience, and home-cooked dinners have been invaluable to me over the years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

Introduction

SIGNIFICANCE OF REGULATED PROTEOLYSIS

Proteins are involved in virtually every cellular function, and it is essential that the amount, folding, processing, and localization of proteins are exquisitely controlled. Numerous mechanisms regulate the protein levels in the cell, including transcription, translation, post-translational processing, and degradation. Transcription is a critical step in the regulation of intracellular protein levels because control at this step determines which gene products are produced and how many copies are synthesized. Once the cell decides to up-regulate a gene, transcription, RNA processing in eukaryotes, translation, and sometimes localization and post-translational processing events are necessary before attaining the functional protein. Tuning down levels of protein expression typically involves alleviating the activation signal or inducing a repressor. However, control at the step of transcription still allows existing RNA transcripts to be translated into protein. Therefore, if rapid down-regulation of a gene product is required, direct degradation of the existing protein is a convenient strategy often employed by cells.

One study showed that 2.7% of proteins are degraded per generation in *Escherichia coli* growing in logarithmic phase (Fox and Brown 1961). There are many reasons why it would be advantageous for a cell to turn over a protein. A cell assaulted by external stress needs to respond quickly in order to survive. For example, degradation of a cellular regulator could result in transduction of a signal to activate stress-response genes. Protein degradation is also involved in less life-and-death situations in which the cell adjusts its proteome while continuously sampling its environment, such as for nutrients or developmental cues. It can sometimes be deleterious for too much of a particular protein to be present in the cell; proteins not wanted or required for active use can be degraded to provide building blocks for up-regulated gene products. The ability to fine-tune the levels of proteins allows the adaptability to external stimuli as well as the progression of cellular programs. Several examples illustrate the role of proteolysis in these functionalities.

In eukaryotic cells, various cyclin-dependent kinases are responsible for carrying out cell cycle progression. As their name implies, these kinases must bind to their partner cyclin protein for activity. Cyclin levels oscillate with the cell cycle, and it was discovered that cyclins accumulate and then are specifically and periodically degraded,

thereby deactivating their partnered kinases at indicated times during the cell cycle. Different cyclins are degraded at different stages to allow cell cycle progression (Glotzer et al. 1991; Nigg 1995). Thus, these precise proteolytic events are crucial in maintaining the cell cycle clock.

Another example of regulated proteolysis taken from eukaryotes is the response to ethylene by *Arabidopsis*. Exogenous ethylene can result in a range of physiological and morphological changes in plants including leaf and flower senescence, fruit ripening, and adaptation to stress and pathogenic infections. One study reports on the rapid accumulation of a nuclear signaling protein in *Arabidopsis* involved in the ethylene response pathway. It was discovered that this protein, EIN3, is constitutively produced and degraded in the absence of ethylene. The exposure to ethylene blocks EIN3 from being recognized for degradation through an undefined mechanism, thereby allowing the protein to accumulate and turn on its target genes for ethylene response (Guo and Ecker 2003). The employment of protein degradation in this mechanism allows the plant to rapidly respond to environmental signals.

Although protein degradation plays key roles in cellular function, regulated degradation is also essential for clearing aberrant or damaged protein products while ignoring functional ones. The cell needs to prevent irreversibly-misfolded or otherwise damaged proteins from inducing uncontrolled aggregation and disrupting normal processes. The challenge for this protein-clearance process is to determine which proteins should be degraded. In general, the cell uses several strategies for discriminating proteins slated for degradation from stable proteins. One common mechanism for this discrimination is the revealing of a latent degradation signal in the targeted protein. This signal could be a stretch of hydrophobic residues that would normally be in the interior of a properly-folded protein, a degradation motif that is revealed after an upstream processing event, or a sequence in a protein that is hidden when associated with a partner protein (Gottesman and Maurizi 1992). Degradation tags are discussed more thoroughly below.

Regulated proteolysis can broadly be defined as the process by which proteins are selectively degraded. This degradation can be controlled temporally and spatially and is often a result from a physiological change or external stimulus such as a stress signal.

This regulated degradation is important not only for housekeeping functions such as providing new building blocks for synthesis and avoiding accumulation of abnormal proteins but also in reshaping the global proteomic composition of the cell. It is therefore important to understand the various mechanisms that the cell uses to select substrates for degradation as well as the requirements for the proteolysis to be accomplished.

This remainder of this introduction provides an overview of the principal players in the pathway for energy-dependent proteolysis. An emphasis is placed on control mechanisms in order to highlight the strategies the cell utilizes to regulate degradation at multiple steps. In-depth coverage is provided for the protease ClpAP and its accessory protein ClpS, which are central to this thesis. ClpAP and ClpS provide a system in which to study the mechanism underlying the complex role of proteases in the cell.

KEY COMPONENTS OF PROTEIN DEGRADATION

When regulated protein degradation is discussed, it is often in reference to the complete degradation of a protein substrate rather than a single cutting event or the cleavage of peptides, such as by amino- or carboxypeptidases. Our focus is on complete protein processing through processive degradation carried out by energy-dependent proteolytic machines. There are two principal components for protein degradation: the substrate to be eliminated and the complex that carries out the degradation. Tertiary factors, such as adaptor proteins, often provide another layer of regulation to substrate selection and processing.

Degradation Tags

Proteins that are slated for degradation must bear a specific signal that is distinct from stable proteins. In bacteria, proteins recognized for degradation possess tags encoded in their primary structure on their N- and/or C-termini. In eukaryotes, proteins intended for degradation are appended with a small modifier protein called ubiquitin. These tags or modifiers are recognized and bound either directly by the protease or by adaptor proteins which recruit the tagged substrate to the degradation machine. Here, I will address several types of tags, with a broader focus on bacterial systems.

ssrA tags

One type of degradation tag is the ssrA tag. This tagging system, encoded by the ssrA RNA, is involved in protein quality control in all eubacteria (Karzai et al. 2000). In the instance when an mRNA being translated lacks a proper stop codon, the ribosome stalls due to the absence of sufficient signals for release of the mRNA and the growing polypeptide chain. The ssrA RNA can rescue the stalled ribosome through the action of its tmRNA features. The tmRNA enters the A-site of the ribosome and directs the addition of an alanine to the existing polypeptide through the transpeptidation reaction. Subsequently, the tmRNA moves into the P-site of the ribosome and displaces the original mRNA message. The tmRNA then guides the translation of the remainder of the ssrA protein tag onto the C-terminus of the nascent polypeptide until a stop codon is reached in the new RNA message. The ribosome is now free to translate another mRNA, and the protein product is slated for degradation via the freshly-encoded ssrA tag (Keiler et al. 1996). It is estimated that 0.5% of translations in *Escherichia coli* result in ssrA-tagging, and proteins with the ssrA tag are degraded by intracellular ATP-dependent proteases (Gottesman et al. 1998; Lies and Maurizi 2008). The role of the ssrA system in bacteria appears to be largely for quality control, as it prevents build-up of aberrant protein products and keeps the cellular translation machinery running smoothly.

The sequence of the *E. coli* ssrA tag is AANDENYALAA-CO_2^-. The various proteases and adaptor proteins known to interact with the ssrA tag recognize different features and regions of the tag sequence, and proteins that recognize overlapping sequences on the ssrA tag compete for binding. The *E. coli* ATP-dependent protease ClpXP primarily recognizes the final LAA and the C-terminal carboxyl group of the ssrA tag (Flynn et al. 2003). Interestingly, other proteins that are known to be turned over in this bacterium have a similar C-terminal di-alanine sequence to the ssrA tag, such as RecN, a protein involved in DNA double-stranded break repair that is rapidly turned over in the cell when not in use (Neher et al. 2006). The di-alanine C-termini of these tags is a simple degradation motif recognized by intracellular proteases in bacteria, though other residues in the tag and in the overall protein may contribute to enhanced degradation and/or regulation.

Exposed substrate-specific tags

Like RecN, there are proteins in the cell that have a degradation tag encoded into its native sequence. One example, also taken from bacteria, is Dps (DNA-binding protein in starved cells), which can form biocrystals with DNA, protecting it from stresses and damage (Nair and Finkel 2004). Dps has a degradation tag encoded at its N-terminus and is degraded by ClpXP; Dps protein deleted in its first five residues is no longer degraded by ClpXP, and appending the first twelve residues onto a model protein causes that protein to be proteolyzed (Flynn et al. 2003). Low levels of Dps exist during exponential phase of *E. coli* growth, presumably because degradation is occurring soon after the protein is translated. However, Dps levels rise in stationary phase. The mechanism for how Dps is stabilized in stationary phase is unknown, but the DNA-binding and protection functions of the protein necessitate abundant Dps levels in the cell (Almiron et al. 1992). Thus, by constantly producing Dps, the cell is able to rapidly respond to stress and starvation conditions by halting degradation, leading to immediate accumulation of the protein. This scenario exemplifies the use of proteins with degradation signals "hard-wired" into their native sequence.

Internal and revealed degradation tags

Degradation tags can also reside in the interior of a protein's primary sequence. Some studies suggest that ATP-dependent proteases can directly recognize interior degradation sequences in model proteins (Hoskins et al. 2002). However, some proteins bearing an interior degradation tag are internally cleaved, revealing a tag on the new N- or C-terminus. The example of the RseA protein in *E. coli* illustrates this process well. RseA is an inner membrane-spanning protein. The N-terminal portion of the portion extends into the cytoplasmic space of the cell and binds to sigmaE, the sigma factor involved in turning on genes involved in addressing extracytoplasmic stress. RseA acts as an inhibitor of sigmaE activity because it sequesters sigmaE away from its respective promoters on the DNA. When the cell experiences periplasmic stress by way of an accumulation of unfolded or misfolded outer membrane proteins, two site-specific cleavage events occur in the RseA protein, one of which liberates the N-terminal portion bound to sigmaE. This cleavage reveals a degradation tag on the C-terminus of this N-

terminal fragment, denoted NRseA, and this tag bears a C-terminal –VAA motif that is similar to the ssrA tag in *E. coli*. Specific ATP-dependent proteases in the cytoplasm now recognize NRseA as a proteolytic substrate, releasing sigmaE to activate stress response genes (Flynn et al. 2004). Therefore, the specific revelation of a degradation tag is another strategy employed by the cell in regulated protein degradation.

Ubiquitin

In eukaryotes, the major degradation tag directly recognized for proteolysis is ubiquitin. Instead of being encoded into the sequence of the protein being degraded, ubiquitin is a small protein that is post-translationally added to substrates for degradation by the proteasome (discussed below). Multiple proteins are responsible for fusing the ubiquitin protein tag onto its substrate protein. The first enzyme, E1, is the ATP-dependent activating enzyme that renders the C-terminal amino acid of ubiquitin more reactive for subsequent steps through covalent binding to a cysteine residue in E1. The ubiquitin protein is next transferred to E2, a ubiquitin-carrier protein. Lastly, a ubiquitin ligase, E3, fuses the C-terminus of ubiquitin to a lysine residue on the substrate protein. The E2 and E3 steps occur in concert, and due to the need for substrate specificity in ubiquitin ligation and the subsequent degradation, there are many different types of E2 and E3 enzymes in the cell dedicated towards specific substrates. In some cases, adaptor proteins may assist E3 in the transfer of ubiquitin to substrates. Typically, multiple ubiquitins in the form of a polyubiquitin chain are required for recognition of a substrate by the proteasome, in which the C-terminus of a monoubiquitin unit is covalently attached to a lysine residue of the preceding ubiquitin in the chain. The proteasome directly recognizes the ubiquitin degradation tag, but the ubiquitin molecules are not degraded with the substrate protein (Hershko and Ciechanover 1998).

Because the ubiquitin molecule itself is not degraded, it can sometimes be considered an adaptor (discussed below) that delivers the actual substrates for degradation closer to the proteolytic machine. Evidence suggests that once a protein is brought into close proximity to the proteasome, an unstructured and accessible region of the substrate is required for efficient initiation of unfolding and degradation (Prakash et

al. 2004; Inobe and Matouschek 2008). In this light, the ubiquitin is simply the adaptor bringing in the unstructured portion towards the proteasome processing sites.

Ubiquitin-dependent sequences

Further complexity arises when one considers how the ubiquitination enzymes determine which proteins are slated for degradation. Even though the polyubiquitin chain is the principle degradation tag recognized by the proteasome, certain sequence elements in the substrate proteins themselves are necessary for degradation. One example is the recognition of PEST elements, regions in a substrate protein enriched in Pro, Glu, Ser, and/or Thr residues. PEST elements possess phosphorylation sites that are in turn recognized by the ubiquitination enzymes. The timing of some cyclins is regulated by key phosphorylation events at PEST elements (Hershko and Ciechanover 1998). Thus, ubiquitin serves as a very important and versatile degradation tag for a wide variety of proteolytic substrates in eukaryotic cells.

N-degrons

Another broad category of degradation tags that exists in eukaryotes as well as prokaryotes is the N-degron. It was first reported in 1986 by the lab of Alexander Varshavsky that systematically changing the amino-terminal residue of a protein altered the protein's half-life in the cell (Bachmair et al. 1986). The N-end rule was developed, stating that the half-life of a protein is related to the identity of its N-terminal residue, with proteins commencing with large, bulky residues favored for degradation. However, the N-end rule, a phenomenon that exists in all organisms studied, has different sequence specificities in different species (Fig. 1.1). For example, Leu, Phe, Trp, and Tyr are primary destabilizing residues in *E. coli*. Arg and Lys are considered secondary destabilizing residues because a Leu/Phe aminoacyl-transferase is required to append an additional N-terminal Leu to bear a primary destabilizing residue (Varshavsky 1995). In *E. coli*, most N-degrons can be directly recognized by the protease ClpAP, but is assisted in recognition by the adaptor protein ClpS (Erbse et al. 2006). In contrast, in the yeast *Saccharomyces cerevisiae*, the N-end rule is more complex. More residues are considered primary destabilizing residues, and there are both secondary and tertiary

destabilizing residues that require one and two more processing steps, respectively, to become functional recognition signals. Furthermore, unlike *E. coli*, in *S. cerevisiae*, the N-degron is not directly recognized by the protease. Instead the N-degron is recognized by N-recognin – the general class of E3 ubiquitin ligases that recognize N-degrons, and the N-degron-containing substrate is processed along the ubiquitination pathway as described above (Varshavsky 1996; Hershko and Ciechanover 1998).

In *S. cerevisiae*, Varshavsky and colleagues discovered that a subunit of cohesin, the molecule that holds together sister chromatids during DNA replication, is an N-end rule substrate. Before anaphase, the protease separase cleaves the subunit of cohesin, revealing an N-degron on the C-terminal fragment. This fragment is degraded through the ubiquitin-proteasome pathway, and the destruction of the cohesin complex allows the separation of sister chromatids to their respective daughter cells during cell division (Rao et al. 2001). In contrast to eukaryotic cells, no physiological N-end-rule substrates have been fully characterized in bacteria. Because of the existence and conservation of N-degron processing machinery in bacteria, N-degron substrates are likely to exist. One theory for the conservation is that the degradation of N-end rule substrates in bacteria is a quality control mechanism. In *E. coli*, for instance, the initiator methionine is cleaved from the amino-terminus of the newly translated protein only if the second residue is small (Hirel et al. 1989). Thus, the primary and secondary destabilizing residues should never be revealed at the N-terminus of a protein after translation. If one of these destabilizing residues is present at the N-terminus, that protein may be a damaged or aberrant protein that should be cleared from the cell. Nevertheless, it is also possible that the N-end rule is conserved in bacteria for specific regulatory events involving post-translational cleavage of a protein and uncovering of an N-degron, as exemplified in eukaryotes with cohesion. Due to the lack of absolute sequence stringency other than the character of the N-terminal residue, N-degrons are potentially used in a wide range of regulatory and quality-control mechanisms in the cell. A very recent study suggests that Dps and Putrescine-Aminotransferase (PATase) are processed by the N-end-rule pathway (Schmidt et al. 2009), and it will be interesting to follow-up on these substrates.

I have presented several classifications of degradation tags as well as examples illustrating their uses. Understanding how substrate proteins are recognized for

degradation aids in the discovery of new substrates as well as the mechanisms underlying the proteolysis.

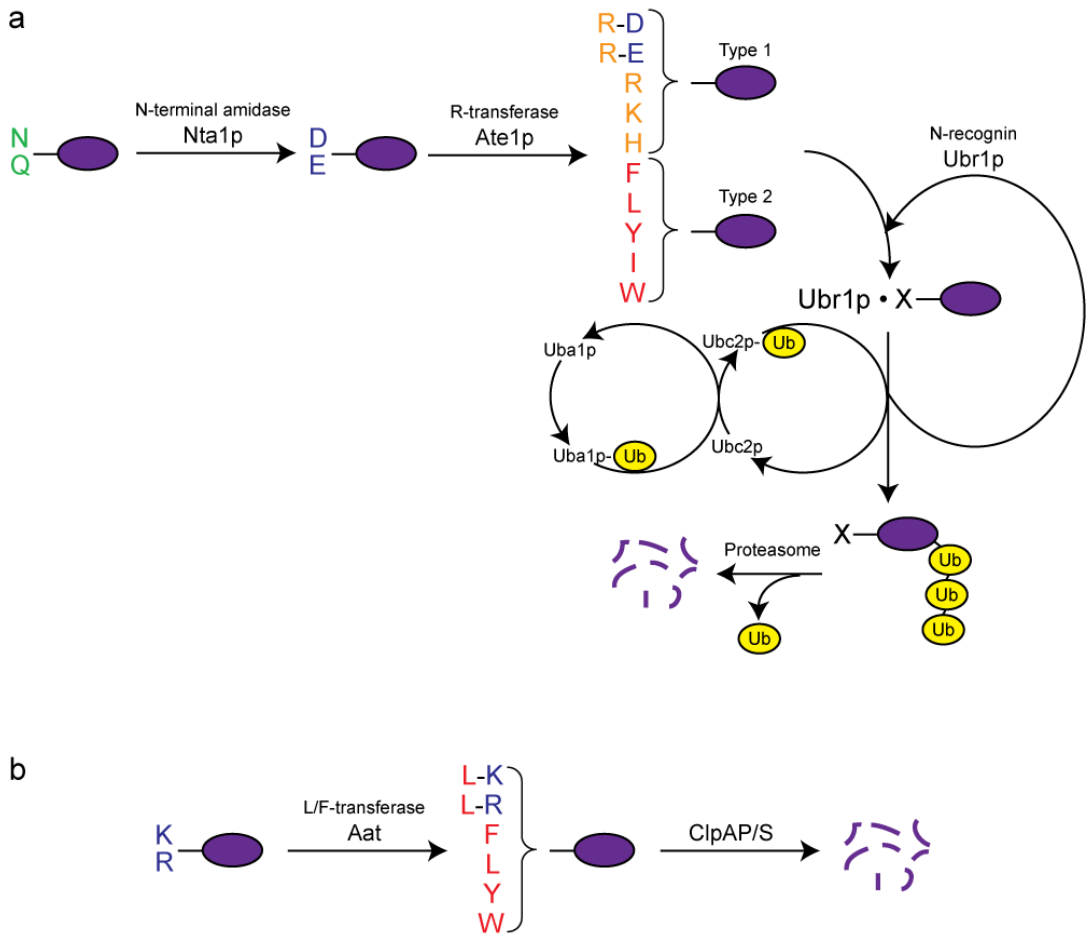


Figure 1.1. The N-end rule pathway.
(a) The ubiquitin-pathway in *Saccharomyces cerevisiae* is shown, with tertiary residues in green, secondary residues in blue, and primary destabilizing residues in orange (Type 1) and red (Type 2). (b) The pathway in *Escherichia coli* with secondary destabilizing residues in blue and primary destabilizing residues in red.
(Adapted from Varshavsky 1996)

Proteolytic machines

A study examining proteolysis in yeast showed that energy was required for the breakdown of intracellular proteins (Halvorson 1958a; Halvorson 1958b). The field now knows that ATP-dependent proteolytic machines are responsible for this degradation.

The archetypical ATP-dependent protease is a barrel-shaped complex with proteolytic active sites hidden in the inside chamber. Substrates are recognized by ring-shaped regulatory components capping the ends of the barrel, unfolded through a pulling motion generated by ATP hydrolysis-induced conformational changes within the machine, and threaded through the pore of the barrel to the active sites where the substrate is degraded. The resultant peptides are then released into the milieu in a mechanism not thoroughly understood.

Basic architecture

The core component of the proteolytic machine is the peptidase barrel. In general, the proteolytic core's barrel is comprised of stacked rings of identical or similar subunits. The interior chamber contains the enzymatic active sites, and structural protrusions at the ends of the pore act as "gates" to the opening of the inner cavity, preventing entire proteins from diffusing inside (Bajorek and Glickman 2004; Jennings et al. 2008). Thus, additional regulation is required to select which substrates are eligible for degradation.

The protein-unfolding components that cap the ends of the core proteolytic components bear the ability to discriminate between substrates and to open the central pore allowing processing of larger polypeptides such as whole proteins. These subunits belong to the AAA+ (ATPases associated with various cellular activities) family of proteins. AAA+ proteins are characterized by ATP-binding motifs that have been conserved throughout evolution. Proteins bearing these motifs have a AAA+ domain, and these domains typically multimerize to form hexameric rings. Most rings of AAA+ proteins are able to bind and hydrolyze ATP, and the resultant conformational changes in the overall protein can exert mechanical force on other molecules. Therefore AAA+ proteins are involved in a wide array of cellular functions requiring movement or the application of force (Fig. 1.2). For example, AAA+ proteins are involved in cytoskeletal mobility, DNA replication, vesicle fusion, and a wide array of ubiquitin-mediated functions (Hanson and Whiteheart 2005). Because AAA+ proteins participate in a multitude of functions, understanding the mechanisms behind regulated proteolysis by AAA+ proteases aids in the knowledge of this general class of proteins.

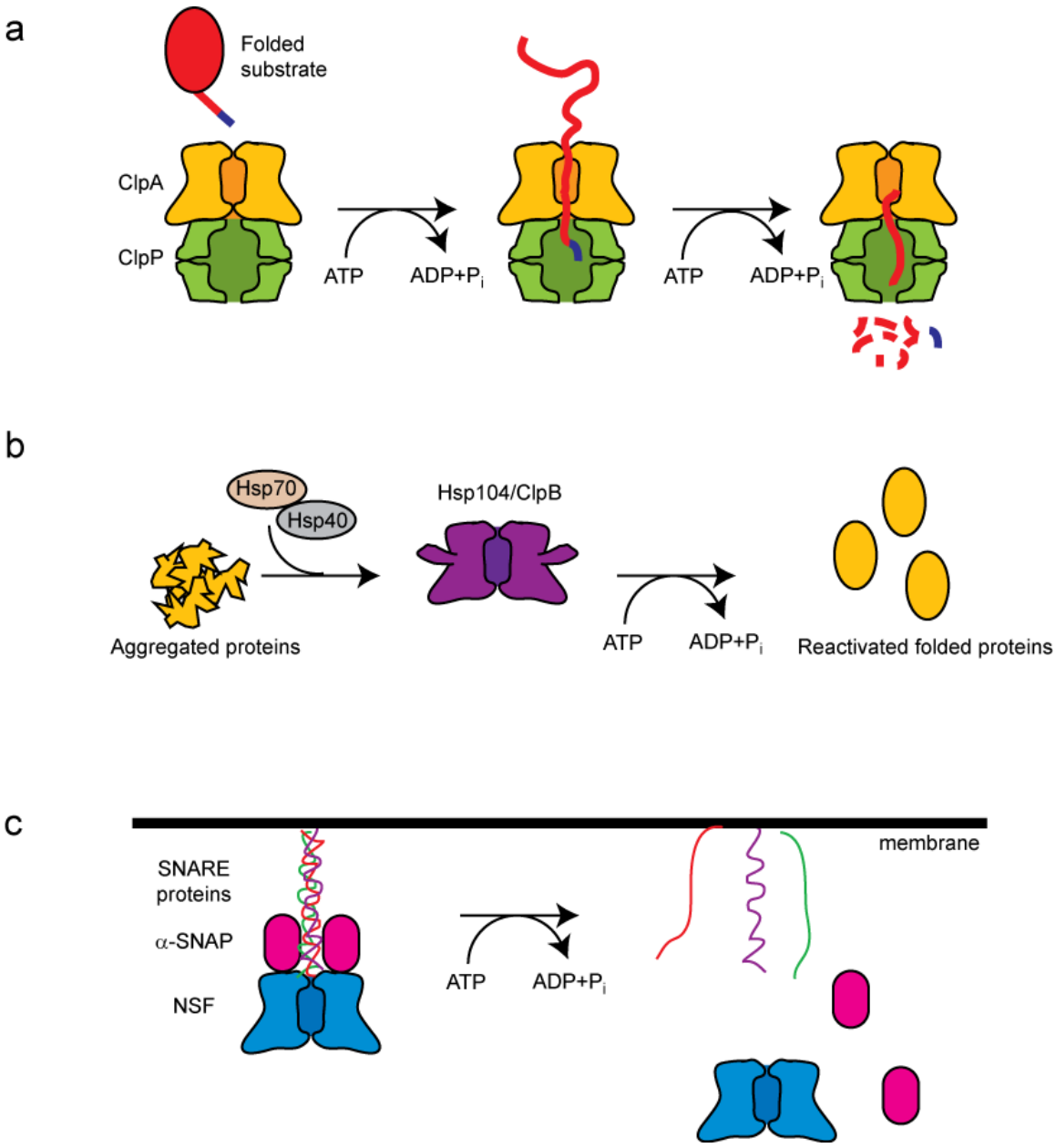


Figure 1.2. Roles of AAA+ proteins.
(a) In bacteria, the AAA+ enzyme ClpA in conjunction with the peptidase ClpP can unfold and degrade tagged protein substrates. (b) The AAA+ unfoldase Hsp104/ClpB, with the assistance of co-chaperones Hsp70 and Hsp40, can unfold and resolubilize aggregated proteins. (c) The AAA+ enzyme N-ethylmaleimide-sensitive factor (NSF) protein and its adaptor protein α-SNAP can disassemble SNARE proteins, allowing vesicle fusion in eukaryotes to proceed.
(Hanson and Whiteheart 2005)

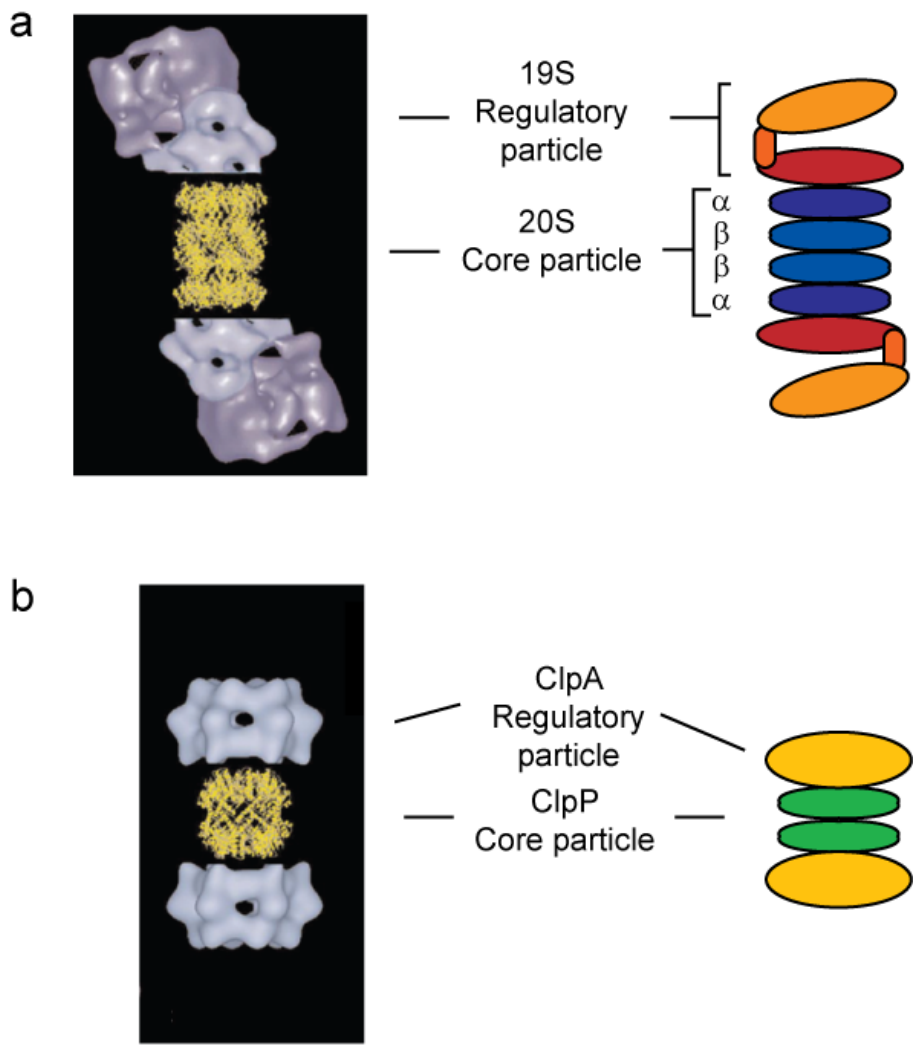


Figure 1.3. The 26S proteasome in eukaryotes and ClpAP in bacteria.
(a) Left, combined electromicrogram and crystal structure of the 19S regulatory particle and the 20S core particle. Right, cartoon depiction of the major structural components of the proteasome. (b) Left, combined electromicrogram and crystal structure of ClpAP. Right, cartoon depiction of the protease.
(Zwickl et al. 2000; Schmidt et al. 2005)

The proteasome

The proteasome in archaea and eukaryotes and the various AAA+-proteases in bacteria bear much resemblance. For simplicity, only the 26S eukaryotic proteasome will be described and used for comparison. The 26S proteasome is comprised of a proteolytic core called the 20S subunit and a 19S regulatory subunit capping each end of the core. The 20S proteasome is composed of four seven-membered rings of alpha and beta

subunits (Fig. 1.3a). In the eukaryotic 20S core particle, only three of the seven beta subunits in each ring have proteolytic activity; the alpha subunits lack protease active sites. The proteasome typifies the canonical protease with its proteolytic active sites sequestered inside the chamber created by the stacked rings of alpha and beta subunits. The 19S regulatory particle is required for the specific degradation of substrate proteins by the proteasome and contains two major regions, the lid and the base. The base is formed by six ATPase subunits and two non-ATPase subunits. The base of the 19S particle directly contacts the 20S core and is responsible for the ATP-dependent unfolding of substrate proteins to be fed into the 20S particle. The lid of the 19S particle is less well understood, but one component of the lid appears to be necessary for the deubiquitination of ubiquitinated substrates (Pickart and Cohen 2004).

Bacterial proteases

Bacterial AAA+-proteases are simpler than the 26S proteasome. There are several different types of AAA+ proteases in bacteria. *E. coli* has five distinct proteases with partially overlapping specificities: ClpAP, ClpXP, FtsH, HslUV, and Lon. Using ClpXP, a well-studied bacterial protease, as an example, a similar proteolytic core and capping unfoldase architecture is observed. ClpXP is comprised of the ClpP core peptidase flanked by one or two ClpX hexamer rings. ClpP is a tetradecamer in which two identical seven-membered rings bind back-to-back (Gottesman 2003). In the absence of an ATPase, ClpP is able to degrade peptides, preferentially cleaving after hydrophobic residues (Katayama-Fujimura et al. 1987; Woo et al. 1989). As with the 26S proteasome, partnering of ClpP with an ATPase such as ClpX allows the regulated degradation of protein substrates. ClpX is a homohexamer, in which each monomer has an N-terminal domain and a AAA+ domain. The AAA+ domain of ClpX forms a hexameric ring as is the case with most AAA+ family members, with the ATP-binding site at the interface between subunits.

The other AAA+ proteases in bacteria have a similar construction to ClpXP. ClpA is similar to ClpX in that it associates with ClpP, but ClpA has two AAA+ domains per monomer instead of one (Fig. 1.3b). HslUV is another two-component protease in which an HslU hexamer is the AAA+ unfoldase and the HslV tetradecamer is the

peptidase. In contrast, Lon and FtsH are both one-component proteases, with the ATPase domain and the proteolytic domain encoded in the same polypeptide per monomer. FtsH is bound to the inner membrane and is the only AAA+ protease in *E. coli* that is essential (Gottesman 2003).

Studying bacterial proteases has shed light on general mechanisms behind AAA+ enzymes. One insight relates to the recognition of degradation tags. Because the AAA+ subunits dictate which proteins are transferred to the proteolytic subunit, there must be recognition elements of substrate degradation motifs on the ATPase. Several electron microscopy studies have shown that the regulatory subunits and the proteolytic subunits bind such that their central pores are aligned (Beuron et al. 1998; Ortega et al. 2000; Ishikawa et al. 2004). Due to the fact that the proteolytic active sites are sequestered inside the enclosed chamber of the proteolytic subunit, the clearest path to the protease active sites is through the central pore. The pore is lined with loops that protrude into the axial space (Wang 2001). Mutation of pore loop residues in the AAA+ subunits of these proteolytic machines impair the function of their respective protease, demonstrating the importance of these residues in the recognition and processing of substrate proteins (Yamada-Inagawa et al. 2003; Siddiqui et al. 2004; Hinnerwisch et al. 2005a; Park et al. 2005; Okuno et al. 2006; Farrell et al. 2007; Martin et al. 2008b). Cross-linking experiments show direct interactions between substrate tags and pore residues (Hinnerwisch et al. 2005a; Martin et al. 2008a). These pore-loop residues are not only important for substrate recognition, but are also likely involved in the ATP hydrolysis-dependent movements shuttling the substrate unidirectionally towards ClpP (Reid et al. 2001; Hinnerwisch et al. 2005a; Martin et al. 2008b).

Even though all ATP-dependent substrates must pass through the pore, some substrates require or are aided by interactions with other regions of the protease. The simplest way in which other regions of the protease can participate in substrate interactions is by "tethering" the substrate on the N-terminal domains of the ATPase subunit in proximity to the pore. For example, the ClpXP protease recognizes the protein-DNA transposition complex of phage Mu. This transpososome is comprised of a tetramer of the phage-encoded protein MuA bound to DNA in a hyper-stable structure that must be disassembled for transposition of phage DNA into host DNA to continue

(Levchenko et al. 1995). ClpX acts upon the Mu transpososome by removing one MuA monomer, thereby destabilizing the structure to allow subsequent steps of transposition (Burton and Baker 2003). Both ClpX and ClpXΔN are able to bind and process isolated MuA monomers. However, only full-length ClpX is able to unfold a MuA monomer in the context of the Mu transpososome. The current model states that there are conformational differences in the MuA protein when it is a monomer versus when it is a tetramer on DNA. This conformational difference could yield new contacts in the tetramer that interact with the N-domain of ClpX while tethering and positioning the core contacts to the pore for processing (Abdelhakim et al. 2008). Thus, the N-terminal domains of AAA+ proteases can also interact with substrates distinctly from pore contacts, adding further control and complexity to regulated proteolysis.

Conversely, other substrates can be processed by ClpXP with or without the N-terminal domain of ClpX. One example is ssrA-tagged substrates, which can be degraded by ClpXΔNP (Neher et al. 2003; Wojtyra et al. 2003). As mentioned earlier, the ssrA-tag is directly recognized in the pore of ClpX and therefore lacks any requirement for the N-terminal domain. Thus, the N-terminal domains of the ClpX hexamer and other AAA+ proteins can participate in additional regulatory roles but are not always essential.

Adaptors

The cell regulates protein degradation at many different points along the proteolytic pathway, such as by the revelation or creation of substrate tags, recognition of degradation motifs by the protease, and differential processing of substrates based on conformation and stability. The use of adaptor proteins provides yet another manner in which recognition is increasingly controlled and intricate. In general, adaptors function by binding to both the substrate and to the enzyme, increasing the local concentration of the substrate to the enzyme. The N-terminal domains of AAA+ proteases serve as docking sites for adaptors, exhibiting another use for these domains. In the cell, not only is there competition between substrates for degradation, but adaptors also compete for binding space on the N-terminal domains. Up-regulating an adaptor can prioritize its

partner substrate for degradation over other substrates and substrate-adaptor pairs. Hence, adaptor proteins can modulate the selection of substrates degraded by a protease.

SspB, a canonical adaptor

E. coli SspB is a well-characterized adaptor for the protease ClpXP (Fig. 1.4). SspB recognizes the ssrA tag and delivers ssrA-tagged substrates to ClpXP, increasing the rate of degradation (Levchenko et al. 2000). The SspB adaptor protein is a dimer, and each monomer is composed of a substrate binding domain, a flexible linker, and a short ClpX-binding (XB) motif at the C-terminus (Wah et al. 2002; Dougan et al. 2003; Levchenko et al. 2003; Wah et al. 2003). The XB motif binds to the ClpX N-terminal domain, and the flexible linker keeps the substrate domain on a short leash in close proximity to the pore of ClpX. Interestingly, ClpX recognizes the extreme C-terminal amino acids of the ssrA tag and SspB recognizes the N-terminal residues, such that the adaptor hands off the substrate to the protease rather than competing for binding to the same residues (Flynn et al. 2001). Due to its simple yet elegant method of delivery, SspB has become the paradigm of a tethering adaptor protein.

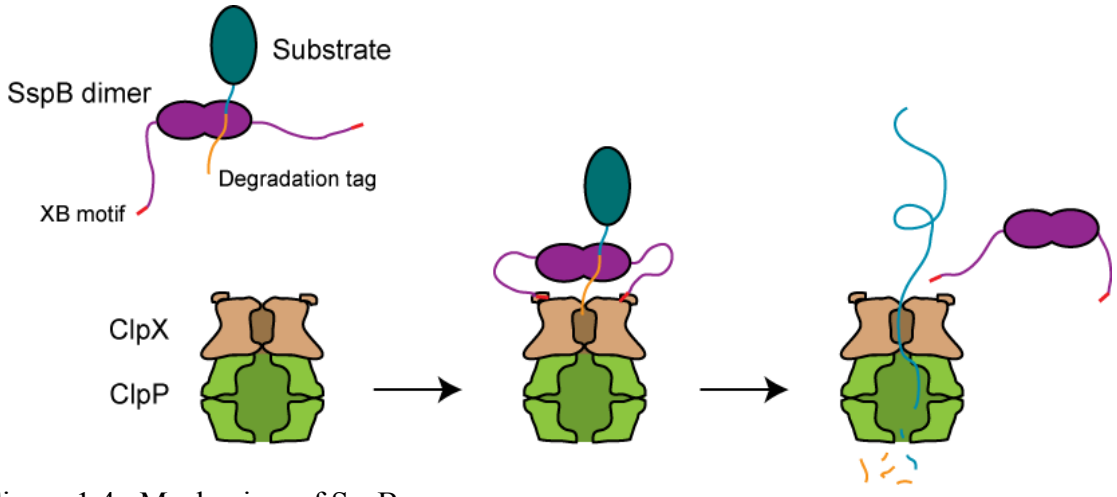


Figure 1.4. Mechanism of SspB.
The SspB dimer binds to the N-terminal section of the ssrA tag (blue) on the tagged substrate and delivers the substrate to the protease ClpXP. The XB motifs on SspB bind to the N-terminal domain of ClpX, tethering the ssrA-tagged substrate in close proximity to the enzyme pore. ClpX recognizes the C-terminal section of the ssrA tag and proceeds to degrade the substrate, releasing the SspB to recruit another substrate.

Other bacterial adaptors

Some substrates require an adaptor for degradation. For example, the stress sigma factor in *E. coli*, sigmaS, is only degraded with the assistance of the adaptor protein RssB. RssB function, in turn, can be controlled by phosphorylation and the presence of anti-adaptors, showing how regulation of the adaptor protein itself can be a component in the regulation of overall proteolysis (Klauck et al. 2001; Zhou et al. 2001; Bougdour et al. 2006; Bougdour et al. 2008). Similarly, the adaptor protein MecA is required for virtually all functions of the *Bacillus subtillis* AAA+ protease ClpCP (Schlothauer et al. 2003). It was discovered that the adaptor MecA is required for ClpC oligomerization (Kirstein et al. 2006). *B. subtillis* ClpC is structurally similar to ClpA in *E. coli*, but ClpA does not require an adaptor to form an oligomer. Hence, adaptors can play essential roles in the degradation of particular substrates rather than simply enhancing the basal degradation by the protease.

THE PROTEASE CLPAP AND THE ADAPTOR CLPS

The protease ClpAP in *E. coli* is a AAA+ protease that still holds many mysteries. ClpAP has several similarities to ClpXP in *E. coli* in that both proteases have identical ClpP proteolytic units, are composed of hexameric AAA+ domains, and degrade some shared substrates. However, the similarities end there, and not all insights into ClpXP structure and function can be translated to ClpAP. There are numerous outstanding questions regarding the physiological significance of ClpAP as well as the detailed mechanisms behind substrate recognition and processing. Furthermore, the field is learning that the mechanism of the adaptor protein ClpS for ClpA is dissimilar from the paradigm of adaptor function that was largely worked out based on the mechanistic studies with the ClpX adaptor SspB.

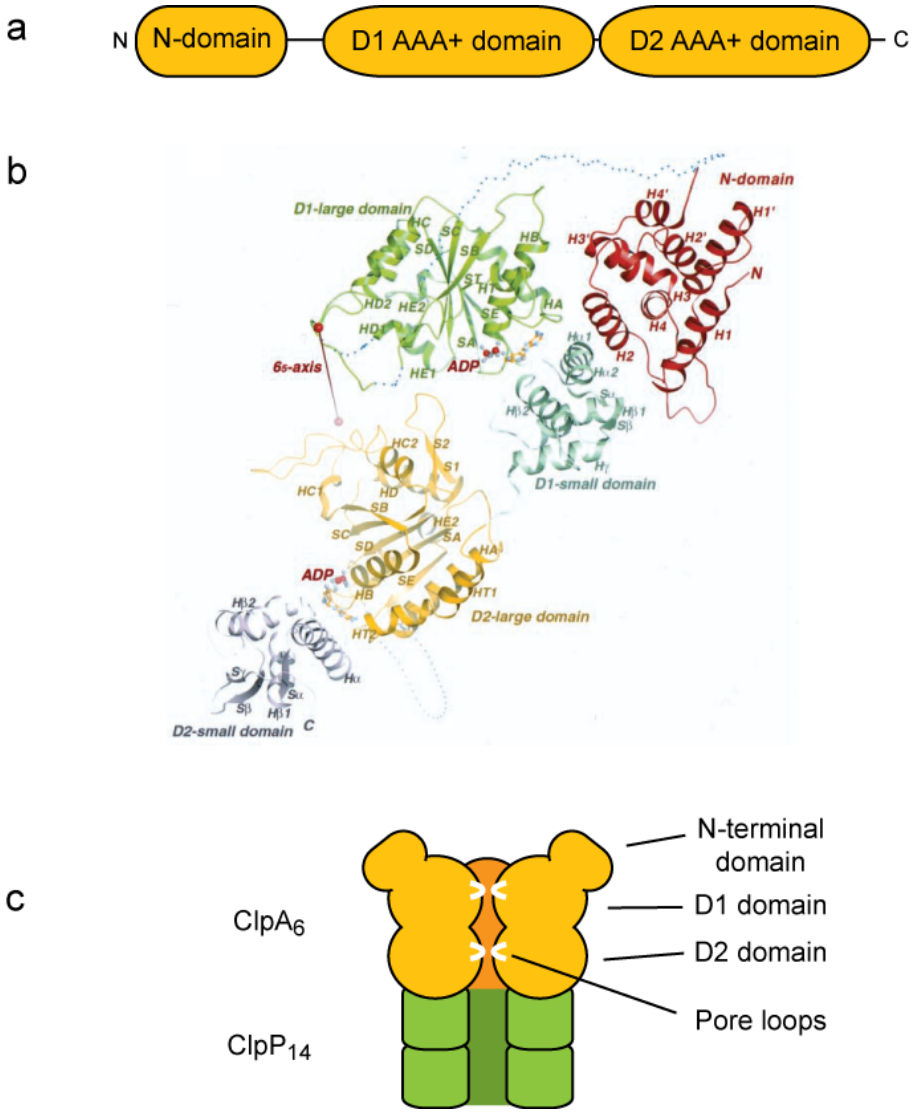


Figure 1.5. The AAA+ enzyme ClpA.
(a) Domain layout of ClpA in *E. coli*. (b) Crystal structure of a ClpA monomer, with unresolved portions depicted as a dotted line (Guo et al. 2002b). (c) Cutaway representation of the ClpA hexamer bound to the ClpP tetradecamer. D1 and D2 pore loops are depicted in white.

The ATP-dependent unfoldase ClpA

ClpAP was discovered in *E. coli* in 1987 as a protein complex that has ATPase and proteolytic capabilities. The discovery that the purified complex could degrade casein led to the name Clp (caseinolytic protease, and later, chaperone-linked protease) (Hwang et al. 1987; Katayama-Fujimura et al. 1987). The milk protein, casein, does not

exist in bacteria, but casein is still used today as a model substrate for purified ClpAP. Alpha-casein is a protein without a well-defined secondary or tertiary structure (Susi et al. 1967), and ClpAP has since been characterized to recognize unfolded proteins.

The physiological role of ClpAP is still undefined. In *E. coli*, the *clpA* and *clpP* genes are both non-essential. *clpA* knockouts by themselves do not exhibit any reported growth phenotype (Katayama et al. 1988). *E. coli* with the *clpP* gene deleted do not have a strong growth defect but are slightly less able to survive in starvation conditions compared to wild-type cells (Damerau and St John 1993; Weichart et al. 2003). It is important to note, however, that ClpP is also a component of the protease ClpXP in *E. coli*, thus deletion of *clpP* produces pleiotropic effects due to its roles in other proteolytic machines. Furthermore, the *clpA* gene is not upregulated by heat shock, unlike other some other genes for protein unfoldases (Katayama et al. 1988).

There are very few known physiological substrates for ClpAP. One known substrate is MazE, part of the MazEF toxin-antitoxin system. MazF is toxic towards cells but is neutralized when partnered with MazE. The *mazEF* cassette is constantly being expressed, but upon encountering stress, *mazEF* transcription is diminished. The labile MazE is degraded by ClpAP, leaving MazF to initiate programmed cell death (Aizenman et al. 1996; Engelberg-Kulka et al. 2005). Under stress conditions, it is sometimes advantageous to kill some cells of the population in order to preserve others, and the degradation of MazE by ClpAP appears to play a role in this rapid regulation.

ClpAP is able to degrade ssrA proteins, and ssrA-tagged proteins have been extensively used as a model substrate (Gottesman et al. 1998). However, studies have shown that although ClpAP independently has the ability to degrade ssrA-tagged proteins, ClpAP is not a large contributor to the degradation of these tagged proteins *in vivo*. In *E. coli*, ClpXP is the primary protease responsible for degrading ssrA-tagged proteins, and only if the *clpX* gene is disrupted does ClpAP show an effect on decreasing ssrA levels in the cell (Farrell et al. 2005; Lies and Maurizi 2008). The adaptor ClpS inhibits the degradation of ssrA-tagged proteins by ClpAP *in vitro* (discussed below), but even disruption of the *clpS* gene does not greatly affect ClpAP's ability to assist in ssrA-tagged protein degradation *in vivo* when ClpX is present (Farrell et al. 2005; Lies and

Maurizi 2008). It is possible that the cell has created a redundancy for quality control mechanisms such as by having multiple proteases clear ssrA-tagged proteins.

Another likely physiological role of ClpAP is the active degradation of substrates bearing an N-degron. Earlier studies have shown ClpAP to be the principle protease involved in degrading this class of substrates. However, these studies have utilized model N-end rule substrates, and so far, there are no known physiological N-end-rule substrates in bacteria. It is not clear whether the N-end-rule in bacteria is for specific regulatory pathways or more of a quality control mechanism, ensuring that the cell eliminates proteins that might be damaged or mis-processed, alerted by sensing unnaturally with large, bulky residues at N-termini.

The gene for *clpA* exists virtually only in bacteria, unlike *clpX* and *clpP* which are encoded in human mitochondrial DNA. ClpA is not even prevalent in all bacteria; some bacteria such as *Bacillus subtilis* have similar Clp proteins such as ClpC. It is unclear whether ClpA and ClpC should be considered orthologs, and it is possible that these two proteins have evolved to play different roles in different organisms.

ClpAP possesses a similar structure to the proteasome and to the other bacterial AAA+ proteases. ClpA is a nucleotide-dependent hexamer, and each ClpA monomer is comprised of an N-terminal domain and two ATPase domains, called D1 and D2 (Fig. 1.5). The D2 domain is most similar to the single AAA+ domain of ClpX, and it has been suggested that in the hexamer, the D2 domain plays the active role in ATP hydrolysis (Singh and Maurizi 1994; Seol et al. 1995). The same studies purport that the D1 domain bound to nucleotide is the main domain responsible for oligomerization. In the absence of ClpA, ClpP can only proteolyze peptides (Katayama-Fujimura et al. 1987; Woo et al. 1989). Like other AAA+ proteases, the addition of the ATPase chaperone is required to process entire proteins. The ssrA-tag attached to a substrate protein has been shown to interact with pore residues in the D2 domain of ClpA through cross-linking (Hinnerwisch et al. 2005a). The cross-linking freezes an interaction of the substrate-enzyme complex, but it does not rule out other transient pore interactions that may occur closer to the ClpA pore opening by the D1 domain. In fact, mutation of pore loop residues in the D1 domain also disrupt binding and/or processing of substrates by ClpAP (Hinnerwisch et al. 2005a).

The ClpA N-terminal domains do not play a role in hexamerization, as ClpA lacking its N-terminal domain (ClpAΔN) can hexamerize and process substrates (Lo et al. 2001; Singh et al. 2001; Hinnerwisch et al. 2005b; Cranz-Mileva et al. 2008; Maglica et al. 2008). Nevertheless, one study suggests that ClpAΔN has lower ATPase activity and a weakened affinity to the ClpP tetradecamer (Hinnerwisch et al. 2005b). The ClpA N-terminal domain also might contain peptide binding sites that assist in substrate recognition and/or processing (Xia et al. 2004; Erbse et al. 2008). Accordingly, the N-terminal domains most likely play important but incompletely characterized roles in ClpAP mechanism in addition to adaptor-binding.

The Adaptor ClpS

To date, only one known adaptor has been found to bind to ClpA, the ClpS protein. ClpS is a small, monomeric protein that binds to the N-terminal domains of ClpA. The *clpS* gene was discovered as a short open reading frame immediately upstream of the *clpA* gene in the same operon (Dougan et al. 2002). It is unclear how the transcription and translation of both of these genes are regulated. A recent study showed that there are around 250-300 molecules of ClpS per *E. coli* cell. ClpA levels, on the other hand, varied by growth stage. In exponential phase, there are approximately 40-50 ClpA hexamers. ClpA levels rise in stationary phase to about 150 hexamers per cell (Farrell et al. 2005). The regulatory mechanism controlling the levels of both proteins is not understood but altering the relative stoichiometries of these proteins could provide an additional layer of regulation in protease activity (see Chapter Two).

ClpS has been crystallized in conjunction with the isolated N-terminal domain of ClpA (Guo et al. 2002a; Zeth et al. 2002). The majority of the N-terminal 24 residues of *E. coli* ClpS is not visible in crystal structures and is presumed to be quite flexible. The remainder of the ClpS protein is folded into a single domain, and the interface with the ClpA N-terminal domain is through the folded domain of the adaptor (Fig. 1.6a). The core folded domain of ClpS is well conserved throughout bacteria; however, the N-terminal extension is not conserved in sequence or in length. The residues in the core domain of ClpS responsible for binding to the ClpA N-terminal domain are relatively

invariant (Zeth et al. 2002), implying that the ClpS-ClpA interaction is also well conserved.

ClpS is able to inhibit ClpA's ability to degrade several types of substrates. ClpA in the presence of ClpP has the ability to degrade itself, and ClpS inhibits this autodegradation (Dougan et al. 2002). It was previously suggested that there was a ClpA degradation tag in the N-terminus of ClpA itself and that ClpS simply masked this signal (Gottesman et al. 1990; Dougan et al. 2002), but recent evidence advocates that the degradation tag is at the C-terminus of ClpA and that ClpS binding results in a substrate specificity switch (Maglica et al. 2008). It is unclear whether ClpA autodegradation is a phenomenon that is relevant *in vivo*, but it is possible that ClpS can affect the levels of ClpA in the cell. Furthermore, ClpS also blocks the degradation of ssrA-tagged proteins (Dougan et al. 2002). As mentioned above, ClpAP is not a major contributor of ssrA-tagged protein degradation in the cell, but ClpS is unlikely to be the only reason for the lack of ClpA participation (Farrell et al. 2005; Lies and Maurizi 2008). Lastly, ClpS appears to inhibit ClpAP's ability to degrade unfolded proteins, such as the model *in vitro* substrate casein (Guo et al. 2002a). Perhaps unfolded proteins, ssrA-tagged proteins, and ClpA itself are recognized for degradation by ClpAP in a similar manner and are consequently inhibited by ClpS binding by a unified mechanism.

ClpS also plays a positive role in stimulating the delivery of N-end-rule substrates to ClpAP for degradation. Conserved residues on the folded domain of ClpS recognize the N-degron, and adaptor binding to ClpAP recruits bound substrates to the protease (Erbse et al. 2006; Wang et al. 2008b). A bioinformatics study established that ClpS is homologous to a domain of the E3 ubiquitin ligase N-recognin (Lupas and Koretke 2003), and together with the knowledge that ClpAP is the protease responsible for N-end-rule degradation, it was logical that ClpS played an important role in this process.

ClpAP alone has the ability to recognize and degrade N-degrons slowly, but the addition of ClpS greatly increases the rate of this degradation (Wang et al. 2007). There are several features that generate a good N-end-rule substrate. The first feature is having an N-terminal destabilizing residue combined with a free α-amino group at the N-terminus (Wang et al. 2008a). A recent crystal structure indicates that the ClpS binding pocket for N-degrons makes molecular contacts with the N-terminal α-amino group as

well as the side chain of the first residue (Fig. 1.6b) (Wang et al. 2008b). The second feature of a good N-degron is the lack of negatively-charged side chains near the amino terminus. Moving a negatively charged residue closer to the N-terminal position decreases the observed rate of substrate degradation (Wang et al. 2008b). Thirdly, to be degraded efficiently, the N-degron must not be too close to the folded region of the substrate protein (Erbse et al. 2006; Wang et al. 2008a). For example, the four-residue N-degron YAAA (YA_3) appended onto the N-terminus of GFP can be recognized by ClpS but cannot be degraded by ClpAP. Similarly, the stability of the substrate at its N-terminus and its propensity to being unfolded by ClpAP determines whether or not it is a good N-end-rule substrate (Wang et al. 2008a).

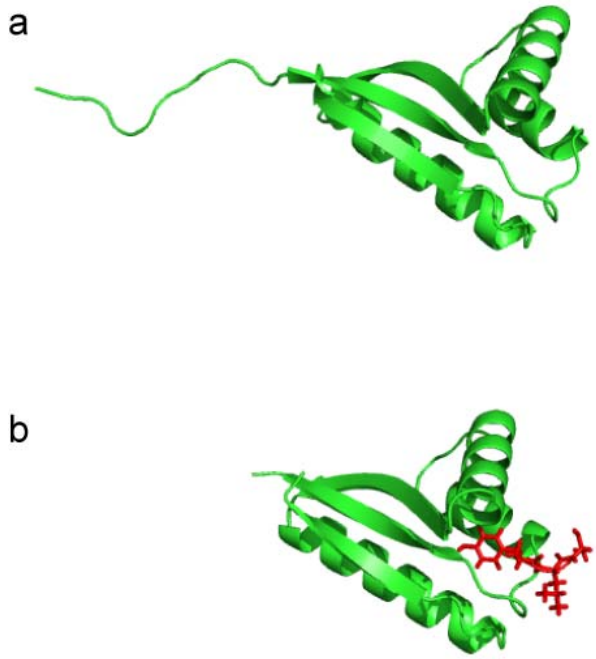


Figure 1.6. Structures of ClpS.
(a) Structure of *E. coli* ClpS. The unstructured region depicted is the N-terminal extension commencing at residue 16. Residues prior to position 16 are completely disordered in this structure (PDB 1LZW). (b) Structure of *C. crescentus* ClpS bound with an N-degron-containing peptide. A truncated form of ClpS from residues 31-119 was used for crystallization, and only residues 35-119 are visible in the resolved structure. Only the first two residues of the N-end-rule peptide (red) are visible (PDB 3DNJ).

Though both ClpAP and ClpS can directly recognize the N-degron, the requirements for binding and processing are not identical. As mentioned above, YA_3-GFP can be bound to ClpS but cannot be recognized by ClpAP. ClpAP seems to require a slightly longer linker length between the N-degron and the folded substrate (Erbse et al. 2006; Wang et al. 2008a). Additionally, ClpAP can recognize the N-terminal motif ILFVQEL fused to GFP, which is not a bacterial N-degron per se. ClpS does not bind a peptide containing the ILFVQEL motif (Wang et al. 2008a; Wang et al. 2008b). Because both ClpS and ClpAP must recognize ClpS-mediated N-end-rule substrates, the differential substrate requirements are intriguing.

THESIS OVERVIEW

My work aims to further characterize the roles and mechanism of the adaptor protein ClpS. Chapter Two describes an initial characterization of ClpS as an adaptor protein, showing that ClpS breaks the mold of how the field previously thought adaptor proteins operated. Chapter Three details further dissection of the ClpS protein, focusing on the N-terminal extension of the adaptor and how it affects the ability to delivery N-end-rule substrates. This work aims to find the minimal required features for a functional adaptor. Chapter Four describes a proteome-wide approach for gaining insight into the *in vivo* role of ClpS and how it alters ClpAP substrate selectivity.

Determining how ClpS molecularly shifts ClpAP substrate selection allows a deeper understanding of the strategies used by the cell in shaping the proteome. Furthermore, understanding how ClpS interacts with ClpA and how it affects the mechanics of substrate processing leads to a broadened appreciation of how adaptor proteins in general influence their AAA+ enzyme. Thus, ClpS and ClpAP expand our knowledge of how AAA+ proteins are regulated has implications beyond bacteria and proteolysis.

CHAPTER TWO

Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP

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

J.Y.H. carried out the experiments. J.Y.H., R.T.S. and T.A.B. contributed to experimental design and wrote the manuscript.

ABSTRACT

Adaptor proteins modify substrate recognition by AAA+ ATPases. We examined how the adaptor ClpS regulates substrate choice by the *Escherichia coli* protease ClpAP. Binding of six ClpS molecules to a ClpA hexamer enhanced N-end-rule substrate degradation and inhibited ssrA-tagged protein proteolysis. Substoichiometric ClpS binding allowed intermediate degradation of both substrate types, revealing that adaptor stoichiometry influences substrate choice. ClpS controls substrate selection using distinct mechanisms. The N-terminal segment is essential for delivering N-end-rule substrates but dispensable for ssrA-protein inhibition. We tested existing models for ClpS action and found that ClpS does not block recognition of ssrA-tagged substrates by steric occlusion and that adaptor-mediated tethering of N-end-rule substrates to ClpAP was insufficient to explain facilitated delivery. We propose that ClpS functions, at least in part, as an allosteric effector of ClpAP, broadening our understanding of how AAA+ adaptors control substrate selection.

INTRODUCTION

AAA+ enzymes (ATPases associated with cellular activities) have evolved to catalyze a broad range of cellular functions, including protein degradation, cytoskeletally based motility, and DNA replication, in organisms ranging from bacteria to mammals (Neuwald et al. 1999; Vale 2000; Hanson and Whiteheart 2005). Many AAA+ proteins remodel macromolecular substrates. In these processes, substrate binding and enzyme conformational changes driven by cycles of ATP binding, hydrolysis and nucleotide release result in application of mechanical force that disrupts native protein structures or dismantles complexes. Substrate recognition by AAA+ enzymes is highly regulated, often by the actions of specific adaptor proteins, which have been classically regarded as molecular matchmakers (Mogk et al. 2004; Baker and Sauer 2006). Elucidating adaptor mechanism is therefore essential for a full understanding of substrate selection by AAA+ proteins.

AAA+ proteases ensure protein quality control and are key in many regulatory circuits (Baker and Sauer 2006). ClpAP, one of five AAA+ proteases in *E. coli*, consists of the ClpA ATPase and the ClpP peptidase (Katayama et al. 1988). Hexamers of ClpA

recognize substrates, unfold these proteins, and then translocate the denatured polypeptide through a central pore and into an interior chamber of the ClpP tetradecamer for degradation (Woo et al. 1989; Beuron et al. 1998; Weber-Ban et al. 1999). As observed for many bacterial AAA+ proteases, the ClpA component of ClpAP typically recognizes short peptide tags located at the N or C terminus of substrates (Baker and Sauer 2006). For example, when ribosomal protein synthesis stalls, the tmRNA system, which has tRNA- and mRNA-like functions, mediates cotranslational addition of an 11-residue C-terminal tag (ssrA) onto the nascent polypeptide, which marks the tagged protein for degradation by several proteases, including ClpAP (Keiler et al. 1996; Gottesman et al. 1998).

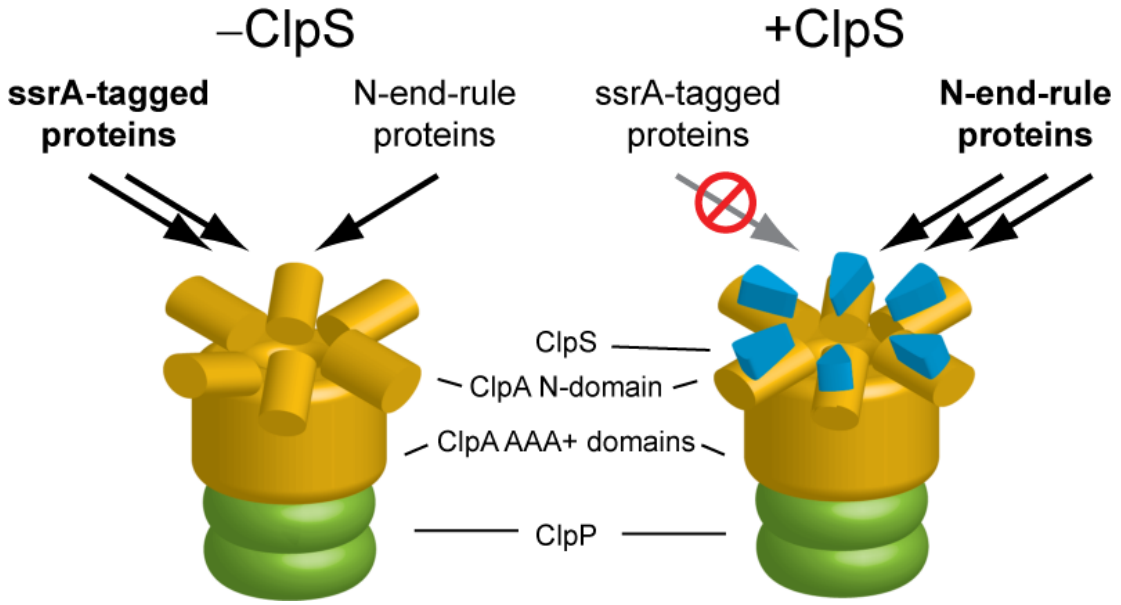


Figure 2.1. Dual roles of ClpS in changing ClpAP substrate preference.
Schematic illustrating the effect of ClpS on the degradation of ssrA-tagged substrates and N-end-rule substrates by ClpAP.

ClpS is a monomeric adaptor protein that binds to the N-terminal domain of ClpA and regulates ClpAP substrate selection (Dougan et al. 2002; Guo et al. 2002a; Zeth et al. 2002) (Fig. 2.1). ClpS inhibits ClpAP degradation of ssrA-tagged substrates and enhances degradation of N-end-rule proteins (Dougan et al. 2002; Erbse et al. 2006; Wang et al. 2007). The N-end rule correlates susceptibility to proteolysis with a protein's N-terminal residue, leucine, tryptophan, tyrosine and phenylalanine being highly

destabilizing at this position (Tobias et al. 1991). ClpS binds directly to N-end-rule substrates and to ClpAP (Erbse et al. 2006), suggesting that simple adaptor-mediated tethering of these substrates to ClpAP might account for their enhanced degradation. Although the structure of ClpS bound to the N-terminal domain of ClpA is known, the mechanisms that allow this adaptor to function both as an enhancer and inhibitor of ClpAP substrate degradation are poorly understood (Guo et al. 2002a; Zeth et al. 2002). The contrasting positive and negative regulatory activities of ClpS intrigued us. Here, we examined how ClpS, by binding to ClpA, inhibits degradation of ssrA-tagged proteins and enhances proteolysis of N-end-rule substrates. We found that substoichiometric binding of ClpS to ClpAP permitted degradation of both substrates. The N-terminal segment of ClpS, which is disordered in several crystal structures (Guo et al. 2002a; Zeth et al. 2002; Xia et al. 2004), was dispensable for inhibition of ssrA-tagged substrate degradation and for directly binding N-end-rule substrates. Unexpectedly, however, N-end-rule protein delivery and ClpS's ability to suppress rates of ClpAP ATP hydrolysis required the N-terminal polypeptide segment, but not a specific sequence, of ClpS. We propose that binding of ClpS to the protease results in conformational rearrangements that allow handoff of N-end-rule substrates to ClpA and prevent binding to ssrA-tagged proteins. Thus, these studies expand the repertoire of molecular mechanisms used by AAA+ adaptor proteins to regulate substrate selectivity.

RESULTS

ClpS binding to ClpA

One ClpS monomer can bind to each subunit in the ClpA hexamer, allowing a maximum stoichiometry of 6:1 (Dougan et al. 2002; Guo et al. 2002a; Zeth et al. 2002). To establish working conditions for subsequent experiments, we explored ClpS binding to hexameric ClpA by immunoprecipitating ClpS–$ClpA_6$ complexes and quantifying bound ClpS (Fig. 2.2a). Half-maximal binding was observed at a ClpS concentration of ~0.4 μM, a value similar to the dissociation constant of 0.33 μM reported for a ClpS complex with the isolated N-terminal domain of ClpA (Zeth et al. 2002). The resulting binding curve revealed an absence of strong cooperativity in the ClpS–$ClpA_6$ interaction (Fig. 2.2b; fit shown uses a Hill coefficient n = 1.5). For comparison, we generated

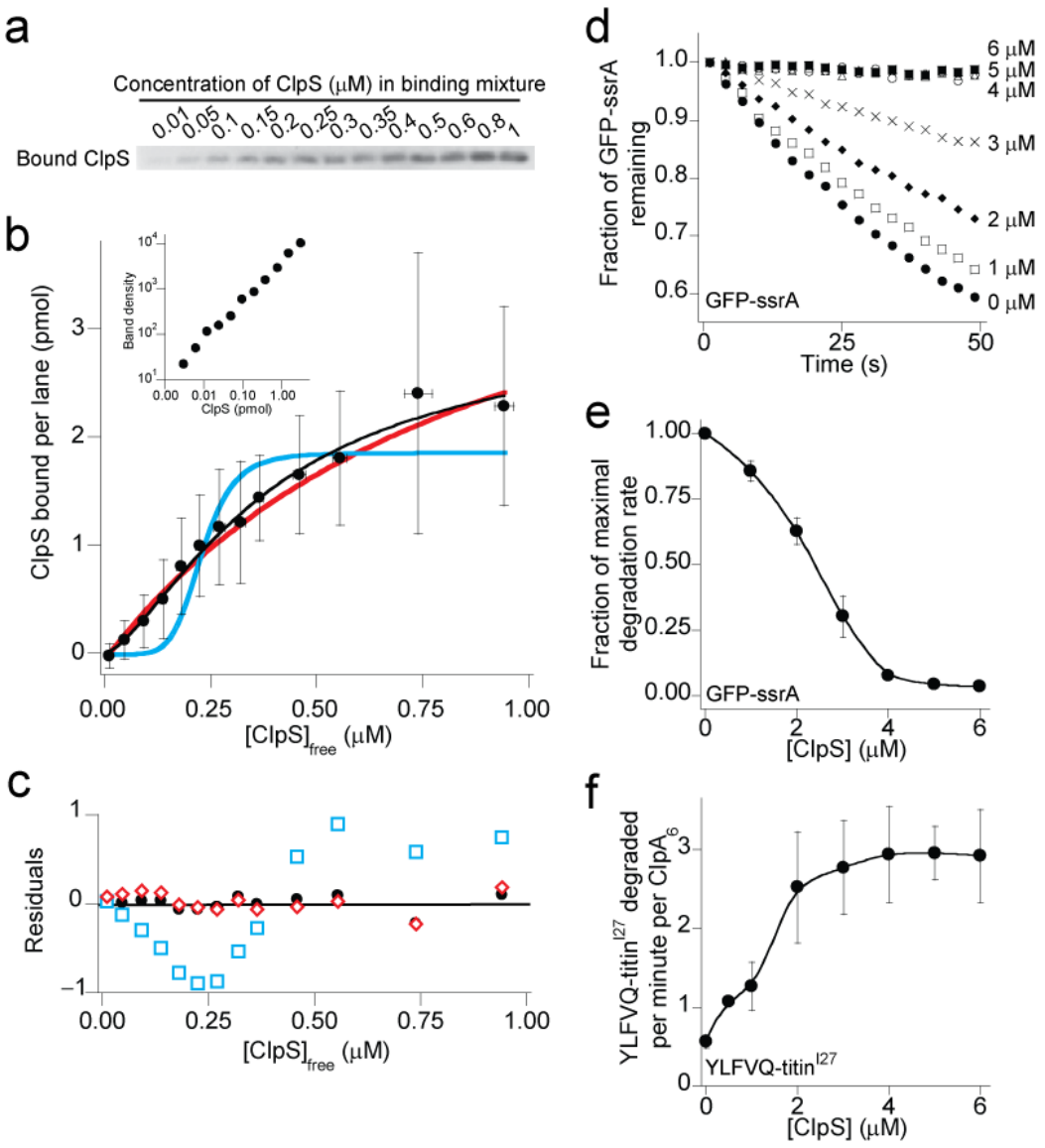


Figure 2.2. Stoichiometry of ClpS.
(a) Western blot of ClpS bound to $ClpA_6$ at increasing concentrations of ClpS. (b) Fitted binding curve of ClpS to $ClpA_6$ from four separate experiments. The red line shows the calculated curve for no cooperativity, and the blue curve is that expected for complete positive cooperativity. Inset, standard curve for the binding assay; all data points fall within the linear range. (c) Residuals showing fit of the data to the various cooperativity models. Black circles, best fit model with cooperativity, n = 1.5; red diamonds, model with no cooperativity, n = 1; blue squares, model with full cooperativity, n = 6. (d) Degradation of GFP-ssrA (5 μM) by ClpAP (1 μM) measured with increasing concentrations of ClpS. (e) Curve showing ClpS stoichiometric effect on GFP-ssrA degradation rates from a with degradation rate in the absence of ClpS set to 1. (f) Curve showing ClpS stoichiometric effect on YLFVQ-titin^{I27} (5 μM) degradation rates by ClpAP (1 μM). Degradation rates were calculated from the increase in trichloroacetic acid–soluble counts in first 110 s of substrate degradation. All error bars indicate ± 1 s.d.

curves of full positive cooperativity (n = 6) and no cooperativity (n = 1). Our data are clearly distinguishable from the binding results expected by a fully cooperative ClpS-$ClpA_6$ interaction (Fig. 2.2b,c). In contrast, the curve for noncooperative binding fit our observed data reasonably well. These results therefore support the idea that ClpS binds to ClpA with little cooperativity, and, as a consequence, many $ClpS_n$–$ClpA_6$ species will be populated at subsaturating levels of ClpS. These data guided our subsequent experiments to study the functional impact of ClpS stoichiometry, which were conducted well above the K_D for the ClpS-$ClpA_6$ interaction.

Effects of subsaturating ClpS on substrate selectivity

ClpS:$ClpA_6$ ratios are known to change under different growth conditions in *E. coli* (Farrell et al. 2005). The stoichiometry of ClpS to $ClpA_6$ could thus provide one level of control on substrate selectivity. To test this model, we monitored ClpAP degradation of an ssrA-tagged protein and an N-end-rule protein, substrates whose degradation ClpS inhibits and enhances, respectively. Degradation reactions were performed at different ClpS:$ClpA_6$ ratios, using ClpA concentrations high enough to ensure that virtually all added ClpS was bound to ClpA. Substantial degradation of green fluorescent protein (GFP) C-terminally tagged with ssrA occurred with one to three equivalents of ClpS, whereas degradation was strongly inhibited when an average of four to six ClpS molecules was bound to $ClpA_6$ (Fig. 2.2d,e). These data further support the conclusion that the ClpS-$ClpA_6$ interaction is not highly cooperative. If ClpS-$ClpA_6$ binding were completely cooperative, addition of 3 μM ClpS would yield precisely 50% GFP-ssrA inhibition because half of the ClpA hexamers would be fully liganded and the other half unliganded. In contrast, we observed ~30% inhibition with this amount of ClpS. Moreover, these functional assays were conducted in the presence of ClpP, indicating that ClpP association does not make binding of ClpS to $ClpA_6$ strongly cooperative. Thus, partial ClpS occupancy of the six sites in a ClpA hexamer permits ClpAP degradation of ssrA-tagged substrates. When $ClpA_6$ is more highly liganded, inhibition becomes stronger and then complete.

Next, we examined how different amounts of ClpS affected ClpAP degradation of a model N-end-rule substrate, the I27 domain of human titin with an appended N-end-

rule signal (YLFVQ-titin^{I27}). To allow observation of stimulatory effects of ClpS, we used a concentration of the YLFVQ-titin^{I27} substrate about six-fold below the Michaelis constant (K_M) determined in the absence of ClpS (Wang et al. 2007). Addition of ClpS enhanced ClpAP degradation of YLFVQ-titin^{I27}, with maximal stimulation occurring when four or more ClpS molecules bound $ClpA_6$ (Fig. 2.2f). However, we also observed substantial increases in degradation with fewer than three equivalents of ClpS. Although decreased degradation of ssrA-tagged substrates and enhanced degradation of N-end-rule substrates occurred over similar ClpS:$ClpA_6$ ratios, the curves were not strictly reciprocal, raising the possibility that different mechanisms could be responsible for the inhibitory and activating activities of ClpS.

Small changes in ClpS levels alter proteolytic preference

Our results suggested that ClpS acts as a variable rheostat rather than a binary switch for ClpA, with substoichiometric levels of bound ClpS yielding intermediate activities. Based on the data in Figure 2.2, we sought to determine whether ClpAP partially filled with ClpS molecules could degrade a mixture of GFP-ssrA and YLFVQ-titin^{I27} at substantial rates or whether one of these substrates would be preferentially degraded. In the absence of ClpS, ClpAP degraded GFP-ssrA faster than YLFVQ-titin^{I27} (Fig. 2.3a). These differences in rate were due to the identity of the tags rather than the use of different attached substrate proteins (Flynn et al. 2001; Wang et al. 2007) (see Fig. 2.3 legend). When two ClpS equivalents were added, this substrate preference was reversed, although both substrates were still efficiently degraded. Moreover, adding two ClpS equivalents had a stronger effect on the ClpAP degradation rate of YLFVQ-titin^{I27} than of GFP-ssrA (Fig. 2.3b). These results demonstrate that addition of a few equivalents of ClpS is sufficient to switch the proteolytic preference of ClpAP to favor N-end-rule substrates without excluding recognition of ssrA-tagged substrates. Substrate preferences can therefore be fine-tuned by modest changes in adaptor-enzyme stoichiometry.

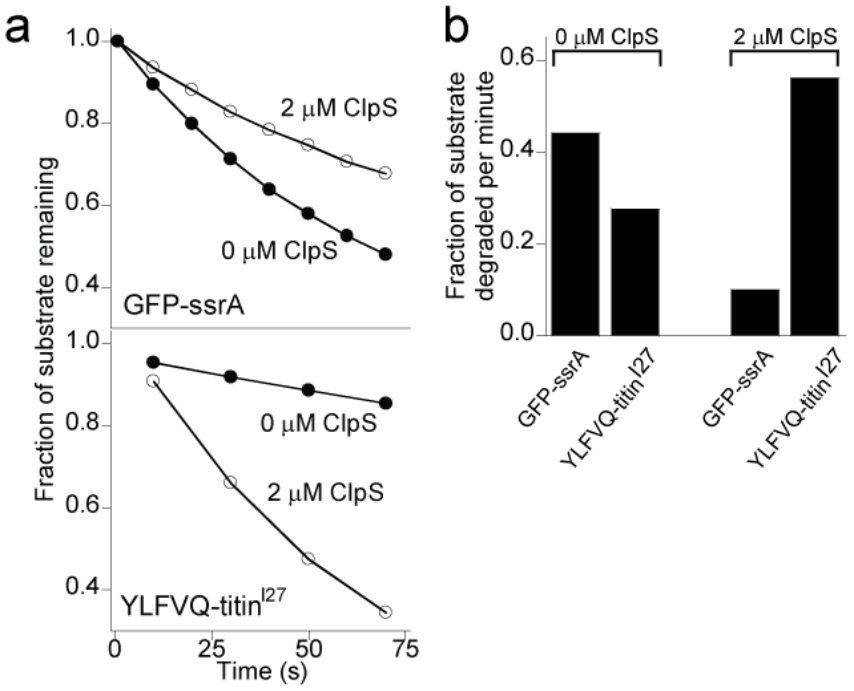


Figure 2.3. Co-degradation of GFP-ssrA and YLFVQ-titin^{I27}.
Comparison of GFP-ssrA and YLFVQ-titin^{I27} degradation rates with substrates combined in the same reaction. (a) Degradation of GFP-ssrA (5 μM) and YLFVQ-titin^{I27} (5 μM), each in the presence of the other substrate, with 1 μM ClpAP and 0 μM or 2 μM ClpS. (b) The addition of two equivalents of ClpS alters the degradation rate of each substrate as well as modifies ClpAP's substrate preference. Degradation rates are from a. Reported V_{max} and K_M values for YLFVQ-titin^{I27} in the absence of ClpS are 11 molecules degraded per minute per $ClpA_6$ and 29 μM, respectively (Wang et al. 2007). Reported V_{max} and K_M values for GFP-ssrA in the absence of ClpS are 4.9 molecules per minute per $ClpA_6$ and 1.5 μM, respectively (Flynn et al. 2001).

The N terminus of ClpS is not a steric inhibitor

Because it seemed unlikely that a common set of interactions was responsible for both the inhibitory and stimulatory effects of ClpS, we sought to identify the distinct structural elements of ClpS that might be important for its dual regulatory roles. ClpS contains a folded core domain, which binds the N-terminal domain of ClpA and is highly conserved phylogenetically. This domain is preceded by an N-terminal extension, which is highly variable in both length and sequence among ClpS orthologs (Dougan et al. 2002; Guo et al. 2002a; Zeth et al. 2002; Xia et al. 2004) (Fig. 2.4a, Fig. 2.5a). Despite the lack of conservation, *E. coli* ClpS with the 17 N-terminal residues removed by proteolysis does not inhibit degradation of GFP-ssrA, suggesting that at least some of the missing residues are functionally important (Guo et al. 2002a; Xia et al. 2004).

Previous experiments indicate that ClpA binding to full-length ClpS and to the ssrA tag is competitive (Dougan et al. 2002). These results, together with the observed effect of removing the first 17 residues, led us to propose a model in which the N-terminal peptide of ClpS and the ssrA-tag compete directly for a common binding site in ClpA (Guo et al. 2002a; Xia et al. 2004) (Fig. 2.4b). To test the importance of the N-terminal sequence of ClpS, we constructed variants in which sequential blocks of four residues were mutated to alanine, as well as a protein with residues 2-17 mutated en masse to an arbitrary assortment of small and positively charged residues (Fig. 2.4c). The methionine at position 1 is removed in vivo in wild-type ClpS and the full-length variants (data not shown). All of the full-length ClpS proteins inhibited ClpAP degradation of GFP-ssrA degradation with activities indistinguishable from wild-type ClpS (Fig. 2.4d). We conclude that the identity of the N-terminal amino acid sequence of *E. coli* ClpS is not important for its function in inhibiting ClpAP's ability to degrade GFP-ssrA.

To further address whether the N-terminal region of ClpS was important for inhibition, we constructed and purified a truncated ClpS variant starting at residue 18 ($ClpS^{M18-106}$). Mass spectrometry showed that most of this truncated protein retained the initiator methionine, making it one residue longer than ClpS with the first seventeen residues removed proteolytically. Unexpectedly, $ClpS^{M18-106}$ inhibited ClpAP degradation of GFP-ssrA to an extent similar to that of wild-type ClpS (Fig. 2.4e). By contrast, when we prepared N-terminally truncated ClpS by protease cleavage in vitro, it had the reported behavior and did not inhibit ClpAP degradation of GFP-ssrA (Guo et al. 2002a) (data not shown). We conclude that contacts between ClpA and determinants in or near residue 17 of ClpS are required to inhibit degradation of GFP-ssrA, whereas the N-terminal 16 residues of ClpS are not needed for this activity.

Overexpression of $ClpS^{M18-106}$ protein probably overwhelmed the cellular processing machinery, resulting in poor cleavage of the N-terminal Met-Val peptide bond by methionine aminopeptidase. To avoid this problem, we constructed $ClpS^{18-106}$ V18A, because Met-Ala is processed more efficiently than Met-Val (Hirel et al. 1989). Mass spectrometry and Edman degradation showed that the N-terminal methionine of this protein was removed (data not shown). $ClpS^{18-106}$ V18A, like the proteolytically

generated ClpS$^{18\text{-}106}$, did not inhibit ClpAP degradation of GFP-ssrA substantially (Fig. 2.4e), but it did bind ClpA, as observed by size-exclusion chromatography (data not shown). Further truncations beginning C-terminal to the seventeenth residue of ClpS did not inhibit degradation of GFP-ssrA, but a variant one residue larger than ClpS$^{M18\text{-}106}$ (ClpS$^{M17\text{-}106}$) showed roughly 80% of the inhibitory activity of wild type (Fig. 2.5b). These results reinforce the importance of a residue at position 17 of ClpS for efficient inhibition of ClpAP degradation of an ssrA-tagged protein. However, based on observations with ClpS$^{M18\text{-}106}$, we can rule out the model (Fig. 2.4b) in which the first sixteen residues sterically block the binding site for the ssrA-tag in ClpA, suggesting that another mechanism must be responsible for inhibition.

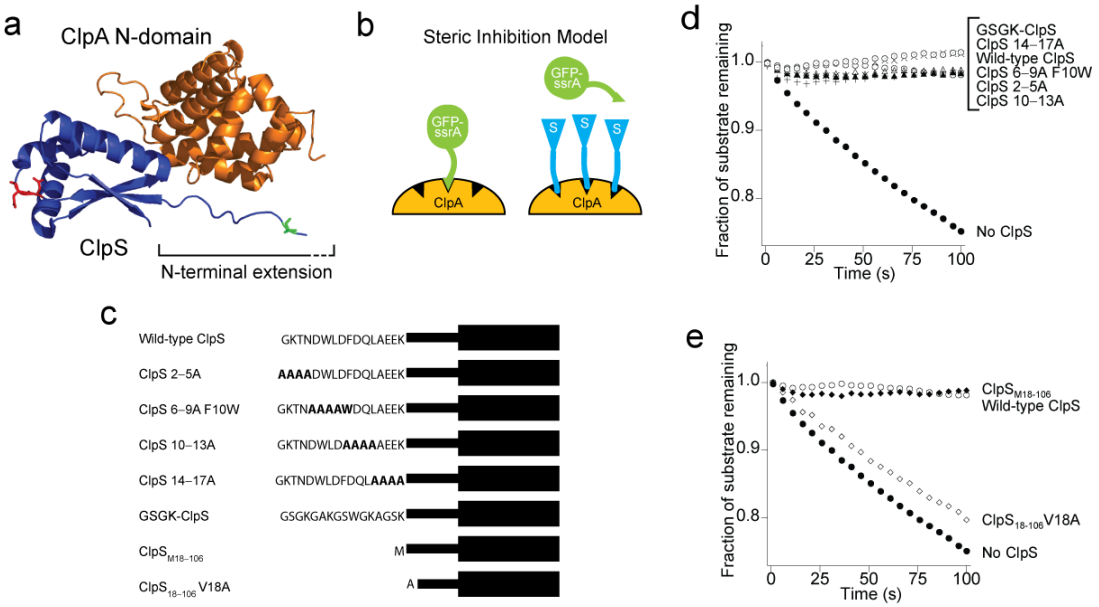


Figure 2.4. ClpS N-terminal extension is not a steric inhibitor of GFP-ssrA.
(a) Structure of ClpS (blue) and ClpA N-domain (gold), PDB accession code 1LZW (Zeth et al. 2002). Residues on ClpS proposed to bind N-end-rule motifs (Erbse et al. 2006) are depicted as red sticks (on left), and the lysine at the seventeenth position is shown with green sticks (on right). The first fifteen residues are not modeled into this structure. (b) Model for ClpS (S, blue) inhibition of ssrA-tagged protein degradation by ClpAP in which the ClpS N-terminal extensions directly occlude ssrA binding sites on ClpA. (c) ClpS variants constructed in this study. ClpS 2-5A, ClpS 6–9A F10W, ClpS 10–13A and ClpS 14–17A are full-length variants of ClpS in which the ranges of residues indicated are mutated to alanines. (d) Comparison of the inhibition of GFP-ssrA (1 mM) degradation by 1 mM full-length ClpS variants (filled circle, no ClpS; open circle, ClpS; filled triangle, ClpS 2–5A; open triangle, ClpS 6–9 F10W; +, ClpS 10–13A; ×, ClpS 14–17A; dotted circle, GSGK-ClpS) and by (e) ClpS N-terminal extension truncations (filled diamond, ClpS$^{M18–106}$; open diamond, ClpS$^{18–106}$ V18A) with ClpAP (0.1 μM ClpA$_6$, 0.27 μM ClpP$_{14}$).

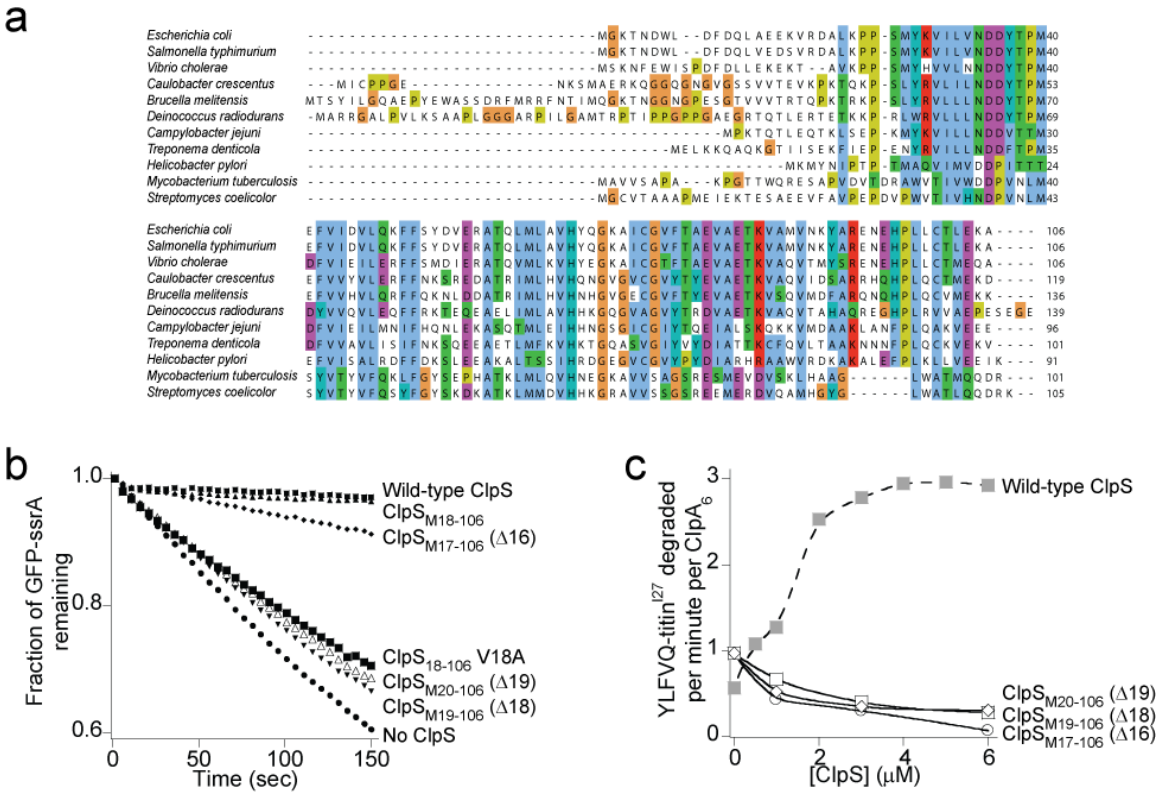


Figure 2.5. Activity of ClpS deletion mutants.
(a) Alignment of ClpS from various bacterial species. Residues similar in nature are colored. (b) Degradation of GFP-ssrA (1 μM) by $ClpA_6$ (0.1 μM) and $ClpP_{14}$ (0.27 μM) with variable length ClpS deletion proteins (1 μM). (c) Concentration-dependent degradation of YLFVQ-titin^{I27} (5 μM) by $ClpA_6$ (1 μM) and $ClpP_{14}$ (1 μM) with ClpS deletion proteins. The wild-type ClpS curve from Figure 2f is shown in gray for comparison.

The ClpS extension is necessary for N-end-rule delivery

We next tested whether ClpS variants with mutations in or truncations of the N-terminal extension could enhance ClpAP degradation of an N-end-rule substrate. Figure 2.6a depicts a model in which enhanced N-end-rule degradation results solely from ClpS-mediated tethering of the substrate to ClpA. The purported N-degron binding site of ClpS lies within the core domain, and the simple tethering model predicts that ClpS variants that bind both to N-end-rule substrates and to ClpA should enhance degradation (Erbse et al. 2006). However, the ClpS$^{M18\text{-}106}$ and ClpS$^{18\text{-}106}$ V18A mutants inhibited ClpAP degradation of YLFVQ-titin^{I27} instead of stimulating turnover (Fig. 2.6b). Both mutants bound YLFVQ-GFP (which is degraded by ClpAP; K. Wang, Massachusetts

Institute of Technology, personal communication) in gel filtration experiments (Fig. 2.6c) but did not bind a GFP variant lacking the N-degron (data not shown). Moreover, both $ClpS^{M18-106}$ and $ClpS^{18-106}$ V18A bound ClpA (data not shown). Hence, the $ClpS^{18-106}$ variants maintained the ability to bind an N-end-rule degradation tag and the ClpA enzyme but could not deliver tagged substrates to ClpAP for degradation. Other ClpS truncations ($ClpS^{M17-106}$ (Δ16), $ClpS^{M19-106}$ (Δ18), and $ClpS^{M20-106}$ (Δ19)) also failed to deliver YLFVQ-$titin^{I27}$. Thus, residues N-terminal to the sixteenth position of ClpS are important for N-end-rule substrate delivery (Fig. 2.5c).

The full-length mutant GSGK-ClpS and even ClpS 14-17A bearing an alanine at the seventeenth residue had activity similar to wild-type ClpS for N-end-rule substrate delivery (Fig. 2.6b and data not shown). Taken together, these results show that the sequence of the ClpS N-terminal extension is not important for N-end-rule protein degradation, but the presence of this polypeptide segment is essential. We conclude that simple tethering of N-end-rule substrates to ClpA by ClpS is insufficient to explain ClpS-stimulated degradation. Rather, the N terminus of ClpS must play an active role in triggering the delivery and degradation of N-end-rule substrates.

To gain further insight into the role of the N-terminal extension of ClpS in substrate delivery, we asked whether wild-type ClpS could stimulate degradation N-end-rule proteins when some sites on ClpAP were occupied by an inhibitory version of ClpS lacking the N-terminal extension ($ClpS^{M18-106}$). Notably, the presence of an N-terminal extension was able to overcome the suppressive effects of $ClpS^{M18–106}$ alone (Fig. 2.7). These data demonstrate that the N-terminal extension of ClpS is a distinct and critical module of the adaptor as well as reinforce the conclusion that substoichiometric levels of the wild-type adaptor are able to influence the substrate preference of the entire ClpA hexamer.

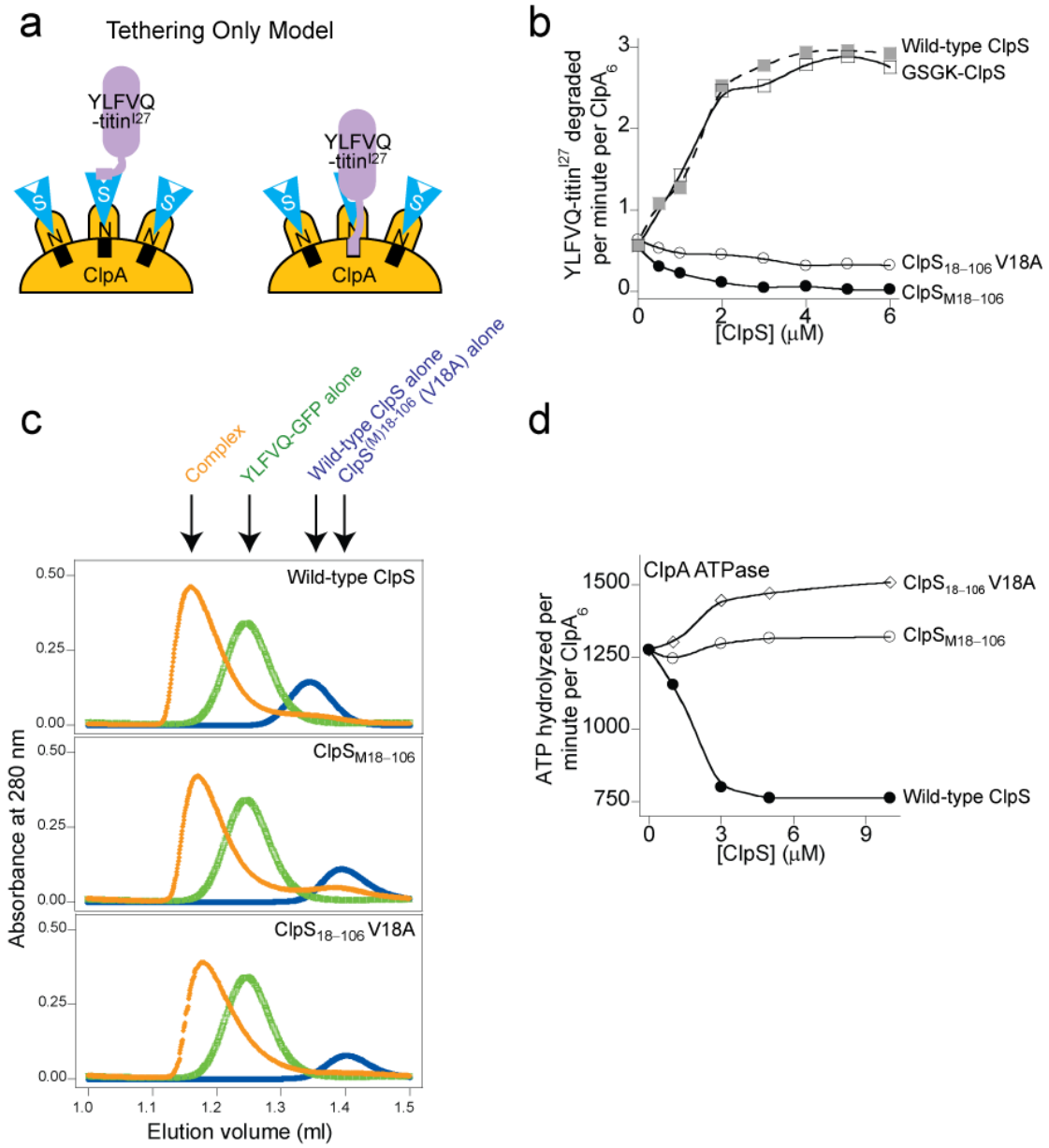


Figure 2.6. ClpS N-terminal extension is required for active delivery of YLFVQ-titin^{I27} and for effect on ClpAP ATPase rate.
(a) Model in which simple tethering of N-end-rule proteins by the folded domain of ClpS to ClpA is sufficient to enhance degradation. The white triangles on ClpS (S, blue) represent the key adaptor residues suggested to bind and tether N-end-rule substrates. (b) Concentration-dependent degradation of YLFVQ-titin^{I27} (5 μM) by ClpAP (1 μM) in the presence of ClpS$^{M18-106}$, ClpS$^{18-106}$ V18A and GSGK-ClpS. Shown for comparison in gray squares connected by the dotted line is the data for wild-type ClpS from Figure 2.2b. (c) Size-exclusion chromatography results showing binding of truncated ClpS variants to YLFVQ-GFP. (d) Comparison of concentration-dependent effects of ClpS and variants on ClpA (1 μM) ATPase rate in the presence of ClpP (1 μM).

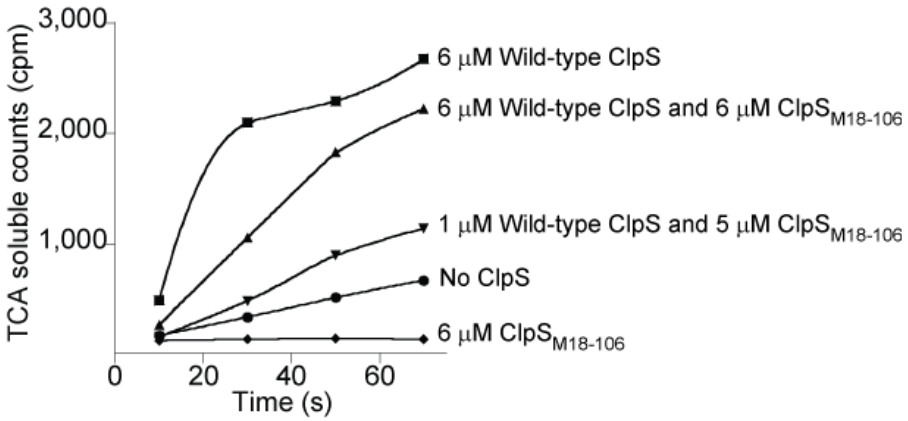


Figure 2.7. The N-terminal extension of ClpS is a separable and critical module for N-end-rule delivery.
Degradation of YLFVQ-titin^{I27} (5 μM) with wild-type or truncated ClpS alone or in mixtures.

The N terminus of ClpS alters ATP hydrolysis by ClpA

ClpP binding increases the rate of ATP hydrolysis by ClpA two-fold (Hinnerwisch et al. 2005b), an effect that is likely to reflect an alteration of ClpA's conformation accompanying ClpP interaction. We asked whether ClpS binding and/or the presence of the N-terminal extension might also change the ATPase activity of ClpAP in a fashion consistent with the promotion or stabilization of a conformational change in the protease. Indeed, ClpS binding depressed the ATPase rate of ClpAP ~40% (Fig. 2.6d), as did the full-length variants depicted in Figure 2.4c (data not shown). In contrast, adding ClpS$^{M18–106}$ caused little change in ATPase rate, whereas adding ClpS$^{18-106}$ V18A increased the rate of ATP hydrolysis (Fig. 2.6d). Both truncated ClpS preparations were free of contaminating ATPase activity (data not shown). We considered the possibility that wild-type ClpS could be lowering ClpAP ATPase rates by weakening the ClpA-ClpP interaction in a manner similar to the weakening of the interaction between ClpP and ClpA observed when ClpA lacks its N-terminal domain (Hinnerwisch et al. 2005b). However, control binding experiments indicated that the ClpA-ClpP interaction was not appreciably altered by ClpS association (data not shown), and it seems unlikely that ClpP would dissociate from the ClpS–ClpA$_6$ complex during degradation, especially for degradation of N-end-rule substrates. We conclude that binding of the core ClpS domain to ClpAP enhances ATPase activity slightly, whereas interactions mediated by the N-terminal 16 residues of ClpS result in a substantial depression of ATPase activity.

Because the rate of ATP hydrolysis by ClpAP must ultimately depend on enzyme conformation and the kinetic barriers between conformations, our results indicate that the N-terminal residues of ClpS cause a structural shift in the ClpA hexamer. This altered conformation of ClpA could, in turn, be an important mechanistic component of ClpS-mediated substrate selection.

DISCUSSION

The ClpS adaptor controls ClpAP using multiple strategies. Previous studies established that saturating ClpS levels prevent ClpAP degradation of ssrA-tagged proteins and stimulate degradation of N-end-rule substrates (Dougan et al. 2002; Guo et al. 2002a; Erbse et al. 2006; Wang et al. 2007). We found that subsaturating ClpS concentrations allowed ClpAP to degrade both types of substrates. With 2 μM ClpS and 1 μM $ClpA_6$ (2:1 ratio), for example, ClpAP degraded ssrA-tagged and N-end-rule substrates at 40–60% of maximal rates in vitro. In contrast, with 5 μM ClpS and 1 μM $ClpA_6$ (5:1 ratio), degradation of ssrA-tagged substrates was minimal, and N-end-rule degradation reached maximal efficiency. In *E. coli*, the ClpS:$ClpA_6$ ratio is ~6:1 during exponential growth but changes to ~2:1 as cells enter stationary phase. During this transition, ClpAP degradation of GFP-ssrA increases substantially (Farrell et al. 2005). Thus, changes in the ClpS:$ClpA_6$ ratio fine-tune the substrate preferences of ClpAP, presumably promoting degradation of N-end-rule and perhaps other classes of substrates during logarithmic growth but expanding the effective substrate repertoire during stationary phase.

The mechanisms that allow ClpS to enhance N-end-rule substrate degradation but inhibit ssrA-tagged protein degradation are clearly different. We found that much of the unstructured N-terminal extension of ClpS was required for stimulation of N-end-rule degradation but was dispensable for inhibition of ClpAP degradation of ssrA-tagged substrates. Specifically, a ClpS variant missing the N-terminal 16 residues and having a Lys17 to Met substitution ($ClpS^{M18\text{-}106}$) inhibited ClpAP degradation of GFP-ssrA but did not stimulate degradation of the model N-end-rule substrate YLFVQ-$titin^{I27}$. Notably, both $ClpS^{M18\text{-}106}$ and $ClpS^{18\text{-}106}$ V18A still bound N-end-rule substrates. Two main conclusions emerge from these observations. First, a model in which the 16 N-terminal

residues of ClpS physically occlude binding of the ssrA tag to ClpA (Fig. 2.4b) is ruled out, as $ClpS^{M18\text{-}106}$ inhibited GFP-ssrA degradation efficiently. The ssrA tag binds in the central channel of the ClpA hexamer (Hinnerwisch et al. 2005a; Piszczek et al. 2005), and steric blockage by the remaining unstructured residues of $ClpS^{M18\text{-}106}$ seems unlikely. Second, simple tethering of N-end-rule substrates to ClpAP by the ClpS adaptor (Fig. 2.6a) cannot explain the enhanced degradation of this class of substrates. ClpA and ClpS both bind N-end-rule substrates independently and therefore must both contain N-degron binding sites (Erbse et al. 2006; Wang et al. 2007). Because both $ClpS^{18\text{-}106}$ variants inhibit rather than enhance degradation of YLFVQ-$titin^{I27}$, a mechanism that involves more than simple tethering must be required. Thus, preexisting models of adaptor function do not explain how ClpS functions either as a positive or negative regulator of protein degradation.

With the simplest mechanistic models ruled out, we turned our attention toward the observation that ClpS binding altered ClpAP ATPase rates. ClpS variants that lacked most of the N-terminal extension were both unable to deliver N-end-rule proteins for degradation and unable to suppress ClpAP ATPase rates, suggesting that the two effects are related. Similarly, the difference between the abilities of $ClpS^{M18\text{-}106}$ and $ClpS^{18\text{-}106}$ V18A to stimulate ClpAP ATPase rates may reflect their differing abilities to inhibit GFP-ssrA degradation. These data suggest that an allosteric mechanism contributes to ClpS's ability to control substrate choice (Fig. 2.8). Although ClpS binding to $ClpA_6$ is almost completely independent, there is likely to be communication between ClpA subunits, resulting in substrate binding and processing that is nonlinear in response to ClpS stoichiometry. Binding of a single ClpS molecule to a ClpA hexamer, for example, could relay a conformational signal within the bound ClpA monomer, which in turn would affect neighboring subunits. We therefore rule out the two simple models for ClpS activity and present an initial version of a unified model consistent with our results.

In this model, the N-terminal extension of ClpS interacts with two regions on ClpA that induce conformational changes in the enzyme. The contact near residue 17 of ClpS is critical for altering the ssrA recognition, whereas the contact made by the more N-terminal residues is important for altering N-end-rule protein recognition (Fig. 2.8a,b). Present evidence indicates that the ssrA tag is recognized by the pore of ClpA

(Hinnerwisch et al. 2005a; Piszczek et al. 2005); similar analysis has not yet been done with N-degron tags, but we have schematically represented these binding sites in the ClpA pore in Figure 2.8b. We speculate that the allosteric change induced upon ClpS binding make the ssrA binding sites less favorable and the N-degron sites more favorable, altering observed degradation rates. Changes in substrate preference would be reflected by the relative changes in the sizes of the substrate binding sites in the central channel of the protease (Fig. 2.8b). Because there is no absolute sequence requirement for the N-terminal extension of ClpS, it seems likely that peptide backbone atoms mediate the interaction with ClpA. In addition to inducing conformational changes in ClpA, the N-terminal ClpS extension may also play an active role in orchestrating handoff of N-end-rule substrates from the adaptor to ClpA. For example, this flexible N-terminal region may be engaged by ClpA to help 'pull in' the adaptor-bound substrate closer to the central pore. More experiments will be needed to clarify the detailed mechanisms underlying ClpS-mediated modulation of ClpAP substrate selection.

Many AAA+ adaptors function by binding to both a substrate and a AAA+ protease, thereby increasing the local concentration of the substrate relative to the enzyme and allowing it to be degraded more efficiently. SspB, an adaptor for the ClpXP protease, represents a paradigm for this class of positive regulation of substrate choice by tethering (Levchenko et al. 2000; Dougan et al. 2003; Wah et al. 2003). Dimers of SspB bind to sites in the N-terminal domains of ClpX and to peptide sequences in specific substrates but allow the degradation tags of these substrates to be engaged in the central pore of the ClpX hexamer (Bolon et al. 2004). When loaded with substrate, only one SspB dimer binds stably to ClpXP (Wah et al. 2002). Thus, SspB delivers substrates in an all-or-none fashion. ClpS represents a new class of AAA+ adaptor. Simple tethering is not sufficient for its function in stimulating ClpAP degradation of N-end-rule substrates. ClpS acts both as a positive and negative regulator of substrate choice, and changes in $ClpS_n$-$ClpA_6$ binding stoichiometry allow variable rather than switch-like regulation. It will be important to determine whether adaptors like MecA, an indispensable activator that binds the *Bacillus subtilis* ClpCP protease with an apparent stoichiometry of 6:1, function using mechanisms analogous to those of ClpS (Kirstein et al. 2006).

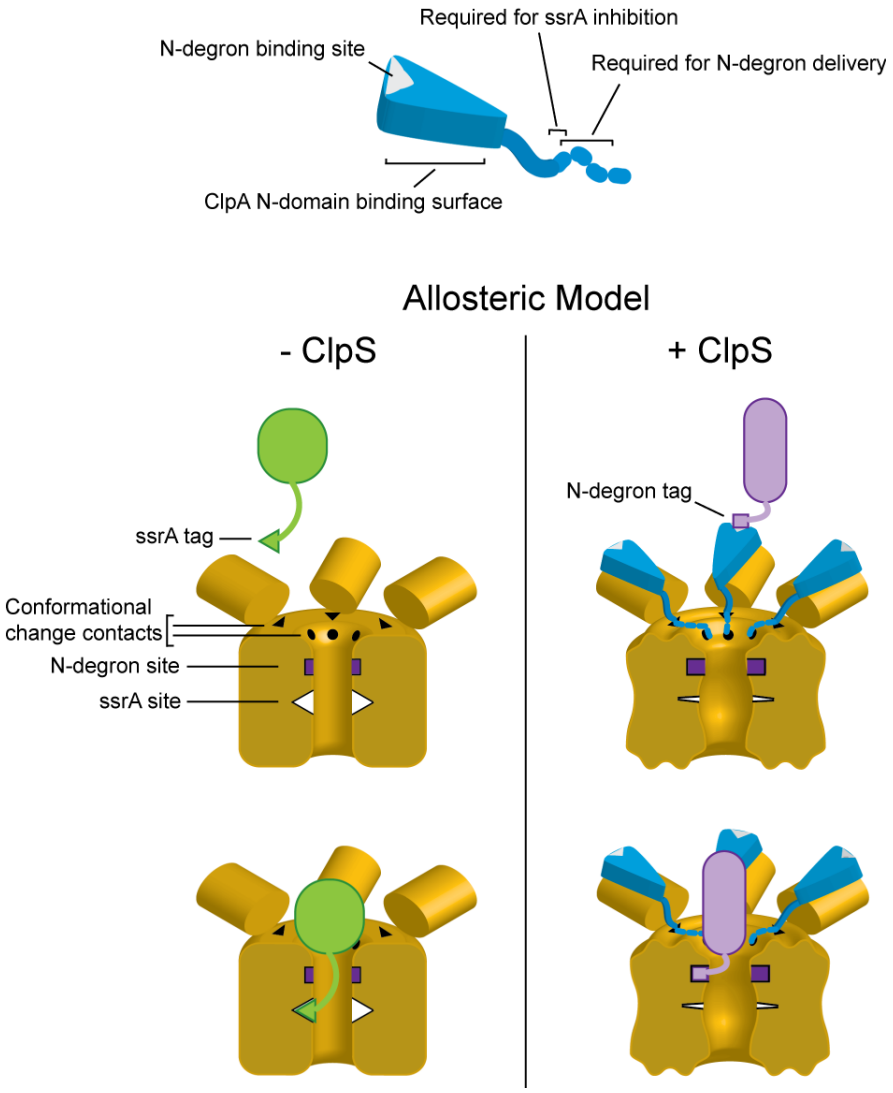


Figure 2.8. Model for ClpS-mediated change in ClpA binding preferences. Top, regions of ClpS conferring distinct activities. Bottom, cutaway views of ClpA hexamers (gold) with representation of N-degron binding sites (purple) and ssrA binding sites (white) in the central pore. ClpS, blue; putative N-degron binding site on the adaptor, white triangles; first sixteen residues of the N-terminal extension, dashed blue line (visible in the top panel); black triangles and circles, points of contact on ClpA for the distinct regions of the N-terminal extension responsible for relaying the ClpS signal and modifying ClpA substrate preference. The initial binding sites for the ssrA and N-end-rule tags are unknown, but because proteins slated for degradation must travel through the axial pore, we model substrate binding sites in the central pore of ClpA. The size differences of substrate binding sites represent the degree of ClpA's affinity toward the respective substrates. Upon adaptor binding, the ClpS N-terminal extensions contact the symbolic black triangles and circles on ClpA, triggering a conformational change in the enzyme. The allostery induced by ClpS binding modifies the affinities of substrates of ClpA, represented by changes in binding-pocket sizes in the pore and overall enzyme shape.

MATERIALS AND METHODS

Strains, plasmids, proteins and peptides

Full-length ClpS alanine variants were generated using the QuikChange Site-Directed Mutagenesis Kit protocol (Stratagene). We chose N-terminal residues of GSGK-ClpS that were mostly small and positively charged to retain solubility and flexibility. We cloned all ClpS constructs (untagged) into pET-23b(+) (EMD Biosciences). $ClpS^{M18\text{-}106}$ and $ClpS^{18\text{-}106}$ V18A were expressed in strain BL21 (DE3) *clpS::kan* pLysS. Other ClpS variants were expressed in BL21 (DE3) pLysS. For expression, cells were grown at 30 °C until OD_{600} 0.4-0.6 and then induced with 0.4 mM IPTG for two hours. All variants were purified as described (Dougan et al. 2002) with the following modifications: we did not use ion exchange purification for $ClpS^{M18\text{-}106}$, $ClpS^{18\text{-}106}$ V18A and ClpS 10-13A; we purified the GSGK-ClpS construct using a MonoS cation exchange column instead of a MonoQ anion exchange column; and ClpS 6-9A F10W protein failed to bind to the MonoQ column and we further purified it by gel filtration as described (Dougan et al. 2002).

ClpA was expressed from plasmid *clpA* M169T/pET9a (see Acknowledgments) in strain BL21 (DE3) pLysS. After French-press lysis, we brought the cleared lysate to 40% (w/v) saturated ammonium sulfate and centrifuged it, then resuspended the pellet in S-Sepharose buffer (25 mM HEPES, pH 7.5, 2 mM DTT, 0.1 mM EDTA, 10% (v/v) glycerol) and centrifuged again. We loaded the supernatant onto an S-Sepharose column and eluted the protein in a gradient from 200 mM to 1 M KCl in S-Sepharose buffer. Peak fractions were combined, brought up to 15% (w/v) saturated ammonium sulfate and chromatographed on a Source 15PHE column. The protein eluted in a linear gradient from 0.6 M ammonium sulfate to 0.6% (w/v) CHAPS detergent in 50 mM sodium phosphate, pH 7.5, 2 mM DTT, and 10% (v/v) glycerol. Peak fractions were dialyzed into 50 mM HEPES, pH 7.5, 20 mM $MgCl_2$, 0.3 M NaCl, 10% (v/v) glycerol and 0.5 mM DTT. Columns and separation media were from GE Healthcare.

YLFVQ-$titin^{I27}$, [^{35}S] YLFVQ-$titin^{I27}$ and YLFVQ-GFP protein were provided by K. Wang and purified as described (Wang et al. 2007). GFP-ssrA and ClpP were purified as described (Kim et al. 2000; Flynn et al. 2001).

Biochemical assays

To test binding of ClpS to $ClpA_6$, we preincubated different concentrations of ClpS for a minimum of 5 min with 10 nM $ClpA_6$ in 50 mM HEPES (pH 7.5), 20 mM $MgCl_2$, 0.3 M NaCl, 10% (v/v) glycerol, 0.1% (v/v) Tween, 0.5 mM DTT and 1 mM ATP-γS at 30 °C. Protein A Dynabeads (Invitrogen) conjugated with purified antibody to ClpA were added to the protein mixture for 5 min at 30 °C. We isolated the beads were isolated and removed the supernatant ('unbound' samples), then added SDS loading buffer to the beads and incubated briefly at 37 °C ('bound' samples). Bound samples and ClpS standards were separated by SDS-PAGE followed by an anti-ClpS western blot. We quantified bands with ImageQuant TL software (GE Healthcare). Nonspecific association of ClpS to the beads in the absence of $ClpA_6$ was quantified and subtracted from binding data.

We assayed degradation with stoichiometric ClpS at 30 °C with 1 μM $ClpA_6$, 1 μM $ClpP_{14}$, 4 mM ATP, an ATP-regeneration system (2.5 mM creatine phosphate and 50 μg ml–1 creatine kinase) and 5 μM substrate in 50 mM HEPES (pH 7.5), 20 mM $MgCl_2$, 0.3 M NaCl, 10% (v/v) glycerol and 0.5 mM DTT. GFP-ssrA degradation was monitored by decreases in fluorescence (excitation, 467 nm; emission, 511 nm) using a Photon Technology International instrument. [^{35}S] YLFVQ-titin^{I27} degradation was monitored by the release of trichloroacetic acid–soluble counts as described (Wang et al. 2007). Other degradation assays were conducted using the same conditions but with 0.1 μM $ClpA_6$, 0.27 μM $ClpP_{14}$, 1 μM ClpS and 1 μM GFP-ssrA substrate.

To assay binding of N-end-rule substrates to ClpS or variants, we chromatographed approximately 20 μM initial concentrations of wild-type ClpS, ClpS$^{M18-106}$ and ClpS$^{18-106}$ V18A alone or in complex with YLFVQ-GFP on a Superdex75 PC 3.2/30 gel filtration column in 20 mM HEPES, pH 7.5, and 150 mM KCl. Elution profiles were monitored by absorbance at 280 nm and 488 nm.

We measured rates of ClpAP ATP hydrolysis using a coupled assay as described (Kim et al. 2001). $ClpA_6$ (1 μM), $ClpP_{14}$ (1 μM) and different concentrations of ClpS were incubated for 1 min at 30 °C before adding ATP mix III.

Alignment of ClpS sequences

ClustalW alignments from ClpS protein sequences of various bacterial species were analyzed using the Jalview alignment editor (Clamp et al. 2004).

ACKNOWLEDGMENTS

We thank K. Wang (Massachusetts Institute of Technology) for generously providing N-end-rule proteins and P. Chien, M. Laub, T. Schwartz, K. Wang, C. Wu and members of the Baker and Sauer Labs for discussion and advice. Plasmid *clpA* M169T/pET9a was a gift from J. Flanagan (Hershey Medical Center). T.A.B. is an employee of the Howard Hughes Medical Institute. This work was supported by US National Institutes of Health grant GM49224 and the Howard Hughes Medical Institute.

CHAPTER THREE

The ClpS N-terminal extension requires a minimal length and is affected by chemical composition

ABSTRACT

Adaptor proteins play key regulatory roles in substrate selection by intracellular energy-dependent proteases. Here, we expand upon previous studies in which we discovered that the N-terminal extension of the *E. coli* adaptor ClpS is essential for its role in delivering N-end-rule proteins for degradation by the AAA+ protease ClpAP. We suggest that the N-terminus of ClpS is an essential allosteric effector of substrate recognition, despite a lack of length and sequence conservation among ClpS orthologs. We uncover a striking length-dependent function for the N-terminal extension; a difference in length of one residue distinguishes functional and non-functional variants of ClpS. Furthermore, drastic sequence mutations of the N-terminal extension reveal a previously unobserved effect on substrate delivery ability. Finally, we localize an interaction site of the ClpS N-terminal extension to the region of the ClpA pore. These studies, therefore, increase our knowledge for how this unique adaptor modulates protease substrate selection.

INTRODUCTION

ATP-dependent enzymes are involved in a multitude of cellular processes. Many of these proteins contain a similar ATP-binding protein domain, called the AAA+ domain. The AAA+ (ATPases associated with various cellular activities) protein fold has been well-conserved throughout evolution, and these domains, whether in isolated form or as part of a larger protein subunit, typically form hexameric ring structures with the ATP-binding site between monomers. The conformational change in the AAA+ hexamer resulting from ATP binding and hydrolysis supplies a mechanical force that can be imparted onto other proteins. Thus, AAA+ proteins are involved in many cellular processes requiring the use of mechanical force such as protein unfolding, protein disaggregation, and complex disassembly (Hanson and Whiteheart 2005).

The processing of substrate proteins by AAA+ enzymes typically involves the recognition and threading of the substrate protein through the central pore of the AAA+ hexamer, resulting in substrate unfolding and potentially substrate complex dissociation. AAA+ proteins have particular substrate recognition properties, but other domains and regulatory proteins can also have a profound effect on the substrate selection of AAA+

enzymes. For the *E. coli* protease ClpAP, composed of the hexameric AAA+ unfoldase ClpA and the barrel-shaped serine peptidase ClpP, the N-terminal domains of ClpA and the adaptor protein ClpS play additional regulatory roles in choosing which proteins in the cell are degraded.

Adaptor proteins have been particularly well-studied with respect to their role in affecting their partner AAA+ enzyme. The adaptor SspB for ClpXP is one of the best studied systems. An SspB dimer functions by binding to both to the ClpX N-terminal domains and to the substrate degradation tag. By binding both the substrate and the enzyme, SspB serves as a tether, increasing the local concentration of the substrate as well as the observed rate of processing (Levchenko et al. 2000; Wah et al. 2002; Dougan et al. 2003; Levchenko et al. 2003; Wah et al. 2003). This system for enhancing degradation of specific proteins by adaptors is an elegant way for the cell to prioritize substrates.

The adaptor ClpS for ClpAP, however, does not function like a simple tethering adaptor. ClpS can enhance the degradation of N-end-rule substrates by ClpAP. The N-end rule relates the half-life of a protein to its N-terminal residue, in which proteins starting with large, bulky residues are degraded the fastest (Varshavsky 1995). This degradation tag, which is comprised of a destabilizing N-terminal residue, is denoted the N-degron. Proteins possessing an N-degron are recognized by ClpAP for degradation, and ClpS enhances the rate of proteolysis (Erbse et al. 2006; Wang et al. 2007). ClpS itself also has a recognition site for N-degrons (Wang et al. 2008b). Thus, the initial assumption was that ClpS could tether N-end-rule proteins to ClpAP for degradation like SspB. Previous work indicated that a variant of ClpS that lacks its first 17 residues, but can still bind to both ClpA and to an N-degron, fails to deliver an N-end-rule protein to ClpAP (Hou et al. 2008). Therefore, the delivery of N-end-rule proteins by ClpS to ClpA cannot be explained by tethering alone. Although ClpS is indeed likely increasing the local concentration of N-end-rule substrate to ClpAP, this recruitment is not sufficient for increasing the rate of substrate degradation. ClpS must use another mechanism to transfer the substrate to ClpAP for processing (Hou et al. 2008).

Because the first seventeen residues of ClpS appeared to greatly alter the behavior of the adaptor protein, we decided to investigate its properties more thoroughly. In *E.*

coli, the first 24 residues of ClpS are disordered in ClpS-ClpA N-domain co-crystal structures (Guo et al. 2002a; Zeth et al. 2002). Because of the lack of structure, it is believed that the N-terminal residues of ClpS are flexible. The remainder of the ClpS protein forms a folded domain, and this folded core possesses the primary ClpA binding interface as well as the N-degron binding pocket. Therefore, the core of ClpS bears all the components to allow ClpS to function as a tethering adaptor. Interestingly, the N-terminal extension of ClpS is not conserved in sequence or in length among ClpS orthologs in bacteria. It is therefore surprising that the N-terminal extension plays such a critical role in ClpS function.

ClpS binds to the N-terminal domains of ClpA, and binding of wild-type ClpS to ClpA causes a suppression in ClpAP ATPase rates (Hinnerwisch et al. 2005b; Hou et al. 2008). Because the N-terminal domains and the ATP-binding sites are separated in the ClpA molecule, we proposed that ClpS binding causes an allosteric change in ClpA. We believe that the conformational change resultant from ClpS binding alters the substrate preference in ClpAP. Our previous work showed that N-terminally truncated ClpS molecules did not suppress ClpAP ATPase rates like wild-type ClpS. Therefore, the data suggested that the truncated ClpS variants do not allosterically affect ClpAP and thus do not cause a modulation in ClpAP substrate specificity. Furthermore, full-length variants of ClpS with different N-terminal sequences behaved like wild-type ClpS in multiple assays, showing that the function is not dependent on a strict sequence.

In this study, the nature of the N-terminal residues of ClpS is further explored. A series of truncations showed a dramatic length dependence of the *E. coli* ClpS N-terminal extension in which adaptors that are too short inhibit the delivery of N-end-rule substrates to ClpAP. Although previous experiments showed that the sequence of the N-terminal extension lacks strict requirements, we have created and tested new ClpS variants in which alteration of the N-terminal residues appears to affect the ability to deliver N-end-rule substrates. Lastly, we determined that the region of ClpA interaction by the ClpS N-terminus is near the central pore of ClpA. These studies expand upon our knowledge of how the ClpS adaptor protein operates, furthering our understanding of the diverse nature of adaptor function.

RESULTS

A minimal length of the ClpS N-terminal extension is required for N-end-rule delivery

Our previous work demonstrated that unlike wild-type ClpS, ClpS lacking its first 16, 17, 18, or 19 residues was no longer able to deliver N-end-rule substrates for degradation by ClpAP (Hou et al. 2008). To gain a better sense of how ClpS interacts with ClpAP, I wished to determine if there was a gradual decline in adaptor activity with shorter ClpS variants or whether there was a specific length of the N-terminal extension that was required for ClpS's delivery function. A series of deletion constructs was created in which one residue was removed at a time from wild-type ClpS. Figure 3.1a shows the first 17 residues of ClpS; in wild-type ClpS, the initiator methionine is removed upon expression. For each ClpS deletion construct, the initiator methionine was designed to be retained to prevent artifacts of exposing N-degrons on ClpS itself. After purification, the integrity of every ClpS protein was verified by mass spectrometry. Each ClpS deletion retained its initiator methionine with two exceptions. A portion of the methionines were cleaved from $ClpS^{M4\text{-}106}$ due to the threonine at the following position. This processing was not unexpected because initial methionines are cleaved when the second residue is small (Hirel et al. 1989). Shockingly, a large portion of the methionines were cleaved from $ClpS^{M7\text{-}106}$, which contains a tryptophan following the methionine. According to the established rules of methionine aminopeptidase, a methionine should not be cleaved when the second residue is as large and hydrophobic as tryptophan (Hirel et al. 1989). A repeat purification of this protein yielded the same partial processing. It has not been determined why this ClpS construct is partially processed, and more investigations to address this issue are being considered.

We tested the series of ClpS truncations in several assays to determine the effect of the deletions on function. It was first necessary to ensure that all of the ClpS variants were correctly folded. We expected that all of the ClpS truncated variants should be able to inhibit degradation of GFP-ssrA by ClpAP because these adaptors all had a longer N-terminal extension than the minimal length established previously as needed for inhibition of degradation of ssrA substrates (Hou et al. 2008). Figure 3.1b shows that these variants all inhibited GFP-ssrA degradation as expected, and these data also

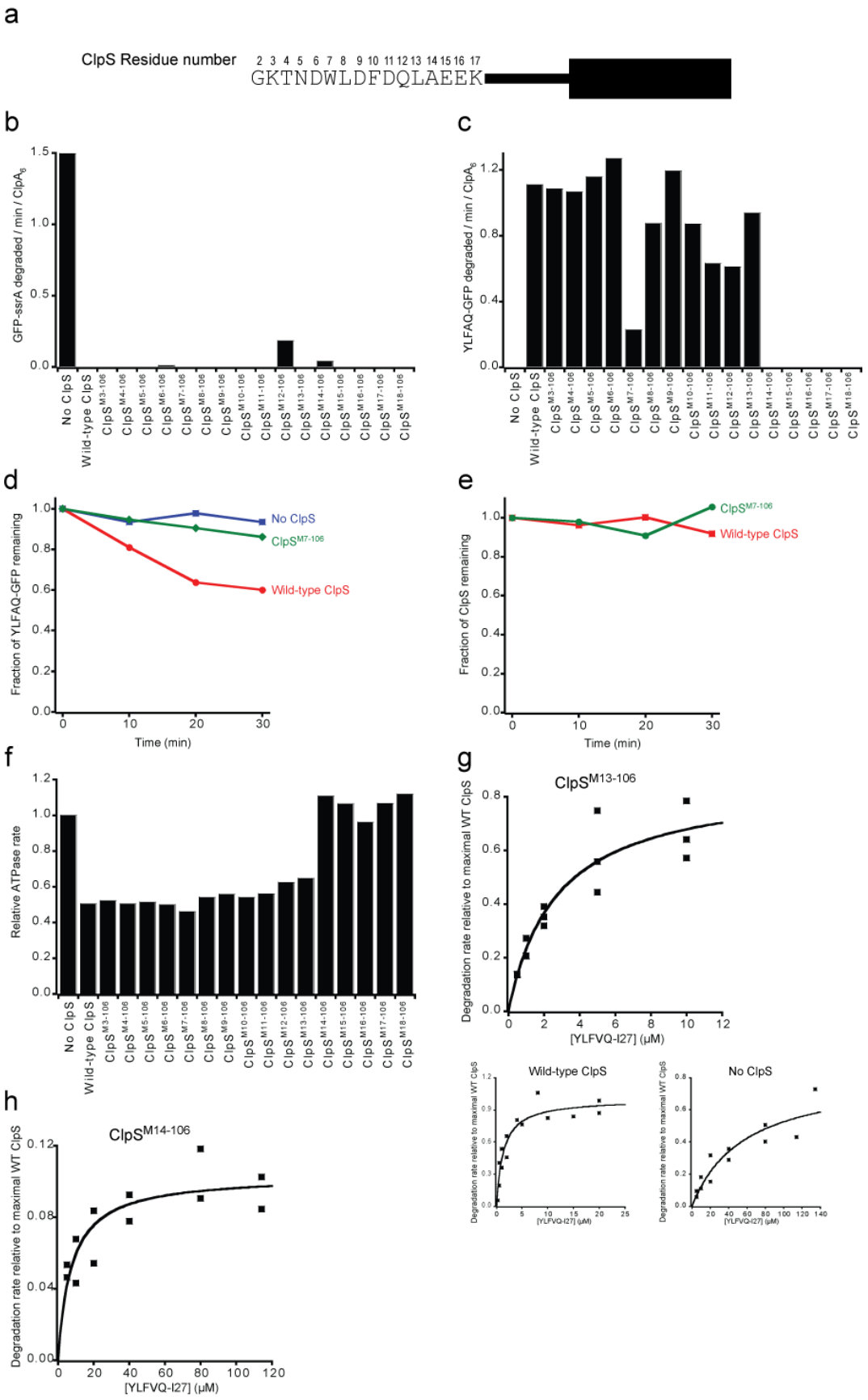
indicate that the variants were correctly folded and interacted with ClpA in order to exhibit this activity. The series of ClpS deletion constructs was then tested for the ability to stimulate degradation of an N-end-rule protein. With the exception of $ClpS^{M7-106}$, there appears to be a definitive drop in activity between $ClpS^{M13-106}$ and $ClpS^{M14-108}$ (Fig. 3.1c). Considering the presence of the initiator methionines, the length breakpoint in ClpS activity lies between residues 12 and 13. Thus, the minimum functional ClpS variant begins at the position of residue 12 in wild-type ClpS. $ClpS^{M7-106}$ behaved aberrantly with respect to N-end-rule stimulation, and this lack of stimulation is likely to due the N-terminal exposure of tryptophan in the majority of those ClpS proteins (Fig. 3.1c,d). Although there was no evidence for degradation of $ClpS^{M7-106}$ (Fig. 3.1e), this ClpS construct may be competing in *trans* with substrate for N-degron binding sites.

The series of ClpS truncation variants was also tested for ability to suppress ClpAP ATPase rates. All of the variants longer than $ClpS^{M14-106}$ behaved similarly to wild-type ClpS in suppressing the rate of ATP hydrolysis (Fig. 3.1f). In contrast, shorter ClpS variants failed to suppress ATPase rates. Short ClpS variants such as $ClpS^{M18-106}$ have already been reported not to lower ClpAP ATPase activity and to even slightly raise the ATPase rate (Hou et al. 2008). Therefore, the sharp cut-off in activity between residues 12 and 13 on ClpS applies to ATPase suppression as well as N-end-rule substrate delivery.

Because a sharp change in activity was observed between variants instead of a gradual decline, the ClpS constructs on either side of this line of activity, $ClpS^{M13-106}$ and $ClpS^{M14-106}$, were further studied. Degradation rates at different N-end-rule substrate concentrations were determined with a fixed concentration of ClpAP and ClpS mutant. The values for K_M (2.7 μM) and relative (to wild-type) V_{max} (0.86) for $ClpS^{M13-106}$ are similar to those of wild-type ClpS (1.3 μM K_M, 1.00 relative V_{max}), reinforcing the trend that variants longer than $ClpS^{M13-106}$ confer similar functional results (Fig. 3.1g). In contrast, $ClpS^{M14-106}$ had a slightly greater K_M (7.4 μM) as well as a much lower relative V_{max} (0.10) compared to $ClpS^{M13-106}$ and wild-type ClpS. In fact, the relative V_{max} of $ClpS^{M14-106}$ is lower than that of ClpAP alone (0.83), suggesting that at high substrate concentrations, ClpAP is more impaired in degrading an N-end-rule substrate with $ClpS^{M14-106}$ than with no ClpS at all (Fig. 3.1g,h). Because the K_M values of $ClpS^{M13-106}$

and ClpS$^{M14-108}$ are similar, one can infer that both ClpS variants bind N-end-rule substrates similarly. It was previously shown that ClpS with truncations in the N-terminal extension can still bind N-degrons (Hou et al. 2008). The main difference between the ClpS variants on either side of the activity divide is the maximal rate. The much lower relative V_{max} of ClpS$^{M14-108}$ suggests a defective hand-off of the substrate from the adaptor to ClpA. These two ClpS proteins differ by only one residue in the N-terminal extension, emphasizing the importance of this region's role in adaptor function.

Figure 3.1. A minimal length of the ClpS N-terminus is required for N-end-rule substrate delivery.
(a) The wild-type sequence of the ClpS N-terminal extension is shown with residue positions numbered. (b) Degradation of GFP-ssrA by ClpAP with series of ClpS N-terminal truncations. Degradation rates less than zero were due to slight upward drift of the fluorescence signal over time, and these negative values were normalized to "0." (c) Degradation of YLFAQ-GFP by ClpAP with series of ClpS N-terminal truncations. Negative degradation rates were normalized as described above. (d) Degradation of YLFAQ-GFP by ClpAP showing inability of ClpS^{M7-106} to greatly stimulate degradation. Samples of the degradation reaction were taken at 10-minute time intervals and run on SDS-PAGE. The substrate band was then quantified. (e) Wild-type ClpS or ClpS^{M7-106} adaptor was not degraded during the degradation assay of YLFAQ-GFP in (d). (f) ClpAP ATPase rates were observed with the series of ClpS N-terminal truncations. (g) (Top) Graph showing degradation of YLFVQ-I27 at different concentrations with ClpS$^{M13-106}$. Data from three independent experiments are shown with the best-fit Michaelis-Menton curve. The degradation rates were normalized to the V_{max} of wild-type ClpS to account for variation between substrate protein preparations. The K_M for ClpS$^{M13-106}$ is 2.7 μM, and the relative V_{max} is 0.9. (Bottom left) Graph showing degradation of YLFVQ-I27 at different concentrations with wild-type ClpS. The K_M of wild-type ClpS is 1.3 μM, and the relative V_{max} is 1.0. (Bottom right) Substrate concentration-dependent experiment showing degradation of YLFVQ-I27 at different concentrations without ClpS. The K_M without ClpS is 57 μM, and the relative V_{max} is 0.8. (h) Graph showing degradation of YLFVQ-I27 at different concentrations with ClpS$^{M14-106}$. Data from two independent experiments are shown. The K_M for ClpS$^{M13-106}$ is 7.4 μM, and the relative V_{max} is 0.1.


a
ClpS Residue number
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
GKTNDWLDFDQLAEEK
b
GFP-ssrA degraded / min / ClpA6
No ClpS
Wild-type ClpS
ClpSM3-106
ClpSM4-106
ClpSM5-106
ClpSM6-106
ClpSM7-106
ClpSM8-106
ClpSM9-106
ClpSM10-106
ClpSM11-106
ClpSM12-106
ClpSM13-106
ClpSM14-106
ClpSM15-106
ClpSM16-106
ClpSM17-106
ClpSM18-106
c
YLFAQ-GFP degraded / min / ClpA6
d
Fraction of YLFAQ-GFP remaining
No ClpS
ClpSM7-106
Wild-type ClpS
Time (min)
e
Fraction of ClpS remaining
ClpSM7-106
Wild-type ClpS
Time (min)
f
Relative ATPase rate
g
ClpSM13-106
Degradation rate relative to maximal WT ClpS
[YLFVQ-I27] (μM)
Wild-type ClpS
No ClpS
h
ClpSM14-106
Degradation rate relative to maximal WT ClpS
[YLFVQ-I27] (μM)

Altering the sequence of a shortened ClpS N-terminus affects function

Prior mutagenesis of the full-length ClpS N-terminus suggested that there was no strict sequence requirement for the first 17 amino acids of *E. coli* ClpS. The previous mutagenesis experiments were conducted in full-length ClpS in which either blocks of four contiguous residues were changed to alanine, or the entire first 17-residue segment was mutated to an assortment of small and positively charged residues (Hou et al. 2008). All of these full-length ClpS variants behaved like wild-type ClpS, and this lack of sequence specificity corroborated the lack of sequence and length conservation of the N-terminal extension. Although we still believe that there is no absolute sequence requirement, we reasoned that the non-mutated regions of the N-terminal extensions, chiefly in the alanine block-replacement experiments, could have compensated for the mutated residues given the extension's expected flexibility. The previous sequence variant in which we generated a wholesale change in the first 17 residues might have been sufficiently complex in sequence to mimic wild-type ClpS. Our current data reveal a length-dependence in the N-terminal extension of ClpS. We therefore decided to revisit mutagenesis of this region, but in the context of a short ClpS variant such that other N-terminal segments could not compensate.

ClpS variants starting at position 10 were created in which a string of residues is mutated to alanine, to glycine, or to proline (Fig. 3.2a). By examining the substrate-concentration dependence on the rate of degradation of an N-end-rule substrate by ClpAP, we observe that 8Ala-ClpS$^{18\text{-}106}$ (Alanine-ClpS) possesses a similar but slightly higher K_M (5.1 μM) to wild-type ClpS but a lower relative maximal rate (0.29) (Fig. 3.2b). Conversely, 8Gly-ClpS$^{18\text{-}106}$ (Glycine-ClpS) demonstrated a relative V_{max} (1.06) similar to wild-type ClpS but had a substantially higher K_M (155 μM) (Fig. 3.2c). Ala-7Pro-ClpS$^{18\text{-}106}$ (Proline-ClpS) is the most impaired variant compared to wild-type ClpS, yielding a low relative V_{max} (0.25) and a high K_M (109 μM) (Fig. 3.2d).

These three ClpS variants were tested in other functional assays. All three ClpS variants inhibited GFP-ssrA degradation by ClpAP as expected by their length, although Glycine-ClpS may be slightly less functional than the other variants (Fig. 3.2e). Furthermore, Glycine and Proline variants of ClpS were not as functional as wild-type and Alanine-ClpS in an assay to test for inhibition of ClpAP ATP hydrolysis rates,

displaying values mimicking the level observed with no ClpS or versions of ClpS starting C-terminal to the fourteenth residue (Fig. 3.2f).

Two control experiments were conducted to ensure that altering the N-terminal backbone did not adversely affect the core domain of ClpS, especially given the inability of Glycine-ClpS and Proline-ClpS to inhibit ClpAP ATPase rates. Fluorescence anisotropy indicated that all three ClpS variants were able to bind to a peptide bearing an N-degron like wild-type ClpS (Fig. 3.2g). Additionally, all three ClpS variants inhibited ClpA autodegradation like wild-type ClpS (Fig. 3.2h). Taken together with the GFP-ssrA inhibition data, these control experiments demonstrate that the core of ClpS in these amino acid variants is still functional and can bind both ClpA and to N-degrons.

Although these three ClpS variants should theoretically be long enough to function like wild-type ClpS and $ClpS^{M13\text{-}106}$, the Glycine and Proline variants are clearly defective in the delivery of N-end-rule proteins and in the suppression of enzyme ATPase rates. It is also unclear why the Alanine variant has a slightly lower V_{max} than wild-type ClpS. The functional debilitation of these amino acid variants has uncovered a sequence dependence not previously observed. It is possible that the use of short ClpS variants eliminated the possibility of functional redundancy of other N-terminal extension segments. However, these relatively severe variants might elucidate a requirement for variation in amino acid composition in the N-terminal extension rather than a sequence strongly biased for one residue.

Due to the nature of our mutations in which we alter a string of residues to alanine, glycine, or proline, we are assuredly altering the backbone flexibility in addition to the side chains. We expected that the alanines would have little overall effect on the backbone but that the glycines and prolines would greatly increase and decrease flexibility, respectively. Even though it is unclear whether the impairment of these ClpS variants is due to changing the side-chains or the backbone, we clearly observe that mutating the amino acids of a minimal-length ClpS variant decreases the ability to deliver an N-end-rule substrate for degradation by ClpAP.

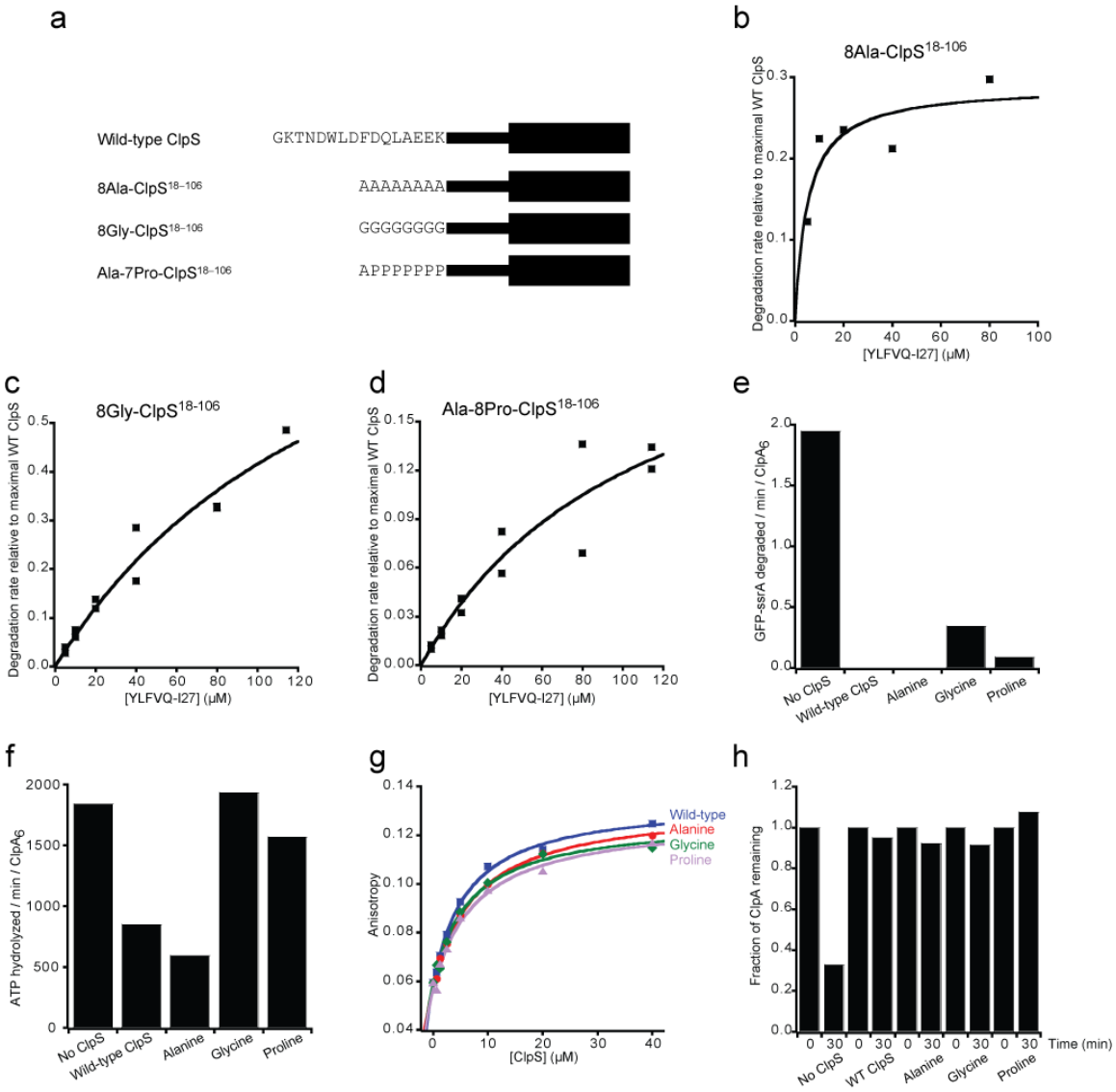

a
Wild-type ClpS
GKTNDWLDFDQLAEEK
8Ala-ClpS18–106
AAAAAAAA
8Gly-ClpS18–106
GGGGGGGG
Ala-7Pro-ClpS18–106
APPPPPPP
b
8Ala-ClpS18-106
Degradation rate relative to maximal WT ClpS
[YLFVQ-I27] (μM)
c
8Gly-ClpS18-106
Degradation rate relative to maximal WT ClpS
[YLFVQ-I27] (μM)
d
Ala-8Pro-ClpS18-106
Degradation rate relative to maximal WT ClpS
[YLFVQ-I27] (μM)
e
GFP-ssrA degraded / min / ClpA6
No ClpS
Wild-type ClpS
Alanine
Glycine
Proline
f
ATP hydrolyzed / min / ClpA6
No ClpS
Wild-type ClpS
Alanine
Glycine
Proline
g
Anisotropy
[ClpS] (μM)
Wild-type
Alanine
Glycine
Proline
h
Fraction of ClpA remaining
Time (min)
No ClpS
WT ClpS
Alanine
Glycine
Proline

Figure 3.2. ClpS requires a normal chemical composition for function.
(a) Depiction of ClpS constructs used in these experiments. (b) Graph showing degradation of YLFVQ-I27 at different concentrations with 8Ala-ClpS$^{18-106}$. Data from independent experiments were variable, but a representative plot is shown. The K_M for 8Ala-ClpS$^{18-106}$ in this experiment is 5.1 μM, and the relative V_{max} is 0.3. (c) Graph showing degradation of YLFVQ-I27 at different concentrations with 8Gly-ClpS$^{18-106}$. Data from two independent experiments are shown. The K_M for 8Gly-ClpS$^{18-106}$ is 155 μM, and the relative V_{max} is 1.1. (d) Graph showing degradation of YLFVQ-I27 at different concentrations with Ala-7Pro-ClpS$^{18-106}$. Data from two independent experiments are shown. The K_M for Ala-7Pro-ClpS$^{18-106}$ is 109 μM, and the relative V_{max} is 0.3. (e) Degradation of GFP-ssrA by ClpAP with sequence variants of ClpS. All constructs can inhibit GFP-ssrA degradation. Negative degradation rates were normalized as described in Figure 3.1b. (f) ATP hydrolysis rates of ClpAP with ClpS sequence variants. Glycine-ClpS and Proline-ClpS were unable to suppress rates like wild-type ClpS and the Alanine-ClpS. (g) Binding of a peptide with an N-degron to ClpS sequence variants. All variants bound the N-degron like wild-type ClpS with a K_D of 6-7 μM. (h) Measurement of the ability of ClpS variants to inhibit ClpA autodegradation. Two time points from N-end-rule substrate reactions with the appropriate ClpS molecule were taken and run on SDS-PAGE. The Coomassie-stained ClpA band was quantified.

The N-terminal extension of ClpS interacts with the pore of ClpA

Using ATPase suppression as a read-out, it is clear that the N-terminal extension of ClpS has a direct effect on ClpAP in the absence of substrate (Fig. 3.1f). Therefore, the N-terminal extension of the adaptor is most likely making a contact with the enzyme. The main ClpS-ClpA binding interface has already been established by x-ray crystallography to be between the core folded domain of ClpS and the N-terminal domain of ClpA (Guo et al. 2002a; Zeth et al. 2002; Xia et al. 2004). The ClpS N-terminal extension is not visible in existing crystal structures, and it is currently unknown whether ClpS makes contacts with the other ClpA domains. We speculate that the extension can bind to full-length, hexameric ClpA, perhaps transiently, as part of ClpS's function in altering ClpAP's substrate preference.

To understand the role of the N-terminal extension of ClpS, we searched for the location on ClpA where it might interact. We employed a protein cleavage reagent, FeBABE, attached to the N-terminal extension of ClpS. Because the region around residues 12 and 13 are critical for the role of ClpS in delivering N-end-rule substrates, we

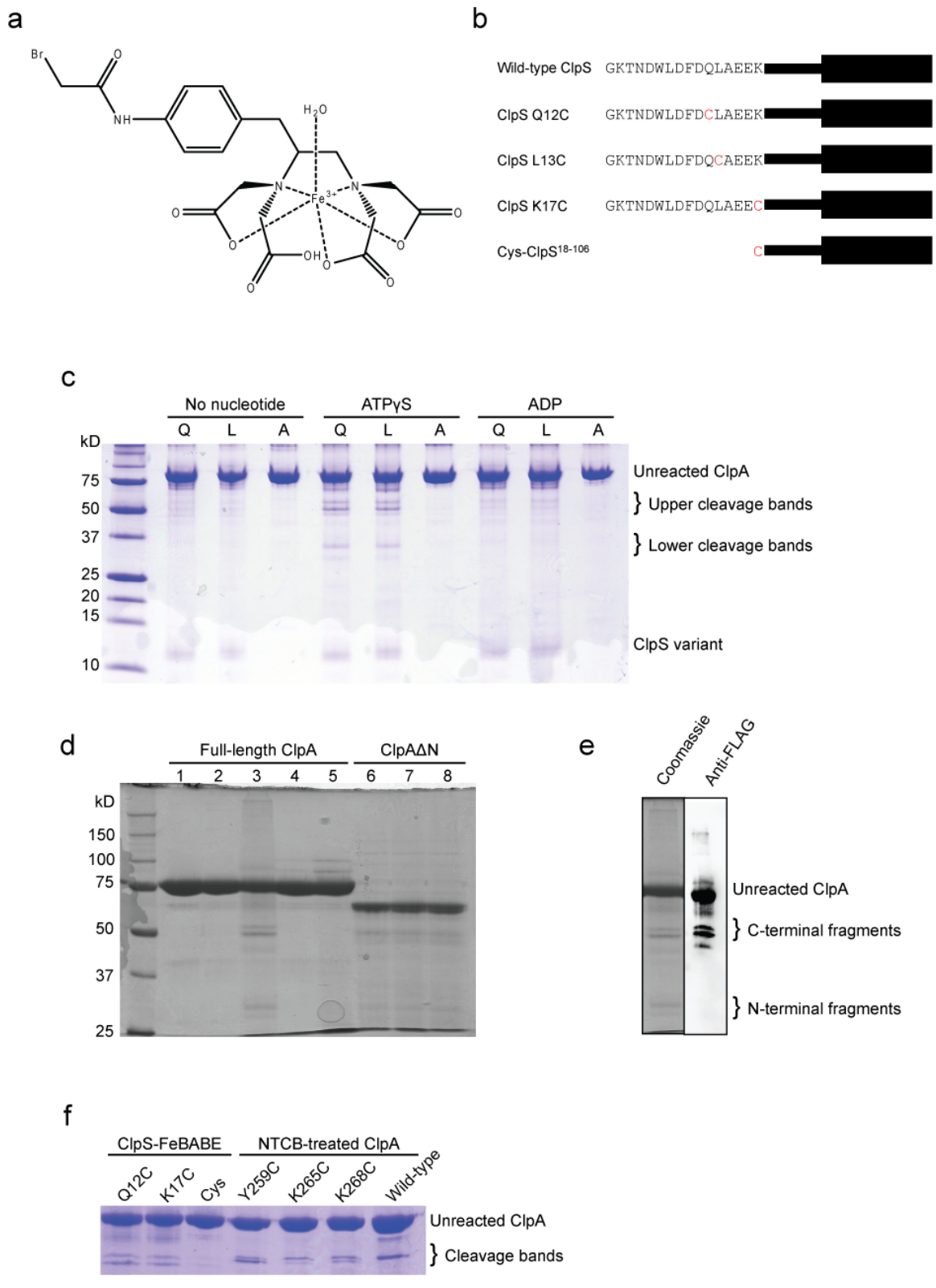
engineered individual cysteines on each of these residues on which to fuse FeBABE (Fig. 3.3a,b). The FeBABE reagent contains a chelated Fe^{3+} atom that when activated by ascorbic acid and hydrogen peroxide, will generate radicals that can cleave protein side chains. The FeBABE reagent also includes a sulfhydryl-reactive group that can covalently bind to cysteine side chains on ClpS. The activated FeBABE attached to the N-terminal extension of ClpS will cleave the region of ClpA in closest proximity, and the cleavage site can be mapped. The FeBABE reagent did not greatly inhibit ClpS function and therefore binding to ClpA (data not shown).

The FeBABE cleavage reaction yielded four primary bands, two bands of approximately 50 kilodaltons and two bands of approximately 29 kilodaltons. Both ClpS Q12C and ClpS L13C gave the same pattern, and ATPγS was required for efficient cleavage (Fig. 3.3c). The sum of the molecular weights of one small band and one large band approximately equals the molecular weight of the 84 kilodalton ClpA monomer, strongly suggesting that each pair of large and small bands resulted from a single cleavage. This cleavage is specific to the ClpS-bound FeBABE reaction because competition with an unlabeled ClpS molecule or the use of ClpA lacking an N-terminal domain for ClpS binding eliminated FeBABE-cleavage of ClpA (Fig. 3.3d). To determine the location of the cuts, it was necessary to determine which band corresponded to which end of ClpA. To address this question, we replicated the cleavage reaction with a C-terminally FLAG-tagged ClpA. The C-terminal FLAG tag does not inhibit ClpA function (data not shown). A western blot against the FLAG tag indicated that the larger cleavage products corresponded to the C-terminal protein fragment, and the smaller bands corresponded to the N-terminal fragments (Fig. 3.3e).

To estimate the location of the cleavage sites, we compared the migration of the resultant bands by SDS-PAGE compared to the molecular size standards (Fig. 3.3f). Our initial estimation of the cleavage site was around residue 260 of ClpA. Based on the crystal structure of the ClpA monomer, the pore of the D1 ATPase domain lies in this region (Guo et al. 2002b). To refine our estimation on the cleavage site and the hence the site of interaction by the ClpS N-terminal extension, we employed variants of ClpA with cysteines engineered in and around the D1 pore. The reagent NTCB was used to chemically cleave ClpA immediately before cysteines. We analyzed the products of

NTCB cleavage on a gel and used the fragments generated from the cleavage before known cysteines in the pore as more precise size standards for the FeBABE reactions. The new estimation indicated that one cleavage site is approximately between ClpA residues 243 and 259. The other cleavage site is C-terminal to residue 272. The use of two other ClpS variants on which to position the FeBABE reagent, ClpS K17C and Cys-ClpS$^{18\text{-}106}$, did not yield different band sizes but yielded less efficient cleavage. The inefficiency of reaction with Cys-ClpS$^{18\text{-}106}$ supports the need for a minimal length of the N-terminal extension to make a proper contact with ClpA.

Figure 3.3. The ClpS N-terminus interacts with ClpA.
(a) Depiction of the FeBABE reagent used for artificial cleavage of ClpA when attached to ClpS. The sulfur group on the cysteine of the bait protein covalently binds to the molecule. The EDTA-chelated iron atom reacts with nearby residues when triggered by ascorbic acid and hydrogen peroxide treatment. (b) Depiction of ClpS variants used in the FeBABE studies. The ClpS cysteine variants were constructed in a C73V C101S background to avoid binding of FeBABE at these cysteines. (c) Cleavage reaction of ClpA by FeBABE with ClpS Q12C (Q), ClpS L13C (L), or with ClpA only. The reaction was most efficient with ATPγS as nucleotide to hexamerize ClpA. (d) Control experiments showing that the FeBABE cleavage is ClpS-specific. Lane 1 – ClpA only, no reaction; Lane 2 – ClpA only, with reaction; Lane 3 – ClpA with ClpS Q12C-FeBABE and reaction; Lane 4 – ClpA with ClpS Q12C (no bound FeBABE), and reaction; Lane 5 – ClpA with ClpS Q12C-FeBABE, excess ClpS Q12C (no bound FeBABE), and reaction; Lane 6 – ClpAΔN, no reaction; Lane 7 – ClpAΔN, with reaction; Lane 8 – ClpAΔN with ClpS Q12C-FeBABE, with reaction. Although some smaller molecular weight bands are seen in lanes with ClpAΔN, because they exist in the lanes without the FeBABE reagent, they were deemed contaminants. "No reaction" refers to lack of ascorbic acid and hydrogen peroxide addition. The full-length ClpA molecules were C-terminally FLAG-tagged in this experiment. (e) (Left) Coomassie blot of FeBABE reaction taken from Lane 3 in (d). (Right) Anti-FLAG western blot showing that the upper cleavage bands correspond to the C-terminal ClpA fragments. (f) Comparison of FeBABE cleavage reactions when bound to ClpS Q12C, ClpS K17C, or Cys-ClpS$^{18\text{-}106}$. The larger cleavage bands are shown next to ClpA fragments generated from NTCB digestions of ClpA cysteines in the pore.


a
b
Wild-type ClpS GKTNDWLDFDQLAEEK
ClpS Q12C GKTNDWLDFDCLAEEK
ClpS L13C GKTNDWLDFDQCAEEK
ClpS K17C GKTNDWLDFDQLAEEC
Cys-ClpS18-106 C
c
No nucleotide
ATPγS
ADP
Q L A Q L A Q L A
kD
75
50
37
25
20
15
10
Unreacted ClpA
Upper cleavage bands
Lower cleavage bands
ClpS variant
d
Full-length ClpA
ClpAΔN
1 2 3 4 5 6 7 8
kD
150
100
75
50
37
25
e
Coomassie
Anti-FLAG
Unreacted ClpA
C-terminal fragments
N-terminal fragments
f
ClpS-FeBABE
NTCB-treated ClpA
Q12C
K17C
Cys
Y259C
K265C
K268C
Wild-type
Unreacted ClpA
Cleavage bands

The exact region of N-terminal ClpS-ClpA contact is still unidentified

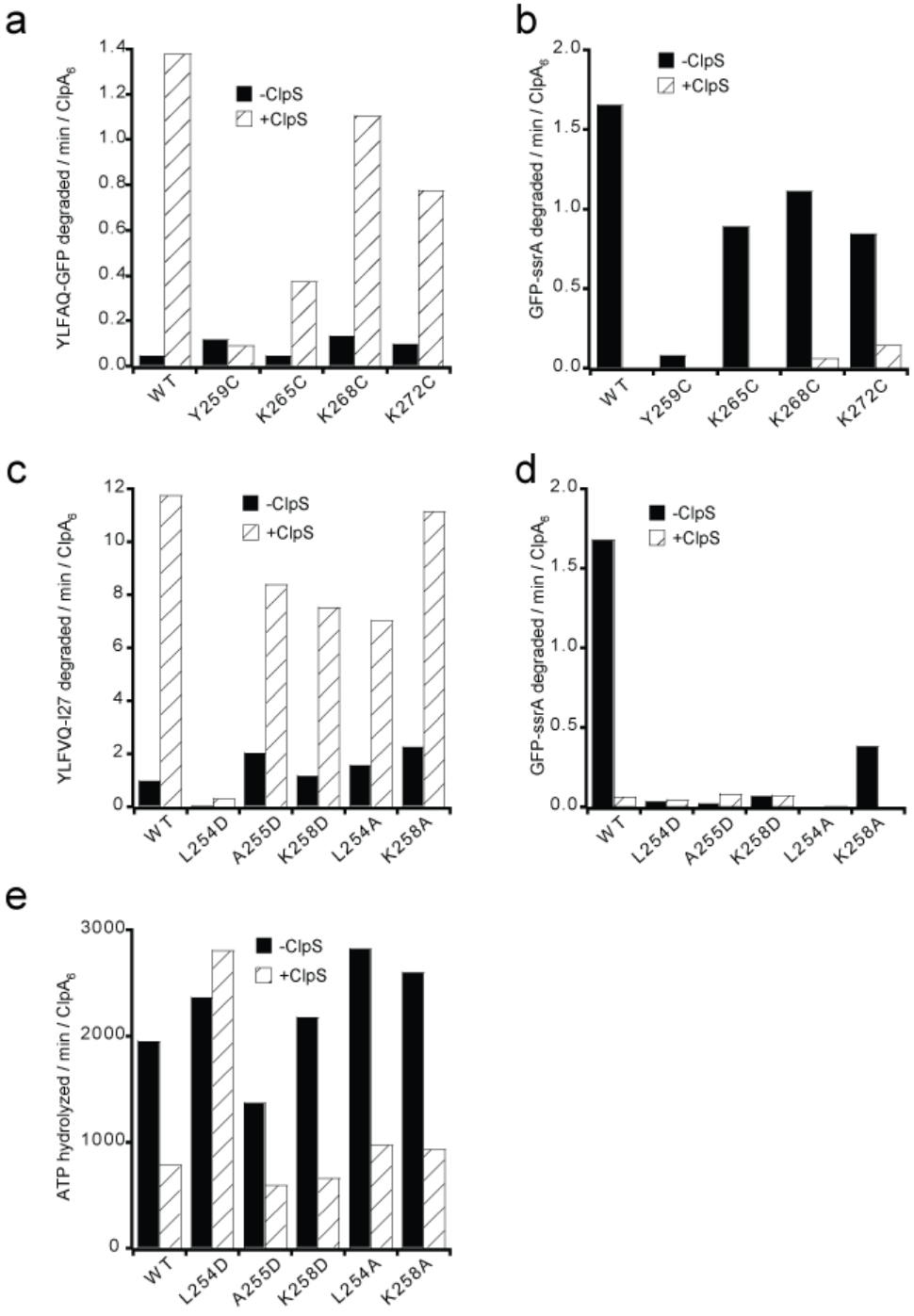
We wished to test the pore region of ClpA to determine if the N-terminal extension of ClpS indeed makes a contact. We made several additional ClpA pore mutations and sought a mutant that would allow the degradation of an N-end-rule substrate by ClpAP alone but be defective with the addition of ClpS. Our assumption was that a defect in ClpS-mediated delivery would indicate that the mutated residue on ClpA was necessary for the interaction of the ClpS N-terminal extension. Two sets of ClpA mutants were utilized. One set included singly mutated residues in ClpA, Y259C, K265C, K268C, and K272C, most of which were used earlier as NTCB-generated size standards. These mutations are in solvent accessible regions on the N-terminal surface of the D1 domain near the pore except for ClpA Y259C, which is directly in the D1 pore. With the exception of ClpA Y259C, all of these ClpA mutants were stimulated by ClpS with respect to N-degron protein degradation (Fig.3.4a). The rate of substrate delivery with the Y259C mutant plus ClpS is not as low as wild-type ClpA alone, suggesting that this difference in degradation is not substantial. Furthermore, Y259C is the only of these ClpA mutants that failed to degrade of GFP-ssrA (Fig. 3.4b). Although it is possible that this residue plays unique roles in the degradation of an ssrA-tagged substrate versus an N-end-rule substrate, we believe that a mutation at Y259 generates a generally inactive ClpA protein independent of ClpS.

The second set of ClpA mutations were made in conserved residues flanking the ClpA D1 pore. We anticipated that the reason for conservation might be because of the necessity of these amino acids for ClpS interaction, although these residues are likely also important for general pore function. Unfortunately, all of these ClpA mutants were also stimulated by ClpS with respect to N-end-rule protein delivery with the exception of L254D, which appeared to be generally inactive (Fig. 3.4c). Interestingly, most of these mutations yielded ClpA proteins defective for GFP-ssrA degradation (Fig. 3.4d). This result was not surprising because we targeted conserved residues likely to be important for the integrity of the D1 pore. Defects in GFP-ssrA degradation by some of these ClpA mutants have been previously reported (Hinnerwisch et al. 2005). It is uncertain why mutations in these ClpA pore residues would allow active degradation of an N-end-rule

substrate but not an ssrA-tagged substrate. Perhaps we have uncovered residues important uniquely for ssrA-like substrates, but this hypothesis requires further testing.

We also tested these ClpA variants for ATPase activity in the presence of ClpP. It is unclear why there is variability in the absolute rate of hydrolysis between variants, but with the exception of ClpA L254D, all of these ClpA proteins showed rates of ATP hydrolysis that could be suppressed by ClpS (Fig. 3.4e). Taken together, all of our ClpA mutants that were not globally defective in substrate processing were still able to be stimulated by ClpS for degradation of N-end-rule substrates.

Figure 3.4. Testing of the ClpA pore as the likely site of interaction with the ClpS N-terminal extension.
(a) Degradation of YLFAQ-GFP by ClpA cysteine mutants with and without ClpS. (b) Degradation of GFP-ssrA by ClpA cysteine mutants with and without ClpS. Negative degradation rates were normalized as described in Figure 3.1b. (c) Degradation of YLFVQ-I27 by conserved residue ClpA mutants with and without ClpS. For this experiment, 5 μM of substrate was used, differing from 1 μM used elsewhere in this figure, to elucidate differences between degradation rates with and without ClpS.
(d) Degradation of GFP-ssrA by conserved residue ClpA mutants with and without ClpS. Negative degradation rates were normalized as described in Figure 3.1b. (e) ATP hydrolysis rates of ClpA mutants with ClpP with and without ClpS.


a
YLFAQ-GFP degraded / min / ClpA6
-ClpS
+ClpS
WT
Y259C
K265C
K268C
K272C
b
GFP-ssrA degraded / min / ClpA6
c
YLFVQ-I27 degraded / min / ClpA6
L254D
A255D
K258D
L254A
K258A
d
e
ATP hydrolyzed / min / ClpA6

DISCUSSION

The elucidation of the complete role and function of the ClpS adaptor protein remains quite complex. This study uncovers more information about the role of the N-terminal extension, specifically with respect to its importance in the delivery of N-end-rule substrates to ClpAP. A striking length dependence in the N-terminal extension has been determined in which ClpS mutants that are too short are unable to deliver N-end-rule substrates. Our previous and current data suggest that the N-terminal extension is required to impose a conformational change upon the ClpA enzyme. This conformational change results in the substrate specificity switch that redirects ClpAP away from ssrA-tagged substrates and towards N-end-rule substrates (Dougan et al. 2002; Zeth et al. 2002). By this model, the N-terminal extension needs to be of a certain length to interact with the putative binding pocket on ClpA.

How do we explain the observation that a short version of ClpS, such as $ClpS^{M14-106}$, results in a lower degradation rate of an N-end-rule protein than with ClpAP alone? Similarly, why do increasing amounts of short ClpS inhibit the rate of N-end-rule degradation (Hou et al. 2008)? One hypothesis is that the ClpS core domain and/or the remaining portion of the N-terminal extension sterically hinders ClpA's ability to directly bind N-end-rule proteins. So, the short ClpS not only does not enhance degradation of an N-end-rule substrate due to the lack of a sufficient N-terminal extension for allosteric activation, but the remainder of the protein blocks ClpA's original ability to recognize N-degrons in the absence of adaptor. Furthermore, in this model, the ClpS folded domains might compete with ClpA for N-degrons, although this potential competition should not factor in saturating substrate conditions.

Another hypothesis for the observed suppression of degradation rates by $ClpS^{M14-106}$ is that two conformational changes in fact occur instead of one: the first change occurs when the ClpS core binds to the N-terminal domains, and the second change occurs when the N-terminal extensions of ClpS bind their respective pocket in ClpA. This distinction in conformational changes might explain why short ClpS constructs reproducibly elevate levels of ClpAP ATP hydrolysis rather than depress rates or keep rates completely unchanged (Fig. 3.1f) (Hou et al. 2008). By this model, the first conformational change exerted by ClpS binding to the ClpA N-terminal domain renders

ClpA unable to recognize ssrA-tags and N-degrons. The second conformational change, which is exerted only by binding of the ClpS N-terminal peptide, re-exposes a higher-affinity N-degron binding site. Thus, the overall effect of wild-type ClpS is the shift of specificity away from ssrA-tagged substrates towards N-degrons. It is currently unknown whether the change in ATPase rate has functional importance, such as shifting the ClpA motor into different "gears" to degrade different classes of substrates, or if changes in hydrolysis rates are simply a tolerable consequence of an enzymatic conformational change. It is important to note, however, that ClpS binding does not suppress ATPase rates to zero; ClpAP in the presence of ClpS must still be able to degrade N-end-rule substrates.

The data from Alanine-ClpS, Glycine-ClpS, and Proline-ClpS provide further insight into the nature of molecular contacts between the ClpS N-terminal extension and ClpA. Although the lack of sequence conservation among ClpS orthologs and results from previous mutagenesis both pointed against sequence-specific contacts, we now see an effect due to drastic alteration of the N-terminal sequence of ClpS. It is possible that sequence complexity, rather than a strict sequence, is necessary for binding of the N-terminal extension into its putative ClpA pocket. However, backbone dynamics may indeed play a role in the ability of these ClpS variants to activate ClpAP. The string of glycines in Glycine-ClpS might be too flexible to stably bind a discrete binding site, and the string of prolines in Proline-ClpS might form a configuration that is not recognized by ClpA.

Repeated proline residues can form helices in two configurations. Form I is a right-handed helix of *cis* residues with a rise of 1.9 Å per residue, and Form II is a left-handed helix of *trans* residues with a rise of 3.1 Å per residue (Cowan and McGavin 1955; Traub and Shmueli 1963). Therefore, assuming that the commencing alanine is unstructured and has a translational length of 3.6 Å, then if the alanine plus the 7 prolines – equivalent to residues 10 through 17 – formed a helix, the overall length could be approximately 16.9 Å (3.6 + 1.9 × 7) or 25.3 Å (3.6 + 3.1 × 7) (Huggins 1943). If fully extended, the distance from residue 12 – the start of the minimal ClpS adaptor – through residue 17 in wild-type ClpS would be 21.6 Å. It is unknown which type of

helix this proline chain could form, if any, but form I would yield a length insufficiently long for N-end-rule substrate delivery based on our deletion studies.

A survey of some ClpS orthologs from other species show that most N-terminal extensions are indeed of minimal length as determined by our studies with *E. coli* ClpS. However, it is intriguing why the N-terminus of the ClpS ortholog from *Helicobacter pylori*, for example, is shorter than the determined minimal length. Perhaps differences in *H. pylori* ClpA can compensate for the shorter ClpS N-terminus. Interestingly, the protein Aat, a leucyl/phenylalanyl-tRNA-protein transferase that is an important and conserved component of the N-end-rule pathway in bacteria (Shrader et al. 1993), is not annotated as being present in *H. pylori*. It will be useful to establish if the N-end rule as we understand it indeed operates in *H. pylori* to determine if ClpS functions as expected.

The FeBABE reagent has permitted the refinement of the interaction zone for the ClpS N-terminal extension. Even though the exact residue or group of residues on ClpA that interacts with the ClpS N-terminus remains elusive, we have narrowed down the region of interaction to near the D1 ring pore of ClpA. Finding the region of interaction is challenging because we expect the ClpS N-terminal extension to be flexible. The length of each residue in an extended polypeptide is 3.6 Å, thus the first 17 residues of ClpS if fully extended could sweep a sphere with a radius of 57.6 Å (Huggins 1943). Furthermore, because the ClpA N-terminal domain on which ClpS sits is thought to be mobile (Ishikawa et al. 2004), the overall spatial possibilities of the ClpS N-terminus with respect to the main ClpA barrel are tremendous. However, because we mapped a region of contact with the FeBABE reagent, which has a length of 12 Å, we can assume that the actual site of ClpS-ClpA contact might be only 3 residue-lengths away in distance. Given the expected flexibility of the ClpS N-terminal extension, our regional refinement greatly facilitates future continued exploration of the ClpS interaction site.

It is logical that the region of interaction between the ClpS N-terminal extension and ClpA has been localized to the pore of the ClpA D1 domain, given that all ATP-dependent substrates of ClpAP are thought to pass through the central pore. We speculate that the N-terminal extension of ClpS binds to a site on ClpA, which results in a conformational change that modifies the substrate preference of ClpAP. However, the role of the D1 AAA+ domain in tandem-AAA+-domain proteins such as ClpA is unclear.

Prior experiments suggest that the D1 ring plays a role in ClpA hexamerization whereas the D2 ring functions more in ATP hydrolysis (Singh and Maurizi 1994; Seol et al. 1995). The ATP-binding rings of single-ring AAA+ proteins such as ClpX correspond most closely to the D2 ring of ClpA, and it is uncertain if the D1 domain participates in ATP hydrolysis and active translocation or whether it serves other regulatory functions. In this case, the D1 domain may act as a link between the N-terminal domain and the D2 domain, relaying a conformational change resulting from ClpS binding the ClpA N-terminal domain to the D2 domain where the chief pulling motion may take place.

How does hand-off of the N-end-rule substrate to ClpAP occur? It was previously suggested that as the ClpS N-terminal extension reveals an enhanced binding pocket for N-degrons in ClpA, other residues in ClpS might also transiently interact with ClpA, destabilizing ClpS's N-degron binding pocket such that the substrate is released from the adaptor toward the ClpA pore (Wang et al. 2008b).

Another attractive and previously untested model for the function of ClpS in N-end-rule protein delivery is that ClpA engages the tantalizingly long N-terminal extensions of the adaptor as if it were a substrate, pulling the ClpS molecule in complex with an N-end-rule substrate in closer proximity to the processing pore after which the N-end-rule substrate is transferred to ClpA. There has been no evidence that ClpS itself gets degraded or even engaged by ClpAP; however, we speculated that ClpS's structural stability might prevent degradation by ClpA. Although we have not observed strong evidence that substrate presence influences enzyme ATPase rates, it is possible that the putative core stability of ClpS may explain the suppression in ATPase rates as ClpAP struggles to unfold the adaptor. ClpXP hydrolyzes ATP at a slower rate when degrading tightly folded proteins compared to unfolded ones (Kenniston et al. 2003). ClpAP may be operating in a similar mechanism with ClpS.

Support for the ClpS "pseudo-substrate" theory lies in evidence that ClpAP can degrade generally unfolded substrates, that ClpS is not degraded by ClpAP despite its long N-terminal extension, that even ClpS bearing an N-degron is not degraded in *trans* as a substrate (Fig. 3.1e), and that precedent exists for the proteasomal adaptor Rad23 being resistant to degradation due to its UBA2 stabilizing domain (Heessen et al. 2005). Incorporating data from this study, it is logical that a minimal length of the N-terminal

extension would be required for ClpA binding if the core of ClpS is anchored onto ClpA's N-terminal domains. Similarly, the ClpA pore might not favor interacting with a string of glycines, prolines, or even all alanines as present in the ClpS variants above. Studies suggest that proteasomal substrates generally have two requirements for degradation: proximity to the processing pore and possession of an unstructured protein region of complex sequence for initiation of degradation (Tian et al. 2005; Inobe and Matouschek 2008). These studies show that a variation in sequence, but not necessarily the identity of the residues themselves, influences whether or not an unstructured region can be captured by the proteasome. A report with ClpXP showed that a variant of GFP-ssrA in which the majority of the ssrA tag was mutated to glycine except for the critical C-terminal LAA was not degraded at concentrations below 50 μM in the conditions tested (Flynn et al. 2001). It was hypothesized that the glycine linker may be too flexible to allow proper recognition and processing of the LAA C-terminal motif by ClpXP. A similar phenomenon might apply to ClpAP in which even the string of eight alanines on the ClpS variant mentioned above is too low in complexity for ClpA to grip, thus resulting in a slightly lower V_{max} compared to wild-type ClpS. If the D2 ring of ClpA were in fact the principal center for substrate binding and pulling, a substrate's degradation tag would have to reach 40 Å into the ClpA's D1 pore to reach D2 (Ishikawa et al. 2004; Hinnerwisch et al. 2005a). As mentioned earlier, the first seventeen residues of ClpS are predicted to reach approximately 58 Å if fully extended, which would support this model.

All of these models require direct testing, but our data reveal more of the intricacies behind ClpS function. Understanding the mechanism for ClpS will enrich our knowledge for how proteolytic substrates are prioritized. Furthermore, ClpS is an example of the diverse adaptors that AAA+ proteins use to modulate their substrate preference.

MATERIALS AND METHODS

Proteins

ClpS truncations with the exception of $ClpS^{M13\text{-}106}$ and $ClpS^{M14\text{-}106}$ were created by amplifying the appropriate gene segment from genomic DNA and cloning into a

pET23b vector (Novagen). These plasmids were transformed into ER2556 (DE3) pLysS cells, and the cells were grown in LB with 0.4 mM IPTG added at OD_{600} 0.4-0.6 to induce protein expression. Cells were harvested and lysed, after which the cleared lysate was precipitated with 35% saturated ammonium sulfate. Precipitated proteins were resuspended, cleared of un-resuspended matter, and applied to a MonoQ anion exchange column followed by a Superdex 75 size exclusion column (GE Healthcare) equilibrated in 20 mM HEPES pH 7.5, 150 mM KCl, 10% glycerol, and 1 mM DTT.

The 8Ala-$ClpS^{18-106}$ (Alanine-ClpS), 8Gly-$ClpS^{18-106}$ (Glycine-ClpS), Ala-7Pro-$ClpS^{18-106}$ (Proline-ClpS), and Cys-$ClpS^{18-106}$ variants were cloned into a SUMO fusion vector bearing a His_6 tag preceding the N-terminal SUMO domain. Cells were grown similarly to above, but the protein was purified over Ni-NTA resin (Qiagen). The SUMO domain of the fusion was cleaved overnight with Ulp1 SUMO protease. The resultant cleavage mix was reapplied over Ni-NTA resin to eliminate full-length fusion protein as well as the isolated SUMO domains, yielding the appropriate ClpS protein. The ClpS proteins were then buffer exchanged into its final buffer (above).

$ClpS^{M13-106}$ and $ClpS^{M14-106}$ were also generated from SUMO cleavage and purification due to lack of expression when cloned with the other deletion panel.

Plasmids expressing ClpS Q12C C73V C101S, ClpS L13C C73V C101S, and ClpS K17C C73V C101S were generated by QuikChange mutagenesis (Stratagene). Proteins were purified similarly to the non-His-tagged ClpS proteins above.

ClpA variants were generated by QuikChange mutagenesis (Stratagene). Proteins were purified similarly to wild-type ClpA. Cleared cellular lysate was precipitated with 40% saturated ammonium sulfate. The resuspended protein was cleared and purified over an SP Sepharose Fast Flow cation exchange column. Peak fractions were pooled, and ammonium sulfate was added to 0.6 M concentration. The salted sample was then purified over a Phe-Source column. The peak fractions were dialyzed into ClpA reaction buffer (50 mM HEPES 7.5, 20 mM $MgCl_2$, 0.3 M NaCl, 10% glycerol, and 0.5 mM DTT).

ClpAΔN ($ClpA^{143-758}$) was cloned into and expressed from pET3a (Novagen) in ER2556 (DE3) pLysS. Proteins were purified similarly to the ClpA variants above

except the protein was purified over a MonoS followed by a MonoQ column. Peak fractions were dialyzed into ClpA dialysis buffer.

Unlabelled and ^{35}S-labelled YLFVQ-I27 protein were purified from the SUMO vector and purified as the ClpS variants indicated above. I27 is a stably-folded domain of the human titin protein. The complete peptide sequence preceding the I27 domain is: YLFVQMSHLA. A FLAG tag is C-terminal to the I27 domain.

Protein degradation

All degradation assays were conducted at 30 ºC in ClpA activity buffer with 0.1 μM $ClpA_6$, 0.27 μM $ClpP_{14}$, 4 mM ATP, 2.5 mM creatine phosphate, and 50 μg/mL creatine kinase. Where applicable, 1 μM ClpS was added. Except where indicated, most assays utilizing a GFP substrate were at a fixed substrate concentration of 1 μM. Most assays measuring GFP degradation were conducted in a PTI fluorimeter. Degradation of YLFVQ-I27 at different concentrations was monitored by removing reaction time points and measuring TCA soluble counts in a liquid scintillation counter. ClpA autodegradation was monitored by quantifying changes in band density on SDS-PAGE among samples taken over a time course.

ATP hydrolysis

ClpAP ATPase rates were monitored with similar conditions as the degradation assays above but with a coupled ATP hydrolysis assay (Kim et al. 2001).

Peptide binding

Binding of the peptide YLFVQY-H_6-C with a C-terminal fluorescein was measured by fluorescence anisotropy in a PTI fluorimeter at 30 ºC. 200 nM peptide was added to a range of concentrations of ClpS protein. No light scattering was evident with the Glycine ClpS variant, therefore all of the anisotropy values for all of the variants were calculated without correcting for light scattering.

FeBABE cutting assay

The assay with FeBABE (Fe(III) (s)-1-(*p*-Bromoacetamidobenzyl) ethylenediaminetetraacetic acid) (Pierce) was conducted according to the manufacturer's protocol with some changes. ClpS cysteine variants were buffer exchanged into a metal removal buffer (30 mM MOPS, 4 mM EDTA, pH 8.1) overnight at 4 ºC. The next day, the protein was buffer exchanged into conjugation buffer (30 mM MOPS, 100 mM NaCl, 1 mM EDTA, 5 % glycerol, pH 8.1) after which FeBABE Cutting Reagent resuspended in conjugation buffer was added for 1 hour at 37 ºC. The ClpS bound to FeBABE as well as ClpA was buffer exchanged into modified cutting buffer (50 mM MOPS, 300 mM NaCl, 10 mM $MgCl_2$, 10% glycerol, pH 8.1). ClpS-FeBABE and ClpA were mixed with 1 mM nucleotide and cutting buffer and incubated at room temperature for 30 minutes. The cutting reaction was initiated by the addition of 40 mM ascorbic acid and 10 mM EDTA immediately followed by 40 mM hydrogen peroxide and 10 mM EDTA. The reaction was quenched by adding electrophoresis sample buffer with 40% glycerol. The reactions were then run on SDS-PAGE.

Bands on Coomassie-stained protein gels were analyzed with ImageQuant TL software (GE Healthcare).

NTCB cleavage

ClpA size standards were created by cleavage with NTCB (2-Nitro-5-thiocyanobenzoate). The ClpA variants were incubated with 50 mM DTT at 37 ºC for approximately 10 minutes to completely reduce cysteines. The proteins were then buffer exchanged into 200 mM Tris-acetate pH 8, 1 mM EDTA, 5 M urea, and 0.1% SDS after which the proteins were incubated at 37 ºC for 20 minutes to react the cysteines with NTCB. Proteins were then immediately buffer exchanged into 200 mM Tris-acetate pH 9 and incubated at 45 ºC for 2 hours to cleave the protein.

CHAPTER FOUR

A proteomic screen yields candidates for ClpS-dependent ClpAP substrates

This work was done in collaboration with Elizabeth Oakes, Kevin Wang, and Judit Villén (laboratory of Steven Gygi, Harvard Medical School). E.O. assisted with initial mass spectrometric analysis, K.W. contributed some initial ClpAP trapping data, and J.V. conducted the mass spectrometry and peptide analysis for trapping using SILAC.

ABSTRACT

ClpAP is one of five ATP-dependent proteases in *Escherichia coli*, and the unique cellular role of each of these proteases is still being uncovered. Here, we used a proteomics screen to search for the substrates of ClpAP in an effort to elucidate its physiological function. Using a "trap" form of the ClpP peptidase, we isolated putative ClpAP substrates in the presence or absence of its adaptor protein ClpS. Quantitative mass spectrometry using SILAC (stable isotope labeling by amino acids in cell culture) has provided a ClpS-dependent comparative identification of putative ClpAP substrates. Although two proteins with ssrA-like C-termini, RecN and the N-terminal fragment of RseA, have shown to be degraded by ClpAP in a ClpS-dependent manner, thorough validation of the trapped substrates has been largely unsuccessful. We believe that substrate-processing steps may be occurring in the cell prior to ClpAP recognition and that i*n vivo* validation will enhance substrate verification.

INTRODUCTION

Although the AAA+ protease ClpAP has been studied for over 20 years, the field does not have a strong understanding of its role in the cell. Only a few physiological substrates are known, and many of these reported substrates are not fully characterized. ClpAP has been most studied in *E. coli*, but several of the putative ClpAP substrates are reported from other bacterial species. It is therefore not determined if all of these proteins are degraded by ClpAP in *E. coli*. For example, although the protease ClpXP is not essential in *E. coli* for viability, this protease is essential in the bacterium *Caulobacter crescentus* and plays a key role in cell cycle control (Jenal and Fuchs 1998). Thus, ClpXP has known dissimilarities in physiological roles, given the non-identical nature of proteomes between organisms.

MazE is one of few proteins reported to be degraded by ClpAP in *E. coli* (Aizenman et al. 1996). It is currently unknown how ClpAP recognizes the 9 kilodalton MazE as a substrate. Because ClpAP has been observed to degrade unfolded proteins lacking specifically-characterized degradation tags, it is possible that ClpAP recognizes an unfolded region of MazE for degradation. A more specific degradation tag has not been identified on this substrate. As described earlier in this thesis, ClpAP can degrade

ssrA-tagged substrates and N-end-rule substrates. However, no physiological N-end-rule substrates have been well-characterized to date, and ClpAP is not a major contributor to ssrA-tagged protein degradation in the cell under the conditions tested (Farrell et al. 2005; Lies and Maurizi 2008).

Overlapping specificities between different proteases also makes understanding the physiological role of ClpAP difficult. Both ClpXP and ClpAP degrade ssrA-tagged proteins, for example (Gottesman et al. 1998). It is very likely that other redundant substrates exist in *E. coli*, and this redundancy may be the cell's mechanism to ensure that priority substrates are degraded, even when the principal protease is incapacitated.

Previously, proteome-wide approaches were used to uncover substrates of the ClpXP protease (Flynn et al. 2003; Neher et al. 2006). In these studies, *E. coli* strains were deleted in the genes for *clpP*, *clpA*, and *smpB*. The wild-type copy of the *clpP* gene was deleted and a peptidase active site mutant (*clpP* S97A, also known as $ClpP^{Trap}$) was introduced on a plasmid. As a result, ClpX was only able to feed substrates into $ClpP^{Trap}$ instead of wild-type ClpP. Instead of being degraded, the substrates were sequestered intact inside of the ClpP barrel. These ClpP variants were then purified, and the encapsulated substrate proteins were identified by tandem mass spectrometry (Fig. 4.1). In these studies, the *clpA* gene was deleted such that the substrates fed into the mutant ClpP were the result of ClpX only, and *smpB*, a protein involved in ssrA-tagging, was deleted such that the isolated substrates did not include stalled translation products slated for destruction. This *in vivo* protein trapping method generated a list of putative ClpXP substrates, many of which have been validated *in vivo* and *in vitro*.

The first of these ClpXP trapping studies also looked for ClpAP substrates in *E. coli*. Only a few proteins were identified, and none were tested further for validation: AceA (aconitase), GapA (glyceraldehyde 3-phosphate dehydrogenase), OmpA (outer membrane protein A), and TnaA (tryptophanase) (Flynn et al. 2003). It seemed unusual that the *clpA* gene would be conserved solely to degrade a few substrates when ClpXP had been attributed dozens of substrates. Therefore, we embarked on a new study to more thoroughly search for ClpAP substrates.

The second ClpXP proteomics study uses the method of SILAC (stable isotope labeling by amino acids in cell culture) (Neher et al. 2006). This labeling method,

originally used in mammalian cell culture, takes advantage of the ability to grow cells in a medium supplemented with heavy-isotope amino acids (Ong et al. 2002). The study conducted by Neher and colleagues aimed to identify ClpXP substrates induced by DNA damage. Leucine auxotrophic strains were employed in which untreated cells were grown with ^{13}C-leucine and cells exposed to DNA-damaging agents were grown with ^{12}C-leucine. After growth and *in vivo* protein trapping with ClpP^{Trap}, the two cellular extracts were combined, and the trapped substrates were purified and identified by mass spectrometry (Fig. 4.2). Peptides used for mass spectrophotometric analysis from substrate proteins trapped in the ^{13}C-leucine medium differ from ^{12}C-leucine substrate peptides by six mass units per leucine. Leucine was selected because it is one of the most abundant amino acids in *E. coli* (Neidhardt and Curtiss 1996), increasing the probability that at least one leucine will exist per peptide analyzed. Due to the differences in mass, SILAC can quantify the ratio of peptides from substrates trapped in one strain versus the other. Thus, this method can reveal whether or not particular substrates are preferentially trapped in one cell type versus another.

In extension of these studies, we have utilized the ClpP^{Trap} to identify *in vivo* substrates of ClpAP. Furthermore, we are specifically interested in how the adaptor ClpS affects ClpAP substrate selectivity. Here, we describe the utilization of the conventional trapping method as well as SILAC to reveal potential ClpAP substrates.

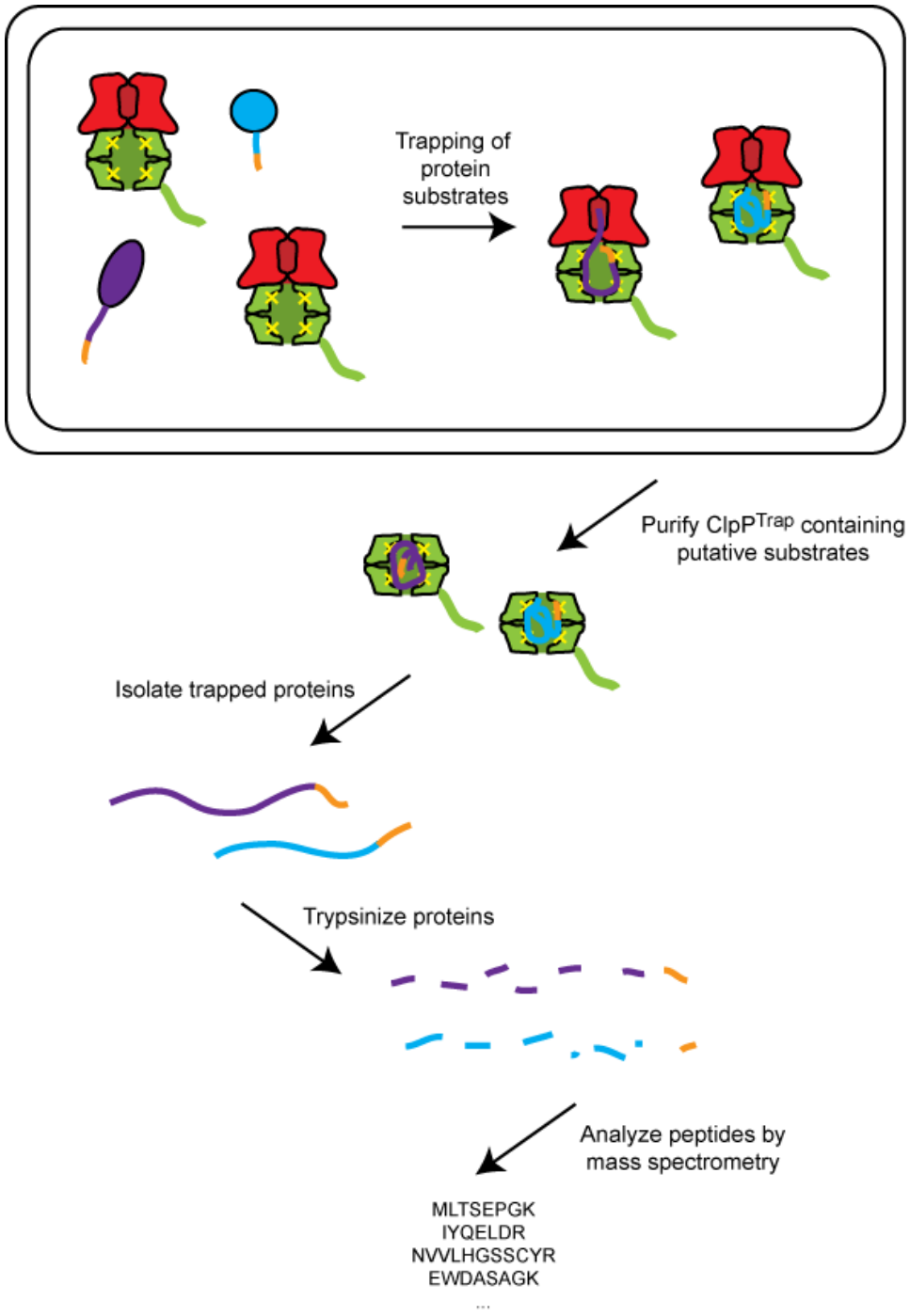


Figure 4.1. *In vivo* ClpP trapping.
The AAA+ chaperone (red) recognizes substrates in the cell by their degradation tag (orange) and feeds them into the ClpP^{Trap} (green). Cells are lysed, and the ClpP^{Trap} proteins harboring the undegraded substrates are purified using its epitope tag (shown as a green tail). After purification, the ClpP^{Trap} is removed, and the isolated substrate proteins are trypsinized and prepared for tandem mass spectrometry. Peptides are then identified and matched to their corresponding protein in the *E. coli* proteome.

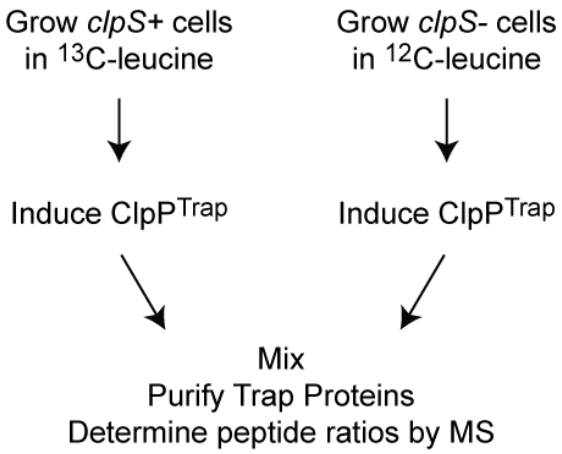


Figure 4.2. SILAC procedure.
Overall schematic of the *clpS*+ versus *clpS*- trapping experiment using heavy-isotope labeling. Proteins are trapped in exclusively light or heavy leucine, and samples are mixed prior to purification of ClpP^{Trap}.

RESULTS

Initial protein trapping in ClpAP

Protease systems have modular components that influence their substrate profiles. For example, the presence or absence of adaptor proteins can play a major role in deciding which substrates are degraded. The N-terminal domains of the AAA+ subunit of the protease typically serve as docking sites for adaptors, but N-terminal domains are also modular entities that have been characterized to bind some substrates. We used ClpP trapping to identify substrates of ClpAP, but more specifically, we sought to examine differences in substrate landscapes between cells with differing substrate modulators. The goal therefore was to identify substrates uniquely trapped by ClpAP in cells with or without *clpS* and with or without the N-terminal domain of ClpA. By probing the role of ClpS, we hoped to determine how the adaptor influences ClpAP substrate selection. The trapping experiments targeted towards the N-terminal domain of ClpA aimed to identify substrates that either require initial binding to ClpA's N-terminal domain or are delivered by unidentified ClpA adaptors.

Multiple repetitions of the trapping experiments followed by mass spectrometry showed little overlap in the substrate profile. It is reported that mass spectrometric

analysis of large samples often only reveals a small percentage of the available proteomic space, often biased toward abundant proteins (Bantscheff et al. 2007). This lack of reproducibility might explain why multiple repetitions of the same trapping experiments would yield results with little overlap. Moreover, most substrates were identified by mass spectrometry by only one or two unique peptides. These identifications are more prone to sampling error and are therefore less confident than if peptides covering the entire length of the putative substrates had been detected, and no particular substrate emerged as a very strong "hit."

With regards to searching for N-end-rule substrates, it is difficult to determine if the putative substrate is an intact protein or one that has been cleaved prior to degradation by ClpAP. It is possible that a subset of substrates requires proteolytic processing before being recognized by ClpAP. Trapped proteins were prepared for mass spectrometry by running the sample on SDS-PAGE, excising gel bands, and performing trypsinization on each gel slice (Fig. 4.1). After mass spectrometry and identification of the trypsinized peptides, the mass of the identified protein was checked to see if it corresponded with the average mass of the proteins in its gel slice. Our results were unclear if any putative substrate was processed prior to targeting to yield a significantly different size.

Because the data lacked sufficiently high resolution to differentiate between ± *clpS* and ± ClpA N-domain substrates, the initial protein trapping studies focused on simply assembling a general list of potential ClpAP substrates irrespective of modulating factors. Control trapping experiments were conducted in which both the *clpA* and *clpX* genes are deleted from the genome. Any proteins isolated in $ClpP^{Trap}$ from the *clpA*- and *clpX*- strain were designated as non-specific proteins. There are several possibilities to explain the presence of proteins in the control trapping experiment. These identified proteins might interact with the exterior of $ClpP^{Trap}$, be fed into $ClpP^{Trap}$ by an unidentified AAA+ partner unfoldase, or co-purify with $ClpP^{Trap}$ after cell lysis as contaminants. A list of proteins from the control experiments was generated, and trace amounts of several proteins such as RpoB, the beta subunit of RNA polymerase, likely persisted through the purification process as contaminants due to their high abundance in the cell (Ishihama et al. 2008).

Several candidate substrates that appeared in multiple runs of trapping experiments were selected for validation of ClpAP degradation. Two of the putative substrates that were tested by *in vivo* degradations were HchA (chaperone Hsp31) and Prs (phosphoribosylpyrophosphate synthetase). Genes encoding these proteins were cloned into plasmids bearing either an N-terminal tag or a C-terminal tag and expressed in cells with and without the *clpA* gene. The epitope tags enabled the use of antibody detection on western blots to monitor substrate degradation, and because most known degradation tags exist on the N- or C-terminus of proteins in bacteria, both N- and C-terminally-tagged constructs were created for each protein tested in the case that one particular tag blocked recognition by ClpAP.

If a protein was indeed a ClpAP substrate, degradation should be stabilized or proceed more slowly in the *clpA* mutant cell. Neither HchA nor Prs appeared to be reproducibly degraded, let alone stabilized in a *clpA*- strain (Fig. 4.3a). We took into consideration that under these growth conditions, ClpAP might be degrading too many other substrates to demonstrate degradation of our target substrates. Our trapping strains were deleted at the *smpB* locus to eliminate the possible competition of and confounding by ssrA-tagged proteins. We thus used strains with the same deletion to test for degradation. However, the elimination of the *smpB* gene did not result in any significant and reproducible changes in the putative substrate degradations (Fig. 4.3b).

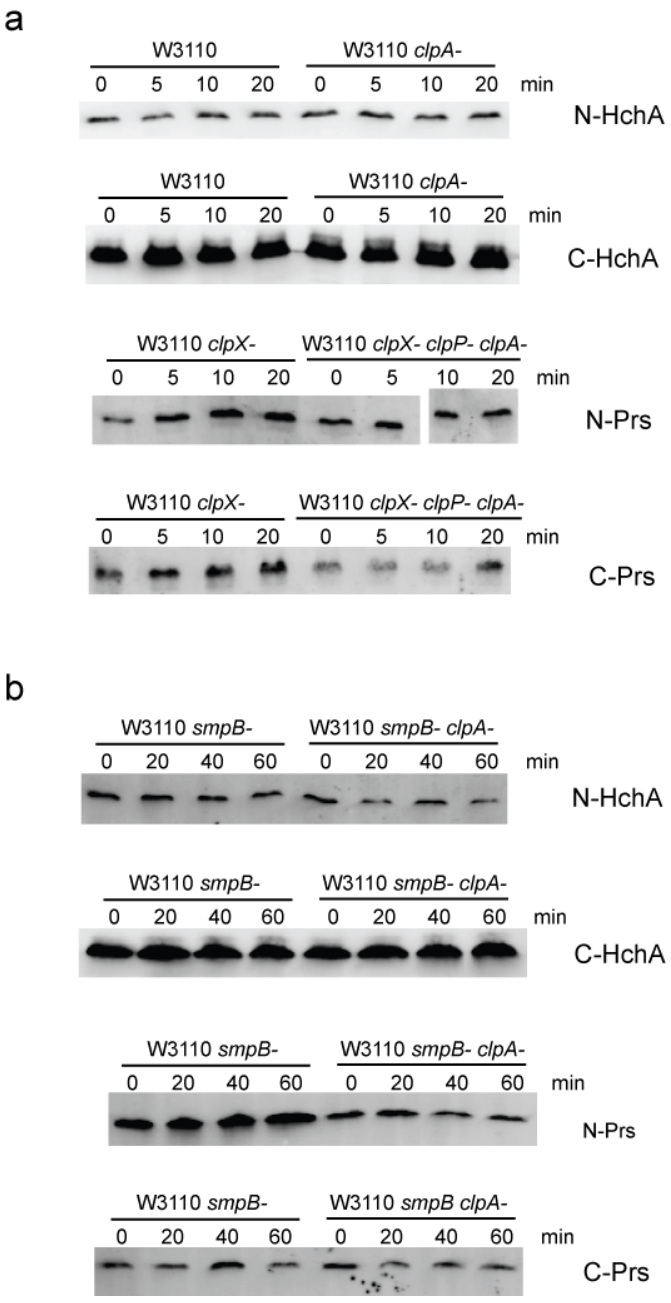


Figure 4.3. Testing of initial trapped proteins.
(a) Western blots showing degradation experiments of N- and C-terminally-tagged HchA and Prs *in vivo* with or without ClpA in a wild-type or *clpX*- background. (b) Western blots showing degradation experiments of HchA and Prs in W3110 *ΔsmpB* cells to test for competition with ssrA-tagged proteins.

Degradation of substrates with ssrA-like C-terminal tags

ClpAP is able to degrade ssrA-tagged proteins *in vitro* (Gottesman et al. 1998), and we noticed the presence of trapped proteins with ssrA-like motifs at their C-termini. HchA has a C-terminal motif with several di-alanine sequences like the ssrA tag, even though this protein did not appear to be degraded under our tested conditions. RecN, a DNA repair protein, ends in a di-alanine sequence and had been characterized to be a substrate of ClpXP (Nagashima et al. 2006; Neher et al. 2006). RecN was identified in several trapping experiments, suggesting that it might be a ClpAP substrate. A plasmid bearing an N-terminally-tagged construct of RecN was transformed into wild-type and *clpA*- strains and tested for degradation. RecN was degraded in wild-type cells and slightly stabilized in *clpA*- cells (Fig. 4.4a). We hypothesized that the degradation of RecN in the absence of ClpAP is likely due to ClpXP. Figure 4.4b shows the degradation of N-terminally-tagged RecN in a *clpX*- background, and although it is clear that ClpAP is making a contribution to the degradation, other proteases appear to participate in RecN proteolysis. Moreover, a version of RecN with two aspartates in place of the final two alanines was not degraded by ClpAP (Fig. 4.4c), showing that, similarly to ClpXP, the C-terminal alanines are important for ClpAP recognition (Neher et al. 2006). These results suggest that ClpAP is one of several proteases able to degrade RecN in the cell, and like ssrA-tagged proteins, further demonstrates the overlapping proteolytic capabilities of AAA+ proteases.

Because RecN was degraded in a manner dependent on its C-terminal di-alanines, we sought other proteins that might behave similarly. Although it did not appear in our trapping experiments, the N-terminal portion of RseA ($RseA^{1-108}$), an inhibitor of sigmaE, is a ClpXP substrate that bears similarities to the ssrA-tag and can be delivered for degradation by the adaptor SspB (Flynn et al. 2004; Levchenko et al. 2005). Multiple proteases have been shown to degrade purified $RseA^{1-108}$, including ClpAP (Chaba et al. 2007). In ClpAP, the degradation of $RseA^{1-108}$ is inhibited by ClpS and by SspB (Fig. 4.4d,e). The inhibition by ClpS is likely due to the ability of ClpS to alter ClpAP's substrate preference away from ssrA-like substrates, and SspB likely inhibits degradation by binding to $RseA^{1-108}$ and sequestering it from ClpA. Because $RseA^{1-108}$ has another set of di-alanines near the C-terminus, we tested whether or not these alanines could

serve as a degradation tag as well. Interestingly, ClpAP is also able to degrade this even shorter truncation of RseA ending with these upstream di-alanines ($RseA^{1-99}$), albeit less well than $RseA^{1-108}$, and this degradation is also inhibited by ClpS (Fig. 4.4f). It is thus unclear whether ClpAP prefers more sets of di-alanine motifs such that $RseA^{1-108}$ is degraded faster than $RseA^{1-99}$, or if $RseA^{1-108}$ simply has a longer unstructured C-terminus for ClpAP to grasp for degradation. Although $RseA^{1-108}$ is clearly degraded by ClpAP *in vitro*, we arise at several possibilities for why RseA was not identified in the ClpAP trap: limited sampling of mass spectrometry, ClpAP not degrading $RseA^{1-108}$ *in vivo*, and lack of extracytoplasmic stress to liberate $RseA^{1-108}$ (see Chapter One).

Although it is reasonable to assume that ClpAP would degrade a substrate more robustly *in vitro* in the absence of competitors than *in vivo*, no substrate tested from our initial list with the exception of RecN was degraded as definitively as model proteins, such as ssrA-tagged proteins that are degraded *in vitro*. Unlike $ClpXP^{Trap}$, which yielded several dozen "hits," some with many detected peptides, $ClpAP^{Trap}$ yielded longer lists of putative substrates with fewer peptides each. From our initial observations, one might conclude that ClpAP does not focus on degrading a small set of specific substrates, but instead degrades a wider range of substrates, each at a lower level. This hypothesis is consistent with *in vitro* experiments showing that ClpAP can degrade many, perhaps most, unfolded proteins. Additionally, with a large list of candidate proteins, the likelihood that the tested substrates would be degraded robustly is lower. Because mass spectrometry is not inherently a quantitative technique, we sought to find a method that could more accurately quantify the level of proteins in one trapping experiment versus another as well as distinguish actual substrates from contaminants. We expect that a quantitative technique should elucidate subtle differences between two traps rather than impose black-and-white standards on substrate qualifications.

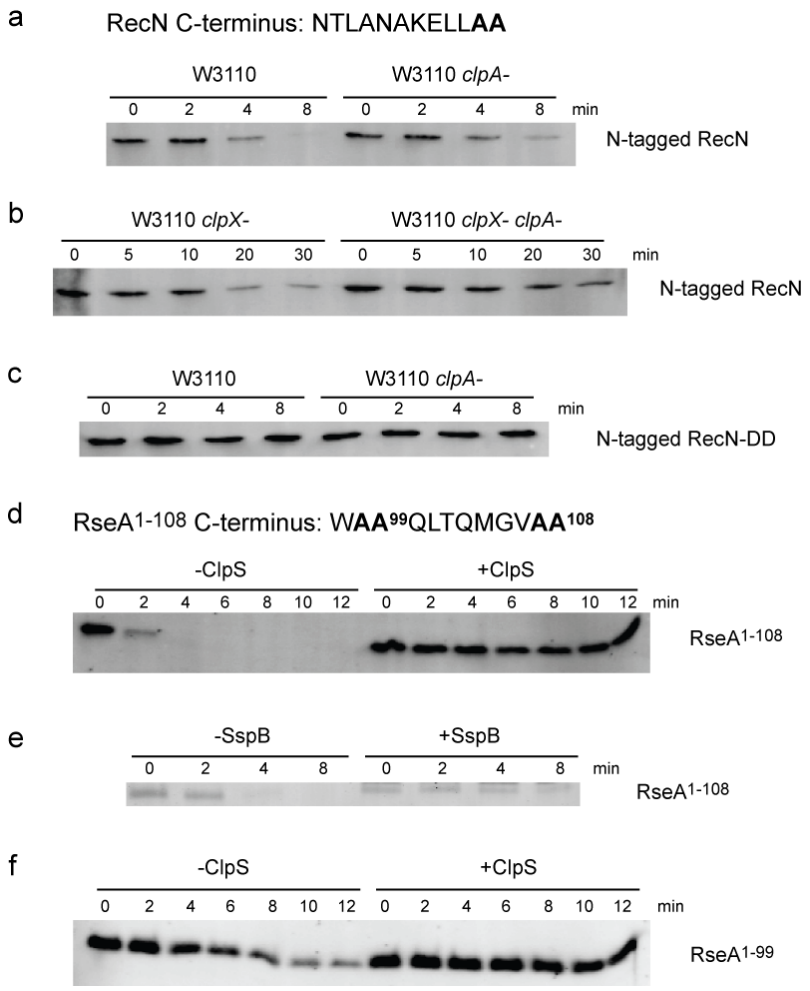


Figure 4.4. Degradation of ssrA-like proteins.
(a) Top: The C-terminal sequence of RecN is shown with di-alanines highlighted in bold. Bottom: Western blot showing *in vivo* degradation of N-terminally-tagged RecN. (b) Blot showing *in vivo* degradation of RecN in a *clpX*- background to highlight role of ClpAP. (c) *In vivo* degradation experiment of RecN-DD, in which the last two di-alanines are substituted by aspartates. (d) Top: C-terminal sequence of $RseA^{1-108}$. $RseA^{1-99}$ ends in –WAA. Di-alanines are highlighted in bold. Bottom: In vitro degradation of $RseA^{1-108}$ by ClpAP with and without ClpS. (e) *In vitro* degradation of $RseA^{1-108}$ by ClpAP with and without SspB as a competitor. (f) *In vitro* degradation of $RseA^{1-99}$ with and without ClpS.

Comparative protein trapping with SILAC

The SILAC approach permits a quantitative comparison of two different samples. Based on the ambiguous data observed from our initial trapping studies, we decided to employ SILAC as an approach to illuminate subtle differences between trapped protein

samples from different strain backgrounds. We primarily aimed to use SILAC to examine the role of ClpS in influencing ClpAP substrate preference in addition to understanding the general role of ClpAP *in vivo*. Furthermore, we made our purification of $ClpP^{Trap}$ slightly more stringent by using a different tagged construct and adding an extra purification step. Additionally, we ran the samples on a more sensitive mass spectrometer to detect low abundance peptides. Two different comparative SILAC experiments were carried out in a *clpX*- *smpB*- background: *clpA*+ versus *clpA*- (both *clpS*+) and *clpS*+ versus *clpS*- (both *clpA*+). The goal of the first experiment is to weed out the substrate proteins that bind to ClpP non-specifically or are contaminants in the purification procedure. The second experiment aimed to study the role of ClpS directly. All strains used in the SILAC protein trapping were deleted in the gene for *leuB* such that leucine was required in the growth media. The *clpA*+ and *clpS*+ strains were grown in ^{13}C-leucine, and the *clpA*- and *clpS*- strains were grown in ^{12}C-leucine.

The SILAC ratio is the ratio of heavy to light (H/L) peptides for a given protein. In our trap, a low H/L ratio indicates that more peptides of a particular protein were from the *clpS*- strain, and a high ratio indicates that more peptides were from the *clpS*+ strain. A SILAC ratio of 1 indicates that equal amounts of heavy and light peptides were detected for a protein. Thus, a low SILAC ratio implies that ClpS directly or indirectly inhibits the degradation of a substrate, and a high SILAC ratio implies that ClpS enhances the degradation of a substrate. Table 4.1 shows a list of substrates detected in the *clpS*+ versus *clpS*- trapping experiment. Substrates with a SILAC ratio of <1 in the *clpA*+ versus *clpA*- experiment have been eliminated from the *clpS*+ versus *clpS*- list as likely contaminants. Figure 4.5 depicts the range of SILAC ratios amongst the proteins trapped in the *clpS*+ and *clpS*- experiment.

Table 4.1. Putative substrates from *clpS*+ and *clpS*- SILAC trapping.[¥]

Gene	NCBI GI	Peptides[*]	SILAC Ratio (H/L)[+]	Description
cysH	16130669	2	0.181799	phosphoadenosine phosphosulfate reductase
lpxC	16128089	2	0.20466	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase
fruK	16130106	2	0.235401	1-phosphofructokinase
asnB	16128650	2	0.249121	asparagine synthetase B

agaR	16131023	2	0.269694	DNA-binding transcriptional dual regulator
yedO	90111355	3	0.30969	D-cysteine desulfhydrase
tgt	16128391	2	0.318358	queuine tRNA-ribosyltransferase
rplJ	16131815	5	0.332085	50S ribosomal protein L10
mog	16128003	2	0.335395	molybdenum cofactor biosynthesis protein
yjtD	16132220	2	0.336357	predicted rRNA methyltransferase
nirB	16131244	13	0.347698	nitrite reductase, large subunit, NAD(P)H-binding
asd	16131307	2	0.372036	aspartate-semialdehyde dehydrogenase
oppD	49176090	6	0.376818	oligopeptide transporter ATP-binding component
accA	16128178	4	0.381838	acetyl-CoA carboxylase subunit alpha
nfo	16130097	2	0.386495	endonuclease IV
galF	16129982	7	0.402537	predicted subunit with GalU
yqiB	16130929	2	0.411119	predicted dehydrogenase
accC	16131144	2	0.413407	acetyl-CoA carboxylase
aceE	16128107	11	0.415422	pyruvate dehydrogenase subunit E1
rfbD	16129980	2	0.418822	dTDP-4-dehydrorhamnose reductase subunit, NAD(P)-binding, of dTDP-L-rhamnose synthase
orn	90111697	2	0.420788	oligoribonuclease
thrS	16129675	3	0.424612	threonyl-tRNA synthetase
pfkA	16131754	6	0.431837	6-phosphofructokinase
wrbA	16128970	2	0.434417	TrpR binding protein WrbA
rpsA	16128878	10	0.435948	30S ribosomal protein S1
pheS	16129670	4	0.438021	phenylalanyl-tRNA synthetase alpha subunit
yhfA	16131235	2	0.440664	hypothetical protein
poxB	16128839	2	0.440869	pyruvate dehydrogenase
iscS	49176235	7	0.444081	cysteine desulfurase
yciW	90111242	2	0.444918	predicted oxidoreductase
yegS	16130026	2	0.464336	hypothetical protein
eda	16129803	2	0.476927	keto-hydroxyglutarate-aldolase/keto-deoxy-phosphogluconate aldolase
rfbC	16129978	5	0.493958	dTDP-4-deoxyrhamnose-3,5-epimerase
ydaO	16129305	3	0.496246	predicted C32 tRNA thiolase
pykF	16129632	8	0.502952	pyruvate kinase
rpoS	16130648	12	0.5056	RNA polymerase sigma factor
gatZ	16130033	5	0.520004	D-tagatose 1,6-bisphosphate aldolase 2, subunit
gmd	16129993	4	0.524833	GDP-D-mannose dehydratase, NAD(P)-binding
fabZ	16128173	3	0.527273	(3R)-hydroxymyristoyl ACP dehydratase
adhE	16129202	10	0.535466	fused acetaldehyde-CoA dehydrogenase/iron-dependent alcohol dehydrogenase/pyruvate-formate lyase deactivase
serS	16128860	2	0.550442	seryl-tRNA synthetase
ymdB	16129008	2	0.552689	hypothetical protein
ycaJ	16128859	3	0.553843	recombination protein
sfcA	90111281	2	0.559554	malate dehydrogenase, (decarboxylating, NAD-requiring) (malic enzyme)
ptsI	16130342	9	0.560333	PEP-protein phosphotransferase of PTS system (enzyme I)
rcsB	16130154	2	0.561389	DNA-binding response regulator in two-component regulatory system with RcsC and YojN
tyrB	16131880	3	0.566613	tyrosine aminotransferase, tyrosine-repressible, PLP-dependent
nadE	16129694	5	0.568537	NAD synthetase
glpK	16131764	2	0.574621	glycerol kinase
ndh	16129072	2	0.57531	respiratory NADH dehydrogenase 2/cupric reductase
cdd	16130081	4	0.580937	cytidine deaminase
lrp	16128856	2	0.582644	DNA-binding transcriptional dual regulator, leucine-

				binding
icd	16129099	2	0.592902	isocitrate dehydrogenase
udp	16131680	4	0.594177	uridine phosphorylase
insC-1	16128345	2	0.610037	IS2 insertion element repressor InsA
aceA	16131841	3	0.621476	isocitrate lyase
lepA	16130494	5	0.624024	GTP-binding protein LepA
nuoB	16130222	9	0.62902	NADH dehydrogenase subunit B
nuoC	16130221	2	0.634109	NADH:ubiquinone oxidoreductase, chain C,D
acnB	16128111	10	0.641847	aconitate hydratase
ahpC	16128588	4	0.64442	alkyl hydroperoxide reductase, C22 subunit
rfbA	16129979	7	0.647576	glucose-1-phosphate thymidylyltransferase
gatA	16130032	6	0.655826	galactitol-specific enzyme IIA component of PTS
nudC	49176450	2	0.662687	NADH pyrophosphatase
oppF	16129208	6	0.663085	oligopeptide transporter subunit
evgA	16130301	2	0.689258	DNA-binding response regulator in two-component regulatory system with EvgS
glf	16129976	3	0.690615	UDP-galactopyranose mutase, FAD/NAD(P)-binding
talA	16130389	8	0.6995	transaldolase A
cytR	16131772	2	0.712274	DNA-binding transcriptional dual regulator
rpsI	16131120	3	0.716812	30S ribosomal protein S9
clpA	16128850	3	0.717764	ATPase and specificity subunit of ClpA-ClpP ATP-dependent serine protease, chaperone activity
yjjI	16132197	2	0.719278	hypothetical protein
yeaK	16129741	2	0.722515	hypothetical protein
miaB	16128644	12	0.728181	isopentenyl-adenosine A37 tRNA methylthiolase
pepQ	16131693	2	0.74442	proline dipeptidase
ycfP	90111213	4	0.748727	hypothetical protein
manX	16129771	3	0.750749	fused mannose-specific PTS enzymes: IIA component/IIB component
rpoA	16131174	6	0.759549	DNA-directed RNA polymerase subunit alpha
clpX	16128423	5	0.761142	ATP-dependent protease ATP-binding subunit
amyA	16129874	2	0.765774	cytoplasmic alpha-amylase
bioB	16128743	2	0.771695	biotin synthase
dapA	16130403	2	0.774613	dihydrodipicolinate synthase
iadA	16132149	3	0.777779	isoaspartyl dipeptidase
murE	16128078	2	0.787788	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase
yejK	16130124	6	0.790699	nucleoid-associated protein NdpA
tufA	16131218	7	0.793572	protein chain elongation factor EF-Tu (duplicate of tufB)
ftn	16129855	2	0.795097	ferritin iron storage protein (cytoplasmic)
nadB	16130499	2	0.797112	L-aspartate oxidase
ydjA	16129719	5	0.813597	predicted oxidoreductase
fruR	16128073	3	0.824745	DNA-binding transcriptional dual regulator
rpsH	16131185	5	0.844545	30S ribosomal protein S8
gyrA	16130166	2	0.858286	DNA gyrase subunit A
bfr	16131215	4	0.86431	bacterioferritin, iron storage and detoxification protein
guaA	16130432	5	0.870947	bifunctional GMP synthase/glutamine amidotransferase protein
cysK	16130340	7	0.872268	cysteine synthase A, O-acetylserine sulfhydrolase A subunit
hdhA	16129577	3	0.873648	7-alpha-hydroxysteroid dehydrogenase
fabA	16128921	3	0.87647	3-hydroxydecanoyl-ACP dehydratase
phoP	16129093	4	0.878044	DNA-binding response regulator in two-component regulatory system with PhoQ
tig	16128421	10	0.881173	trigger factor
prsA	16129170	3	0.885314	ribose-phosphate pyrophosphokinase

yhdH	16131141	2	0.890545	predicted oxidoreductase, Zn-dependent and NAD(P)-binding
dhaH	90111232	3	0.894986	fused predicted dihydroxyacetone-specific PTS enzymes: HPr component/EI component
rne	16129047	10	0.907033	fused ribonucleaseE: endoribonuclease/RNA-binding protein/RNA degradosome binding protein
ygcF	16130684	2	0.909408	hypothetical protein
groL	16131968	8	0.914014	chaperonin GroEL
pflB	16128870	21	0.917315	pyruvate formate lyase I
nuoG	49176206	4	0.923445	NADH dehydrogenase subunit G
yegQ	16130021	2	0.943288	predicted peptidase
atpH	16131603	2	0.944706	F0F1 ATP synthase subunit delta
gyrB	49176395	9	0.961725	DNA gyrase subunit B
rpoB	16131817	21	0.96186	DNA-directed RNA polymerase subunit beta
pheT	16129669	3	0.972833	phenylalanyl-tRNA synthetase beta subunit
gpmA	16128723	2	0.973159	phosphoglyceromutase
ydgA	16129572	7	0.980119	hypothetical protein
yigI	90111652	5	0.984798	hypothetical protein
gldA	90111668	5	0.989556	glycerol dehydrogenase
frdB	16131978	2	1.000946	fumarate reductase (anaerobic), Fe-S subunit
rpoC	16131818	29	1.00198	DNA-directed RNA polymerase subunit beta'
rho	16131639	10	1.01238	transcription termination factor Rho
hepA	16128053	2	1.016214	ATP-dependent helicase HepA
ahpF	90111152	10	1.025201	alkyl hydroperoxide reductase, F52a subunit, FAD/NAD(P)-binding
rplB	16131196	10	1.028697	50S ribosomal protein L2
rplT	16129672	2	1.031948	50S ribosomal protein L20
mpl	16132055	6	1.034562	UDP-N-acetylmuramate:L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase
hslO	90111586	2	1.035115	Hsp33-like chaperonin
oxyR	16131799	2	1.037555	DNA-binding transcriptional dual regulator
mcrB	90111737	2	1.049078	5-methylcytosine-specific restriction enzyme McrBC, subunit McrB
rpsB	16128162	2	1.055113	30S ribosomal protein S2
sspA	16131119	2	1.059902	stringent starvation protein A
mreB	90111564	5	1.066764	cell wall structural complex MreBCD, actin-like component MreB
ydcF	16129375	2	1.0774	hypothetical protein
mrp	90111388	4	1.077511	antiporter inner membrane protein
rbsK	16131620	2	1.07937	ribokinase
talB	16128002	7	1.083969	transaldolase B
ribB	16130937	2	1.086943	3,4-dihydroxy-2-butanone 4-phosphate synthase
ybhA	16128734	4	1.093995	predicted hydrolase
pyrG	16130687	5	1.094551	CTP synthetase
ftsA	16128087	4	1.099742	cell division protein
ribC	16129620	2	1.100347	riboflavin synthase subunit alpha
folC	16130250	4	1.105252	bifunctional folylpolyglutamate synthase/ dihydrofolate synthase
cysN	16130658	2	1.109652	sulfate adenylyltransferase subunit 1
glpR	16131297	2	1.110796	DNA-binding transcriptional repressor
dnaJ	16128009	5	1.111569	chaperone Hsp40, co-chaperone with DnaK
yhbG	16131091	2	1.124241	predicted transporter subunit: ATP-binding component of ABC superfamily
hisB	90111373	3	1.12522	imidazole glycerol-phosphate dehydratase/histidinol phosphatase
narP	16130130	5	1.127565	DNA-binding response regulator in two-component

				regulatory system with NarQ or NarX
yajL	90111131	2	1.13211	hypothetical protein
ydhD	16129612	3	1.13322	hypothetical protein
rfaD	16131490	3	1.135183	ADP-L-glycero-D-mannoheptose-6-epimerase, NAD(P)-binding
iscR	16130456	2	1.136605	DNA-binding transcriptional repressor
yhdN	16131172	3	1.142832	hypothetical protein
pdxB	16130255	6	1.17485	erythronate-4-phosphate dehydrogenase
ygfB	90111511	2	1.179514	hypothetical protein
cysA	16130348	2	1.179913	sulfate/thiosulfate transporter subunit
spoT	16131521	2	1.185143	bifunctional (p)ppGpp synthetase II/ guanosine-3',5'-bis pyrophosphate 3'-pyrophosphohydrolase
cysC	16130657	2	1.194343	adenylylsulfate kinase
yeiE	16130095	5	1.215159	predicted DNA-binding transcriptional regulator
ansA	16129721	4	1.224707	cytoplasmic asparaginase I
prmB	90111419	2	1.225542	N5-glutamine methyltransferase
lipA	16128611	6	1.22884	lipoyl synthase
acrA	16128447	10	1.251609	multidrug efflux system
kbl	16131488	4	1.255954	2-amino-3-ketobutyrate coenzyme A ligase
glmS	16131597	28	1.256188	D-fructose-6-phosphate amidotransferase
mprA	16130596	2	1.261835	DNA-binding transcriptional repressor of microcin B17 synthesis and multidrug efflux
ytfP	16132044	2	1.265386	hypothetical protein
yeeN	16129927	2	1.267921	hypothetical protein
yneJ	16129485	3	1.272252	predicted DNA-binding transcriptional regulator
htpG	16128457	7	1.291656	heat shock protein 90
fusA	16131219	10	1.311558	elongation factor EF-2
rsuA	16130121	6	1.318663	16S rRNA pseudouridylate 516 synthase
ompA	16128924	8	1.319347	outer membrane protein A (3a;II*;G;d)
pspF	90111246	2	1.322085	DNA-binding transcriptional activator
dnaK	16128008	52	1.331164	molecular chaperone DnaK
hslU	16131769	8	1.346604	ATP-dependent protease ATP-binding subunit
crp	16131236	15	1.349575	DNA-binding transcriptional dual regulator
clpB	16130513	2	1.356743	protein disaggregation chaperone
yqjI	16130966	3	1.357565	predicted transcriptional regulator
dnaX	16128454	2	1.38914	DNA polymerase III subunits gamma and tau
grpE	16130533	8	1.389407	heat shock protein
katE	49176140	8	1.396297	hydroperoxidase HPII(III) (catalase)
rfaE	16130948	7	1.400256	fused heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase
clpP	16128422	37	1.402548	ATP-dependent Clp protease proteolytic subunit
wbbI	16129974	2	1.41023	hypothetical protein
trkA	16131169	18	1.410882	potassium transporter peripheral membrane component
hypB	16130634	11	1.41572	GTP hydrolase involved in nickel liganding into hydrogenases
glnD	16128160	3	1.430993	PII uridylyl-transferase
lon	16128424	10	1.465546	DNA-binding ATP-dependent protease La
fnr	16129295	2	1.468049	DNA-binding transcriptional dual regulator, global regulator of anaerobic growth
tpx	16129285	2	1.470169	thiol peroxidase
dps	16128780	7	1.478395	DNA protection during starvation conditions
ybeZ	16128643	3	1.497608	predicted protein with nucleoside triphosphate hydrolase domain
greA	90111554	4	1.528812	transcription elongation factor GreA
speG	16129542	3	1.539196	spermidine N1-acetyltransferase
rplN	16131189	7	1.555762	50S ribosomal protein L14

radA	16132206	5	1.620587	predicted repair protein
yahK	16128310	3	1.650819	predicted oxidoreductase, Zn-dependent and NAD(P)-binding
hscA	16130451	5	1.686969	chaperone protein HscA
rbfA	16131059	2	1.690763	ribosome-binding factor A
nagC	16128652	2	1.69563	DNA-binding transcriptional dual regulator, repressor of N-acetylglucosamine
moaB	16128750	2	1.734879	molybdopterin biosynthesis protein B
sthA	90111670	3	1.743317	soluble pyridine nucleotide transhydrogenase
recN	49176247	11	1.749762	recombination and repair protein
hns	16129198	2	1.800124	global DNA-binding transcriptional dual regulator H-NS
ykgG	90111113	5	1.820142	predicted transporter
ybaB	16128455	2	1.838957	hypothetical protein
tdk	16129199	3	1.845321	thymidine kinase
gadB	16129452	7	1.889412	glutamate decarboxylase B, PLP-dependent
glgA	16131303	8	1.924133	glycogen synthase
sucB	16128702	6	2.005071	dihydrolipoamide acetyltransferase
efp	16131972	2	2.065575	elongation factor P
atpA	16131602	9	2.098482	F0F1 ATP synthase subunit alpha
znuA	90111346	5	2.153056	high-affinity zinc transporter periplasmic component
hha	16128444	2	2.374178	modulator of gene expression, with H-NS
slyB	49176129	3	2.424912	outer membrane lipoprotein
yifE	16131624	2	2.530014	hypothetical protein
yceB	16129026	2	2.71146	predicted lipoprotein
znuC	16129811	4	2.808923	high-affinity zinc transporter ATPase
degQ	16131124	2	3.027239	serine endoprotease, periplasmic
murA	16131079	3	3.158488	UDP-N-acetylglucosamine 1-carboxyvinyltransferase
degP	16128154	4	3.187974	serine endoprotease (protease Do), membrane-associated
ibpB	90111637	2	3.28724	heat shock chaperone
pgi	16131851	2	3.407882	glucose-6-phosphate isomerase
hemA	16129173	2	3.488775	glutamyl-tRNA reductase
nhaR	16128014	2	3.600085	DNA-binding transcriptional activator
pyrB	16132067	2	3.612842	aspartate carbamoyltransferase catalytic subunit
mqo	16130147	2	3.715075	malate:quinone oxidoreductase
yhjK	90111607	3	3.773987	predicted diguanylate cyclase
fis	16131149	4	4.043683	DNA-binding protein Fis
atpF	16131604	3	4.267656	F0F1 ATP synthase subunit B
cbl	16129929	2	4.459872	DNA-binding transcriptional activator of cysteine biosynthesis
ftsH	16131068	5	5.517936	protease, ATP-dependent zinc-metallo
cysJ	16130671	4	6.224329	sulfite reductase, alpha subunit, flavoprotein
ibpA	16131555	2	6.272969	heat shock chaperone
yhjJ	16131399	2	6.778832	predicted zinc-dependent peptidase
carA	16128026	2	8.4552	carbamoyl-phosphate synthase small subunit
yrbC	16131082	2	9.575311	predicted ABC-type organic solvent transporter
carB	16128027	2	10.44367	carbamoyl-phosphate synthase large subunit
rseB	16130496	2	15.51535	periplasmic negative regulator of sigmaE
yiiD	16131728	3	17.75486	predicted acetyltransferase

[¥] Proteins identified in the *clpA*+ versus *clpA*- trapping experiment with a SILAC ratio <1 (enriched in the *clpA*- trap) are removed from this list. Proteins 3-fold enriched in the *clpS*- or the *clpS*+ trap are highlighted in bold.

[*] This column refers to the number of unique peptides identified that were quantified (both heavy and light peptides detected), with outliers removed.

[+]The SILAC ratio shown reflects the average SILAC ratio of the peptides included in the previous column.

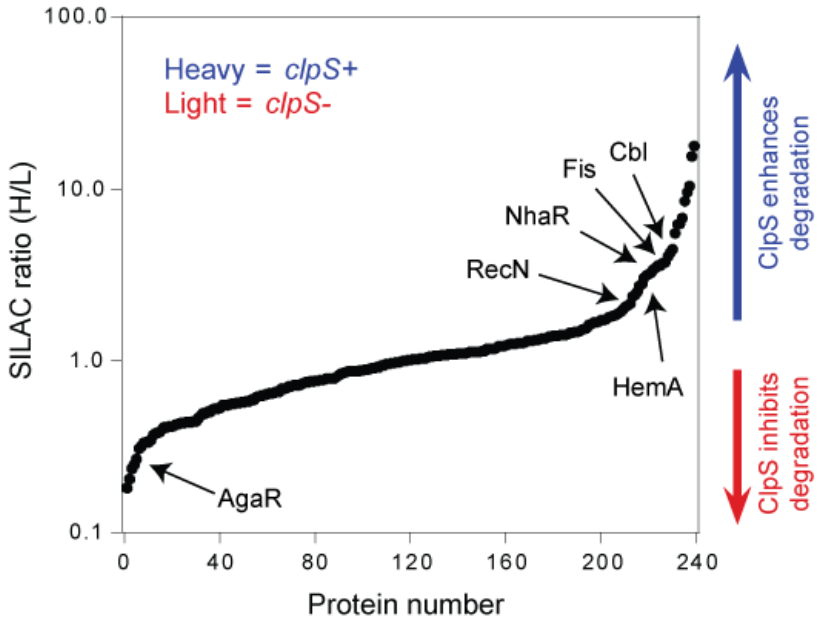


Figure 4.5. SILAC protein trapping.
Depiction of the protein trapped in the *clpS*+ versus *clpS*- experiment, with proteins trapped in *clpA*- cells removed. On the x-axis, proteins are numbered according to their SILAC ratio.

Some proteins in this list of over 200 genes are recognizable from lists generated from our initial trapping studies. Even though the list is weeded for proteins enriched in the *clpA*- trap, it is possible that non-specific substrates are still present in the *clpS*+ versus *clpS*- list due to experiment-to-experiment sampling limitations for mass spectrometry. For example, RpoB (ratio 0.96186) was detected in the *clpA*- control traps from our initial studies, but escaped screening in the SILAC experiment. HchA is not found in this experiment, but Prs (PrsA) is listed with a SILAC ratio of 0.885314. However, proteins like RecN (ratio 1.749762) and HemA (ratio 3.488775), a gene for glutamate synthesis characterized to be a ClpAP substrate in *Salmonella typhimurium*, appear in the SILAC list, providing one level of substrate verification.

We selected several SILAC substrates for further validation. We chose proteins with relatively extreme SILAC ratios because these substrates are more likely to exhibit the expected behavior with respect to ClpS. Furthermore, we decided to focus our attention on transcriptional regulators because transcription factors are generally in low abundance in the cells, and their presence in $ClpP^{Trap}$ is more likely to be specific

(Ishihama et al. 2008). Lastly, the proteins that we selected for further testing were annotated as being cytoplasmic, as proteins reported to be in the membrane or periplasm are more likely to be artifactual. The four transcription factors we chose for *in vivo* and *in vitro* verification are Cbl, NhaR, Fis, and AgaR. Cbl (ratio 4.459872), NhaR (ratio 3.600085), and Fis (ratio 4.043682) have high SILAC ratios, and AgaR (ratio 0.269694) has a low ratio.

Validation of trapping candidates identified by SILAC

First, N- or C-terminal epitope tagged versions of these proteins were expressed in cells for *in vivo* validation. Tagged Cbl, NhaR, and Fis did not show reproducible and robust evidence of proteolysis under these growth and induction conditions (Fig 4.6a). AgaR was visibly degraded *in vivo*, but neither N- or C-terminally tagged AgaR appeared to be stabilized in the *clpA*- strain (Fig. 4.6a). Because ClpAP and ClpXP share some substrates, the degradation of AgaR was tested in *clpX*- cells (Fig. 4.6b). However, there was no strong observable change in the degradation pattern, although C-terminally-tagged AgaR may be slightly stabilized in *clpX*- *clpA*- cells. As with validation of the first round of trapping studies using non-quantitative mass spectrometry, we considered testing the degradations in strains with *smpB* and other protease genes deleted to highlight a potentially subtle ClpA-dependent role. However, to directly monitor if ClpA recognizes these proteins, we decided to test for degradation *in vitro*.

Purified His-tag putative substrates were tested *in vitro* with and without ClpS, with the exception of N-terminally-tagged Fis which did not overexpress for purification. No degradation was observed under the tested conditions with or without ClpS (Fig. 4.7a). Degradation of these purified proteins was also tested with ClpXP because there is some overlap in substrate selectivity between ClpAP and ClpXP. However, neither Cbl, NhaR, Fis, nor AgaR was proteolyzed by ClpXP under these conditions, although there might be some clipping of C-tagged AgaR, as indicated by the increase of a slightly smaller truncation band (Fig. 4.7b).

None of the proteins from this small validation pool appeared to be ClpAP substrates under conditions tested. Furthermore, IbpA and IbpB, two small heat-shock proteins enriched in the *clpS*+ trap also did not appear to be degraded by ClpAP *in vivo* or

in vitro, with or without ClpS, under a number of surveyed conditions (S. Bissonnette, personal communication). The lack of ClpAP/ClpAPS-dependent degradation of all the tested trapping substrates might occur for several reasons. First, it is possible that the proteins identified in ClpAP^{Trap} are substrates because they are recognized as unfolded proteins in the cell. For example, a small percentage of a putative substrate may be unfolded and recognized, and the purified protein and plasmid-overexpressed substrates may be primarily folded and therefore not visibly degraded. However, *in vitro* experiments have shown that ClpS inhibits ClpAP's ability to recognize unfolded substrates, so this explanation is not consistent with putative substrates enriched in the *clpS*+ trap. Second, ClpAP may exist in the cell primarily for N-end-rule substrates, despite being a back-up protease to ClpXP for ssrA-like substrates. Perhaps most proteins found in the ClpAP^{Trap} are proteolytically processed to reveal an N-degron. Because N-end-rule substrates can also be degraded by ClpAP in the absence of ClpS, most of the proteins found in both *clpS*+ and *clpS*- traps may be N-end-rule substrates (Wang et al. 2007). Depending on the nature of the N-degron, ClpS may still exert a substrate preference to ClpAP, as indicated by the putative substrates that are detected with very high or very low SILAC ratios.

For the validation experiments, the plasmid-overexpressed proteins and proteins that are greatly overexpressed prior to being purified might overwhelm the processing machinery that reveals the N-degrons and therefore evade recognition by ClpAP(S). For the *in vivo* degradations, lowering the arabinose concentrations used for protein induction may alleviate a potential saturation. Unfortunately, insufficient peptides for each putative substrate were identified to ascertain whether or not a portion of the protein was trapped as opposed to the entire gene product. Moreover, proteins that are processed prior to trapping may yield an N-terminal peptide upon tryptic digestion (prior to mass spectrometry) that does not correspond to an expected mass and fragmentation pattern calculated by *in silico* protein trypsinization. Consequently, the truncated N-terminal peptide of these trapped proteins will likely be overlooked by the algorithms used to identify peptides.

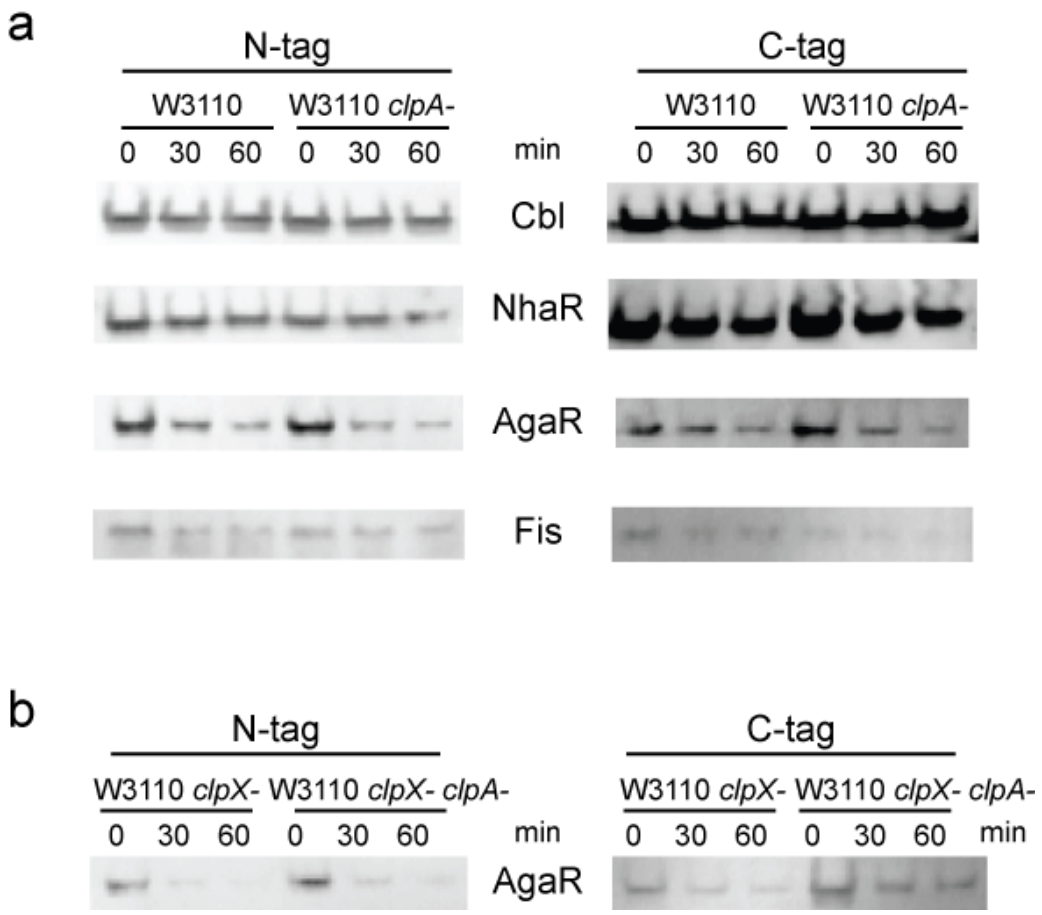


Figure 4.6. *In vivo* validation of transcription factors.
(a) Western blot showing *in vivo* degradation experiments of N- and C-terminally-tagged Cbl, NhaR, AgaR, and Fis. (b) Degradation of AgaR in a *clpX*- background.

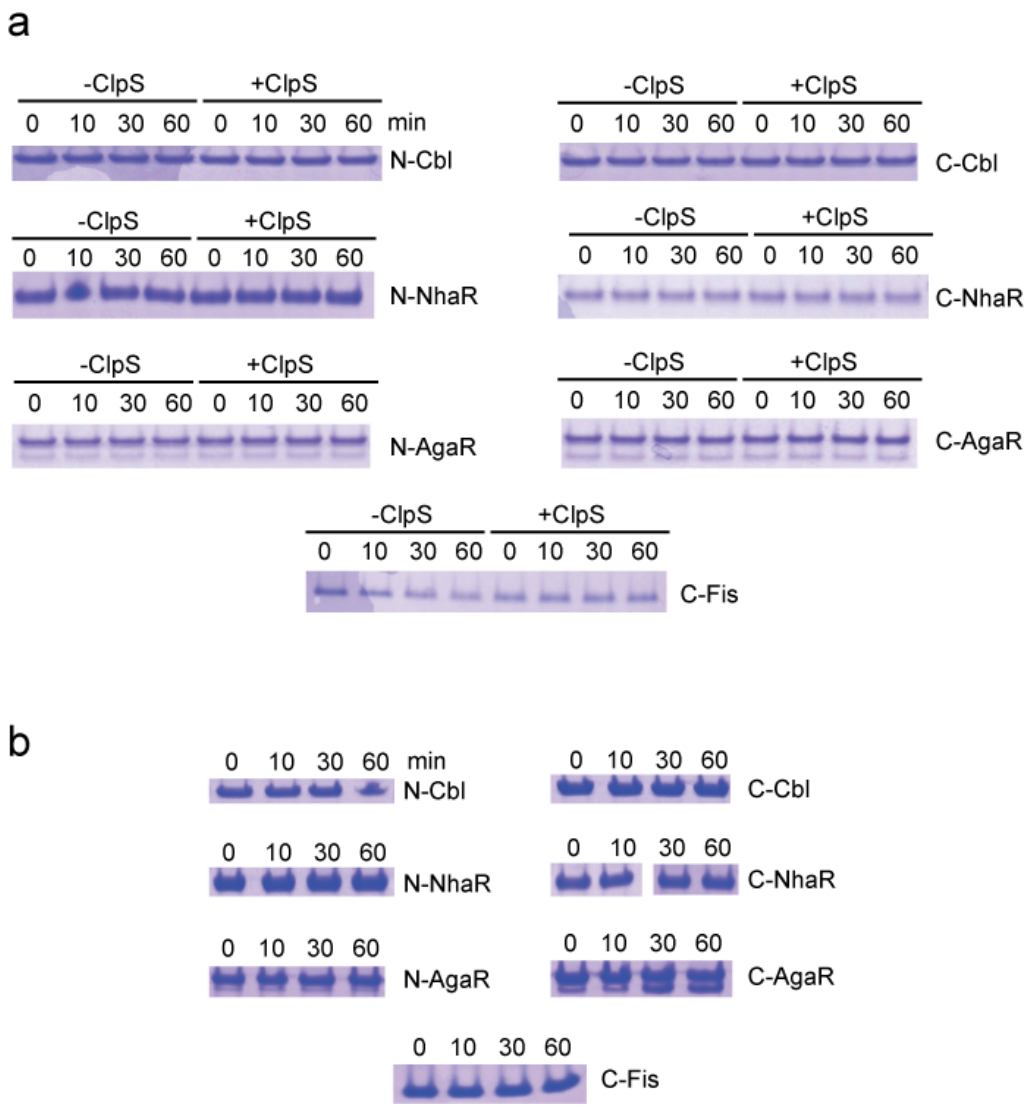


Figure 4.7. *In vitro* validation of transcription factors.
(a) *In vitro* degradation of purified N- and C-terminally-tagged proteins by ClpAP, with and without ClpS. (b) *In vitro* degradation of purified N- and C-terminally-tagged proteins by ClpXP. The accumulation of a truncation product is noticeable in the degradation of C-AgaR.

Re-trapping of SILAC-generated trapping candidates

To determine whether the entire or a part of the protein was trapped, we sought to "re-trap" the four candidate ClpAP substrates into $ClpP^{Trap}$. First, we conducted a control experiment to determine if plasmid-expressed GFP-ssrA could be trapped by ClpAP in the cell. GFP-ssrA was indeed enriched in the $ClpAP^{Trap}$; a slight band for GFP-ssrA appears in the western blot for both *clpS*+ and *clpS*- trapping experiments (Fig. 4.8a). Because ClpS inhibits the ability of ClpAP to degrade ssrA-tagged proteins, we expected GFP-ssrA to be more enriched in the *clpS*- strain trap. However, because the ClpS and ClpA ratios change during the transition from exponential to stationary phase, and substoichiometric levels of ClpS permit some degradation of ssrA-tagged proteins, it was not completely unexpected that some GFP-ssrA would also be trapped in a *clpS*+ strain (Farrell et al. 2005; Hou et al. 2008).

N- and C-terminal epitope-tagged constructs of the putative substrate genes were co-expressed with $ClpP^{Trap}$. Only N-terminally-tagged Cbl did not visibly appear by western blotting to be re-trapped (Fig. 4.8b). All of the other constructs appeared to be trapped in both *clpS*+ and *clpS*- cells. Unexpectedly, both N- and C-terminally tagged AgaR appeared as two bands, with one band at a much higher molecular weight than expected. The nature of this larger band is unknown, but the differential enrichment of the upper band is not identical to the differential enrichment of the smaller band. None of the detected bands appeared smaller by gel than predicted, implying that these trapped species either did not require a proteolysis step prior to ClpAP recognition or were processed very close to a terminus. Several constructs such as C-Cbl and C-NhaR yielded trapping patterns consistent with their SILAC ratio in which protein is enriched in *clpS*+ traps as opposed to *clpS*- traps. However, the C-Fis appeared to be slightly enriched in the *clpS*- trap, which contradicts its SILAC ratio. Although it is difficult to quantify the level of substrates trapped in each strain, this experiment shows that some of the proteins are able to be re-trapped into ClpP. However, these results cannot definitively address whether or not these proteins are being recognized by ClpAP as N-end-rule substrates after a processing step.

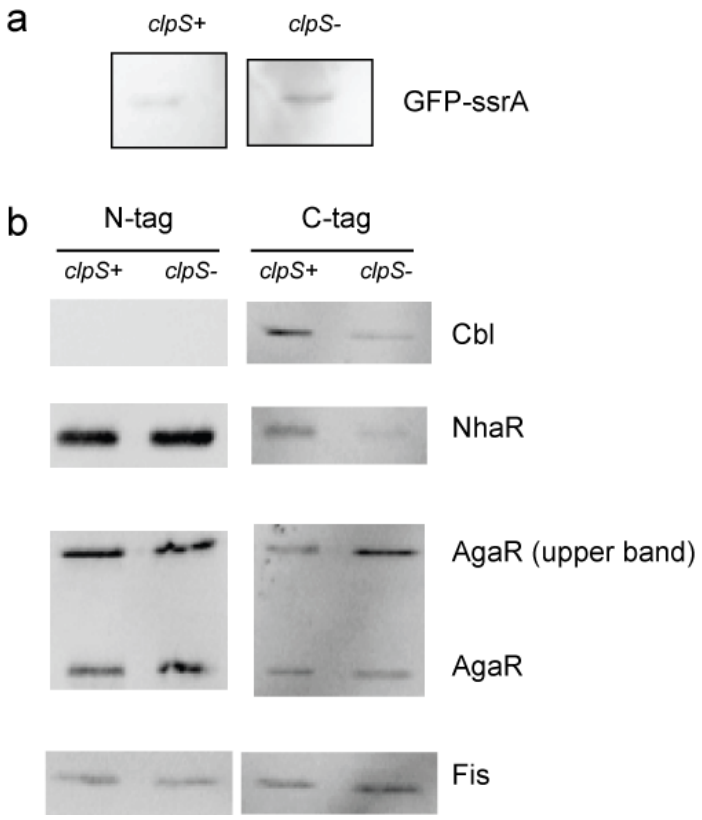


Figure 4.8. Re-trapping of putative ClpAP substrates.
(a) Western blots showing the trapped GFP-ssrA in *clpS*+ and *clpS*- cells. $ClpP^{Trap}$ was purified over streptactin beads and washed thoroughly before elution with biotin. (b) Western blots showing the relative re-trapping levels of tagged proteins using an antibody against the N- or C-terminal epitope tag. A higher molecular weight band reactive appeared in western blots for both N- and C-AgaR. The nature of this higher molecular weight band is unknown.

Fis, an endogenous candidate

Because there could be many explanations behind the lack of degradation with purified or overexpressed proteins, we decided to direct our attention to studying the endogenous levels of the candidate substrates. Fis was chosen for further study because previous reports indicate that Fis levels change according to the growth conditions of the cell (Ball et al. 1992). Data from *E. coli* and *Salmonella* demonstrate that Fis (factor for inversion stimulation) protein is involved in many roles such as site-specific DNA recombination, replication initiation, and transcriptional control (Osuna et al. 1995). Fis levels change dramatically according to the cellular growth conditions. When stationary phase bacteria are inoculated into fresh, rich media, Fis levels rise dramatically. After the

cells begin to divide in exponential phase, Fis levels fall to nearly undetectable levels by stationary phase (Ball et al. 1992; Osuna et al. 1995). Although the majority of these changes in Fis level can be attributed to changes in mRNA levels, one report suggests that the levels of Fis decline slightly more rapidly than would be estimated by cell division and dilution alone (Ball et al. 1992). Therefore, it may be possible that an active proteolytic mechanism is functioning at low levels in exponential phase.

We conducted experiments in which we back-diluted wild-type, *clpA*-, and *clpS*- cells grown overnight into fresh media. The growth curves showed that the *clpS*- cells grew consistently more slowly in exponential phase than wild-type or *clpA*- cells (Fig. 4.9a). However, we were unable to clearly detect Fis levels using our antibody. One piece of data suggests a similar growth pattern as reported in which Fis levels rise immediately upon back-dilution into fresh media and then drop as the cell proceeds into stationary phase (Fig. 4.9b). Furthermore, Figure 4.9b suggests that Fis levels may be slightly more stabilized in *clpA*- cells, but these results were generally too unclear and non-reproducible for certainty. New antibodies to Fis have recently been generated to address these experiments.

Based on the studies of Ball and colleagues in 1992, if Fis is being degraded in exponential phase, this degradation must be at low levels because simple dilution explains most of the decline in levels. Thus, if ClpAP is indeed degrading Fis, then the rate of degradation might also be slow. We returned to *in vitro* degradations of un-tagged Fis (a gift from R. Johnson) but with the use of higher enzyme concentrations. After 20 minutes, about 34% of Fis is degraded by ClpAP under these conditions. ClpS and Fis co-migrate on the gel, but quantification of the combined band shows that Fis is not detectably degraded (Fig.4.9c). These results contradict the trapping results. However, we return to our hypothesis that Fis, and other putative substrates, may indeed require processing to be efficiently degraded by ClpA, and that the observed degradation *in vitro* may simply be a fraction of protein that is transiently unfolded and engaged by the enzyme. Thus, although our results from Fis are currently inconclusive, we believe that studying putative substrates at their normally-expressed levels *in vivo* is the clearest way to determine if proteins are substrates of ClpAP.

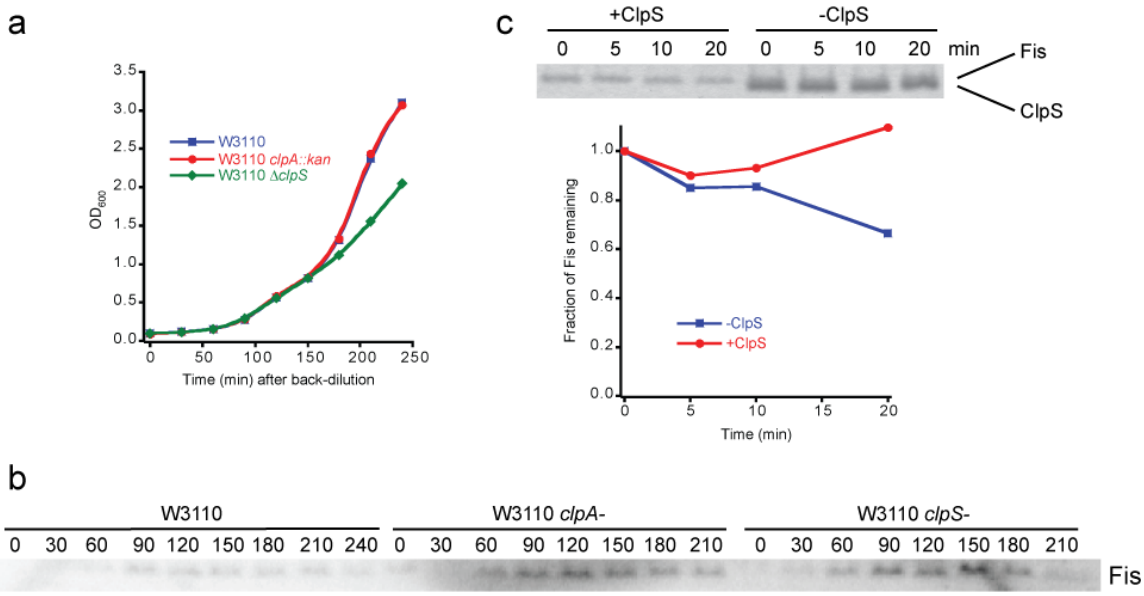


Figure 4.9. Validation of Fis degradation.
(a) Growth of cells after back-dilution into fresh, warmed LB used for monitoring endogenous Fis degradation. Cells lacking *clpS* consistently grew more slowly than wild-type or *clpA*- cells. (b) Western blot using anti-Fis antibodies to monitor Fis levels during the course of cell growth shown in A. Slight stabilization might exist in *clpA*- and *clpS*- cells, but blots were highly variable, and Fis detection was weak. (c) *In vitro* degradation of purified, un-tagged Fis protein by ClpAP with and without ClpS. Fis and ClpS co-migrate on SDS-PAGE (top), and the ClpS-Fis band was quantified together, assuming ClpS levels are stable (bottom). Slight degradation of Fis occurred without ClpS, and ClpS stabilized this degradation.

DISCUSSION

Despite extensive mechanistic study, the physiological role of ClpAP has been largely a mystery. Advances in mass spectrometry have assisted in proteomics experiments and have provided a powerful and sensitive tool to learning about the substrates of AAA+ proteins. Here, we have explored the substrate landscape of ClpAP in *E. coli*. Although our picture on the physiological role of ClpAP is still unclear, our mass spectrometry data suggests that ClpAP is not dedicated to degrading a few key substrates in the cell. Rather, ClpAP likely degrades a wide variety of proteins *in vivo*. The degradation of most of these proteins may be largely for quality control as opposed to being a key step in a regulatory pathway. The degradation of unfolded proteins and proteins possessing an N-degron may fall into the category of quality control. In theory,

any protein in the cell can potentially be unfolded or aberrantly cleaved to reveal an N-terminally destabilizing residue. This potential generality in substrate selection may explain why the trapping experiments for ClpAP yield a long list of substrates but few peptides per substrate. No singular protein stands out from the mass spectrometry data as a specific degradation target.

The majority of the proteins detected in the SILAC experiment were identified by five or fewer unique peptides (Table 1). Many of the proteins with high levels of unique peptides are involved in transcription or translation and are generally abundant in the cell and were thought to be non-specific contaminants (Ishihama et al. 2008). For instance, proteins such as RpoS, Crp, and GlmS have been identified in early *clpA*- control traps. We chose the transcription factors Cbl, NhaR, Fis, and AgaR for further testing because they had relatively extreme SILAC ratios, and the levels of transcription factors are generally low in the cell, suggesting that their presence in the trap may be more specific. Furthermore, transcription factors often gave positive results in validation experiments for ClpXP trapping studies (Flynn et al. 2003; Neher et al. 2003; Neher et al. 2006; Pruteanu et al. 2007). However, Cbl, NhaR, Fis, and AgaR were each only identified by 2-4 unique peptides. It is interesting to note that RecN, a validated substrate from our initial traps, was detected by 11 peptides. Though, RecN was enriched in the *clpS*+ trap yet seems to be inhibited by ClpS when tested directly *in vitro*. The correlation between peptide number and likelihood of being a substrate is not precise because smaller proteins will yield fewer peptides. Nevertheless, changing the strategy of validation toward testing proteins with the highest number of unique peptides, even if it is thought that the reason for the high numbers are due to contamination by abundant cellular proteins, might yield more success than guessing which proteins are the most specific by their function and generally low cellular abundance.

Because of the conservation of N-end-rule machinery, it is an attractive hypothesis that we have trapped some N-end-rule substrates. Substrate validation using a plasmid-overexpressed or purified full-length protein may not reveal degradation in this case. How do we determine which and how proteins are processed? One approach to determining if proteins are processed prior to degradation is by examining putative substrates individually. A set of specific antibodies would greatly assist in conducting *in*

vivo experiments, either by monitoring the size and degradation pattern of a putative substrate *in vivo* with and without ClpA and ClpS as well as detecting the size of a substrate from purified trap samples. Alternatively, a pull-down approach using ClpS as bait might isolate physiological N-end substrates (see Appendix).

Complementary SILAC studies might also elucidate the ClpAP substrate landscape. For example, the *clpS*+ versus *clpS*- experiment could be repeated but with ClpS overexpressed in the *clpS*+ strain to emphasize a difference caused by the presence of the adaptor. Protein trapping can be done under a variety of stress conditions such as heat-shock to determine if there are specific substrates under stressful environmental states. Much remains to be learned about the role of ClpAP and ClpS in the cell. Understanding the role of this protease system will not only expand our knowledge of this particular machine, but we will also understand more fully how the cell handles and prioritizes protein degradation.

MATERIALS AND METHODS

Strains and plasmids

All strains and plasmids used in this study are described in Table 4.2. JH43, JH41, and JH42 are strains used for the initial trapping experiments. Strains JH105, JH106, and SN26 were used for SILAC trapping experiments, in which pJF105 was transformed into SN26. For *in vivo* degradations, the appropriate pBAD plasmids were transformed into W3110-based test strains. Re-trapping experiments were conducted with putative substrate genes encoded into a fusion of pBAD-pACYC184 plasmids (similar to pJF122) and transformed into W3110-based test strains.

Table 4.2. Strains and plasmids used in this study.

Strains/plasmid	Relevant genotype	Source/reference
Strains		
W3110	F- λ- IN*(rrnD-rrnE)1 rph-1*	(Bachmann 1972)
JH8	W3110 *clpA::kan*	This work
JH15	W3110 *ΔclpS*	This work
TB380	W3110 *clpX::kan*	J. Flynn
JH11	W3110 *clpA::kan clpP::cat clpX::kan*	This work
TB371	W3110 *ΔsmpB*	J. Flynn
JH32	W3110 *ΔsmpB clpA::kan*	This work
JH43	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39	This work

JH41	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39	This work
JH42	W3110 *ΔsmpB clpA::kan clpP::cat clpX::kan*, pJH101, pSBN39	This work
JH45	W3110 *ΔsmpB clpA::kan clpP::cat clpX::kan*, pSBN39	This work
JH105	W3110 *ΔleuB ΔsmpB clpP::cat clpX::kan*, pJF105	This work
JH106	W3110 *ΔleuB ΔsmpB ΔclpS clpP::cat clpX::kan* pJF105	This work
SN26	W3110 *ΔleuB ΔsmpB clpA::kan clpP::cat clpX::kan*	(Neher et al. 2006)
JH243	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39, pMS30	This work
JH244	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39, pMS30	This work
JH237	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39, pJH190	This work
JH238	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39, pJH190	This work
JH239	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39, pJH195	This work
JH240	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39, pJH195	This work
JH245	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39, pJH191	This work
JH246	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39, pJH191	This work
JH247	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39, pJH196	This work
JH248	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39, pJH196	This work
JH249	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39, pJH192	This work
JH250	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39, pJH192	This work
JH251	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39, pJH197	This work
JH252	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39, pJH197	This work
JH253	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39, pJH193	This work
JH254	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39, pJH193	This work
JH255	W3110 *ΔsmpB clpP::cat clpX::kan*, pSBN39, pJH194	This work
JH256	W3110 *ΔsmpB ΔclpS clpP::cat clpX::kan* pSBN39, pJH194	This work
Plasmids		
pQE70	Vector, ColE1 ori	Qiagen
pACYC184	Vector, p15 ori	(Chang and Cohen 1978)
pJF122	Vector, pACYC ori tet^R / pBAD MCS	J. Flynn
pJH101	$clpA^{169-758}$ / pJF122	This work
pSBN39	$clpP^{14-207}$ S97A-His_6-TEV-Strep·Tag / pQE70	(Neher et al. 2006)
pJF105	$clpP^{14-207}$ S97A-His_6-TEV-Myc_3 / pQE70	(Flynn et al. 2003)
pBADHisA	Vector, ColE1 ori	Invitrogen
pBADMycHisA	Vector, ColE1 ori, NcoI mutated to AvrII	Invitrogen/S. Neher
pJH112	prs / pBADHisA	This work
pJH116	prs / pBADMycHisA	This work
pJH113	hchA / pBADHisA	This work
pJH114	hchA / pBADMycHisA	This work
pSBN74	recN / pBADHisA	(Neher et al. 2006)
pSBN80	recN-DD / pBADHisA	(Neher et al. 2006)
pJH151	cbl / pBADHisA	This work
pJH152	cbl / pBADMycHisA	This work
pJH153	nhaR / pBADHisA	This work
pJH154	nhaR / pBADMycHisA	This work
pJH155	agaR / pBADHisA	This work
pJH156	agaR / pBADMycHisA	This work
pJH157	fis / pBADHisA	This work
pJH158	fis / pBADMycHisA	This work
pET28b	Vector, f1 ori	EMD Biosciences
pET23b	Vector, f1 ori	EMD Biosciences
pJH167	cbl / pET28b	This work
pJH168	cbl / pET23b	This work
pJH160	nhaR / pET28b	This work

pJH161	nhaR / pET23b	This work
pJH169	agaR / pET28b	This work
pJH170	agaR / pET23b	This work
pJH171	fis / pET23b	This work
pMS30	GFP-ssrA / pACYC184	(Flynn et al. 2003)
pJH190	cbl / pBADHisA-pACYC184	This work
pJH195	cbl / pBADMycHisA-pACYC184	This work
pJH191	nhaR / pBADHisA-pACYC184	This work
pJH196	nhaR / pBADMycHisA-pACYC184	This work
pJH192	agaR / pBADHisA-pACYC184	This work
pJH197	agaR / pBADMycHisA-pACYC184	This work
pJH193	fis / pBADHisA-pACYC184	This work
pJH194	fis / pBADMycHisA-pACYC184	This work

Proteins

The genes for Cbl, NhaR, AgaR, and Fis were cloned into overexpression plasmids. pET28b was used to create N-terminal His-tag fusions, and pET23b was used to create C-terminal His-tag fusions. The overexpressed proteins were purified with 8 M urea and 50 mM sodium phosphate over Ni-NTA resin (Qiagen). The bound proteins were washed and eluted with the same buffer and 500 mM imidazole. Eluted proteins were then step dialyzed into 50 mM sodium phosphate (pH 8), 300 mM NaCl, 10% glycerol, and 1 mM DTT.

Purified Fis was a gift from R. Johnson (UCLA).

Non-quantitative protein trapping

Overnight cultures of the trapping strains were inoculated into fresh medium and grown at 30 ºC. At OD_{600} 0.4, the cells were induced with 0.2 mM IPTG for 3 hours. Cells were harvested and lysed, and 200 µL of 2 mg/mL avidin was added per initial liter of starting culture to bind free biotin. Cells were spun and the lysate added to Ni-NTA beads. The bound protein was washed and eluted with high imidazole buffer, and the eluate was combined with IBA streptactin-sepharose beads. After additional washing, the protein was eluted with 10 mM biotin. The eluted protein was then dialyzed into buffered 8 M urea, and the $ClpP^{Trap}$ was removed by binding to Ni-NTA beads. The flow-through containing the trapped proteins was concentrated and prepared for mass spectrometry. The samples were either subjected to in-solution trypsinization or in-gel trypsinization after running the samples on a denaturing polyacrylamide gel, staining with

Coomassie, and excising gel slices. The resultant peptides were subjected to reverse-phase HPLC followed by tandem mass spectrometry.

SILAC protein trapping

The preparation of SILAC media and growth of trapping strains are as described, and pairs of strains were trapped for equal amounts of time (Neher et al. 2006). The final cell densities were measured by absorbance at 600 nm light, and equal amounts of cells were harvested from each trapping strain, mixed, and the $ClpP^{Trap}$ was purified as described (Flynn et al. 2003). The purified trapped proteins were dialyzed into buffered 8 M urea, and the $ClpP^{Trap}$ was removed by binding to Ni-NTA resin. The protein trapping samples from the *clpA*+ versus *clpA*- experiment were run on a 12.5% tris-glycine polyacrylamide gel. The *clpS*+ versus *clpS*- samples were run on a 12% tris-glycine gel (Bio-Rad). Both gels were stained with Coomassie after which gel slices were excised and the proteins were trypsinized and prepared for mass spectrometry (Neher et al. 2006).

In vivo degradations

pBADHisA or pBADHisMycA plasmids containing the candidate substrate gene of interest were transformed into wild-type W3110 and W3110 with the appropriate deletion(s). A single transformed colony was grown at 37 ºC in a 3-4 mL culture until reaching mid-exponential phase. Cells were then back-diluted into 20 or 25 mL fresh media, starting at an OD_{600} 0.2 or 0.3. 0.2% fresh L-arabinose was added to the culture, and the cells were grown to OD_{600} 0.3-0.4. 100 μg/mL chloramphenicol was added to stop protein synthesis, and 1 mL samples of cells were TCA-precipitated at various time points. 150 μg/mL spectinomycin was used when cells contained a chloramphenicol-resistant *clpP::cat* disruption. The TCA-precipitated cells were spun, wash with acetone, and resuspended with sample buffer before SDS-PAGE. The proteins were transferred to a PVDF membrane (Millipore) and probed with 1:5000 dilution of either rabbit anti-His (Santa Cruz Biotechnology) or mouse anti-Xpress antibody (Invitrogen) followed by a 1:5000 dilution of goat anti-rabbit or goat anti-mouse (Bio-Rad). The western blots were developed with ECF substrate (GE Healthcare) according to the manufacturer's instructions and scanned on a Typhoon 9400 (GE Healthcare).

In vitro degradations

RseA. 4 μM $RseA^{1-108}$ or $RseA^{1-99}$ was degraded with 0.1 μM $ClpA_6$, 0.27 μM $ClpP_{14}$ and either 0.6 μM ClpS or 6 μM $SspB^{1-119}$ monomer in 1X HO Buffer (50 mM HEPES pH 7.5, 20 mM $MgCl_2$, 0.3 M NaCl, 10% glycerol, 0.5 mM DTT) at 30 ºC with 4 mM ATP and an ATP regeneration system (2.5 μM creatine phosphate, 50 μg/mL creatine kinase). $RseA^{1-108}$ migrates closely with ClpS on SDS-PAGE, so a western blot was conducted to visualize $RseA^{1-108}$ degradation, probed with 1:10,000 rabbit anti-cytoplasmic RseA followed by 1:10,000 goat anti-rabbit. The protein blots were developed as described above.

N- and C-terminally-tagged Cbl, NhaR, AgaR, and Fis. 5 μM substrate was degraded with 0.1 μM $ClpA_6$, 0.27 μM $ClpP_{14}$, and 1 μM ClpS in 1X HO buffer or 0.3 μM $ClpX_6$, 0.8 μM $ClpP_{14}$ in 1X PD buffer (25 mM HEPES pH 7.6, 5 mM $MgCl_2$, 5 mM KCl, 15 mM NaCl, 0.032% [v/v] Nonidet P-40, and 10% glycerol) at 30 ºC with 4 mM ATP and the ATP regeneration system described above.

Fis. 5 μM un-tagged Fis was degraded with 1 μM $ClpA_6$, 1 μM $ClpP_{14}$, 10 μM ClpS, 5 mM ATP, 0.5 mg/mL creatine kinase, and 20 mM creatine phosphate at 30 ºC in 1X HO Buffer.

Protein re-trapping

Fusion pBAD-pACYC184 plasmids bearing N- or C-terminally-tagged genes for Cbl, NhaR, AgaR, and Fis were transformed into W3110 *ΔsmpB clpP::cat clpX::kan* pSBN39 or W3110 *ΔclpS ΔsmpB clpP::cat clpX::kan* pSBN39. GFP-ssrA was constitutively expressed, but Cbl, NhaR, AgaR, and Fis were induced with 0.2% arabinose. The $ClpP^{Trap}$ (His-TEV-Strep·tag) was induced with 0.2 mM IPTG during cell growth. Cells were harvested and lysed, after which the lysate was bound to streptactin beads, washed, and eluted with 10 mM d-biotin. The eluates were then run on a 12.5% tris-glycine gel and probed against the Xpress tag (for N-terminally tagged constructs), the Myc-tag (for C-terminally-tagged constructs) or GFP (for GFP-ssrA). Band densities were analyzed by ImageQuant TL software (GE Healthcare).

Fis growth curves

500 μL of W3110, W3110 *clpA::kan*, and W3110 *ΔclpS* cells grown overnight were inoculated into a total of 25 mL fresh, warmed Luria-Bertani (LB) broth. Cell densities were measured and cell samples were analyzed by anti-Fis western blotting using 1:20,000 rabbit anti-Fis primary antibody (a gift from R. Johnson [UCLA]) and 1:10,000 goat anti-rabbit secondary antibody.

CHAPTER FIVE

The future outlook of ClpS and ClpAP

Intense research has been conducted on the AAA+ protease ClpAP and its adaptor protein ClpS in recent years. Most of the work involving these proteins has been done *in vitro* in order to biochemically probe structure-function relationships. However, there are more mechanisms behind this system yet to be uncovered. Less work has been done examining the role of ClpAP and ClpS *in vivo*, and there are many lessons to be learned in understanding the role of this proteolytic system and why it has been conserved in so many bacterial species. Here, I address some future directions of ClpAP and ClpS research, some stemming immediately from the work presented in this thesis and others arising from gaps in knowledge in the field.

ClpS mechanism

Evidence supports a model in which ClpS serves to allosterically alter ClpAP substrate preference. The various parts of ClpS, comprised of the folded core with the N-degron binding pocket and the proximal and distal regions of the N-terminal extension, have distinct functions. What is unclear is how these distinct parts of ClpS work together to form a functional adaptor with respect to ClpAP and its substrates. Future work on the mechanism of ClpS will aim to understand the step-wise roles of the adaptor.

One of the most exciting areas to research is uncovering how ClpS hands-off N-end-rule substrates to ClpAP. ClpS docks onto the N-terminal domains of ClpA, which are thought to be highly mobile around the AAA+ rings of ClpA (Ishikawa et al. 2004). It is unknown whether ClpS binding affects the conformation and/or the mobility of the N-terminal domains. How can the ClpS N-terminus bind the ClpA pore regions if the ClpA N-terminal domains are still mobile? Perhaps the interaction of the ClpS N-terminus to ClpA is very transient, allowing only enough time for the N-end-rule substrate to be passed along to ClpA. This dynamic motion might be a component of the ClpS-induced rate increase for N-end-rule substrate degradation.

Furthermore, because ClpS can bind at up to six copies per ClpA hexamer, it is unclear how this stoichiometry affects the role of the adaptor. The work in this thesis demonstrates that titrating ClpS levels did not cause a linear effect on the inhibition of ssrA or the stimulation of N-end-rule degradation rates. For N-end-rule substrates, for example, the degradation rate enhancement maximizes at approximately 3-4 molecules of

ClpS per ClpA hexamer, even though up to six molecules of ClpS can bind. Does this substoichiometric rate maximization suggest that only 3-4 ClpS molecules loaded with an N-end-rule substrate can bind ClpA at any time, perhaps due to steric clashing at the pore? Or, is ClpA's ability to process N-end-rule substrates the rate-limiting step, in which more ClpS molecules bound does not increase the observed rate?

The N-degron binding pocket is another region of the ClpS molecule that is of immense interest in understanding the role of the adaptor. The recent crystal structure of an N-degron bound to ClpS greatly enhances our understanding of substrate recognition (Wang et al. 2008b). Future research needs to be dedicated to understanding if there is an active mechanism for ClpS to dissociate the N-degron from this binding pocket or if the natural on and off movements of the substrate allow it to be transferred over to ClpA for degradation.

ClpA mechanism

Although the majority of this thesis is focused on the role and mechanism of the adaptor ClpS, the protease ClpAP is also of great interest. Understanding the operational rules for ClpAP will aid in the understanding of ClpS and vice versa. The ClpA hexamer is a large molecule of over 500 kilodaltons in which each monomer is comprised of an N-domain and two ATPase domains. The N-terminal domains have been reported to be a site of substrate contact as well as the docking site for ClpS (Ishikawa et al. 2004; Xia et al. 2004; Erbse et al. 2008). Although ClpA lacking its N-terminal domain has shown to be impaired in recognizing model substrates, its role in substrate recognition is still unclear (Lo et al. 2001; Hinnerwisch et al. 2005b). With regards to the ATPase domains, it is also unclear why some AAA+ proteins have two AAA+ rings and some only have one. Some differing roles have been attributed to the two ATP-domains of ClpA, but it is not understood why some AAA+ enzymes, such as ClpX, can accomplish similar and specific proteolytic tasks with one.

Because I propose that ClpS induces a conformational change in ClpA, a future direction could be to identify what features of the ClpA hexamer change upon adaptor binding. Preliminary studies monitoring tryptophan fluorescence, bis-ANS binding, and protease sensitivity have been attempted but with inconclusive results. The ClpA

hexamer is a large molecule in which subtle changes may be difficult to detect. Therefore, more sensitive assays will be required to probe these changes such as small angle x-ray scattering or crystallography, if possible.

As mentioned in Chapter Three, we have narrowed down a putative region of interaction on ClpA where the ClpS N-terminal extension interacts. Cross-linking studies to achieve the same goal were conducted, but with ambiguous results. However, different techniques with different probes can be explored to triangulate the results gathered with the FeBABE reagent. Once a sharper picture of the interaction zone is created, more ClpA mutants can be generated with higher confidence of success. The discovery of a unique point of contact on ClpA will be a milestone for understanding the role of ClpS in the modulation in substrate selection.

In vivo roles and regulation

As described above, little information is known about the physiological role and regulation of ClpAP and ClpS. Therefore it will be important to understand why these proteins are conserved in cells and how they are regulated.

The immediate future directions I propose are a continuation on the proteomic-based work described in this thesis. We have generated a list of putative ClpAP substrates that are affected by ClpS. Because validation has been inconsistent, I propose validating substrates by looking at their endogenous levels as opposed to validation upon overexpression or purification. The principle reason for this strategy is that many putative substrates, particularly those enhanced for degradation in the presence of ClpS, may be processed prior to ClpAP recognition. Furthermore, we cannot rule out that the effect on substrate degradation by ClpS may be indirect, that is, requiring other factors in the cell. Fis is one protein described herein that may be a substrate for ClpAP. Immediate future work monitoring Fis levels *in vivo* with an improved antibody will greatly assist in this validation.

Moreover, the assay for ClpS pull-downs (described in Appendix) in which we searched for novel ClpS-binding proteins, particularly N-degron-bearing proteins, can be further refined. These pull-downs are a straightforward method to directly look for naturally occurring N-degrons in the cell that can be recognized by ClpS. If ClpS-

binding proteins are found, they can be tested for degradation by ClpAP to determine if they are also proteolytic substrates. No physiological N-end-rule protein has been well-characterized in *E. coli* to date, but a very recent study suggests that the DNA protection protein Dps and Putrescine-Aminotransferase are being recognized as N-end-rule substrates (Schmidt et al. 2009). Although the N-end rule may be part of a general quality control mechanism in bacteria, there are likely to be proteins that are processed more frequently. The degradation of some proteins by the N-end-rule pathway may be part of regulatory strategies in the cell.

Both the comparative $ClpP^{Trap}$ and pull-down assays can be further conducted under a variety of environmental conditions. For example, ClpXP has a reported role in degrading specific proteins under DNA-damaging conditions (Neher et al. 2006). ClpAP and ClpS may also have undiscovered roles in a particular stress response that is undetectable under normal conditions.

Similarly, the field does not have an understanding on the regulation of ClpAP and ClpS at the transcriptional level. ClpS and ClpA lie in the same operon (Dougan et al. 2002). However, it was shown that ClpA levels rose from exponential to stationary phase while ClpS levels remained constant in the cell (Farrell et al. 2005). The reasons why ClpA levels rise are nebulous, although we speculate that ClpAP may degrade a different profile of substrates when its levels are raised above saturating ClpS levels. It is also unknown what factors cause this increase in ClpA production. Likewise, the role of ClpA autodegradation, if it exists robustly *in vivo*, is unclear. Thus, understanding how components of this proteolytic machine are themselves regulated will assist in the understanding of how ClpAP and ClpS affect the cellular proteome under different growth conditions and environmental states.

In this thesis, I have described the multi-faceted role of ClpS in ClpAP substrate selection. With time, the field has adopted a greater understanding of the diverse make-up of AAA+ enzymes and adaptor proteins. ClpS represents a new class of adaptor that does not function like the canonical molecule. Different stoichiometries of ClpS can bind to a ClpA hexamer, offering one level of substrate regulation. Secondly, distinct parts of the adaptor are responsible for affecting different types of substrates. We provide a

deeper examination of the function of the N-terminal extension of ClpS, showing that the length and chemical characteristics are important, despite its lack of conversation between ClpS orthologs. Furthermore, I have focused the area of interaction of the ClpS N-terminal extension to the pore of ClpA, which will facilitate the understanding of how ClpA is affected by adaptor binding. Finally, we have generated a list of putative ClpAP substrates that are stimulated or inhibited by the presence of ClpS in the cell. Hopefully these discoveries will encourage further work on ClpAP and ClpS.

APPENDIX

Preliminary search for ClpS-interacting proteins

The cross-linking part of this work was done in collaboration with K. Wang.

INTRODUCTION

Little is known about the physiological role of ClpS in the cell. ClpS's role in delivering N-end-rule substrates to ClpAP has not been verified *in vivo*, and it is unclear if ClpS only binds to ClpA and potential N-degron-bearing substrates in the cell. Thus, we decided to explore the ClpS interactome through several co-precipitation experiments. We expected that the proteins pulled-down, if any, by ClpS may yield novel ClpS-dependent substrates or other currently identified binding partners. This project is unfinished, and future directions and considerations are described below.

RESULTS AND DISCUSSION

To investigate potential binding partners of ClpS, we implemented pull-down experiments. Two different methods of attachment of ClpS to a solid support were initially used: covalent linkage of ClpS primary amines to N-hydroxysuccinimide esters on an Affigel-15 matrix and non-covalent linkage of ClpS with a C-terminal His-tag to Ni-NTA resin. The advantage of the primary amine linkage is that ClpS can presumably bind to the solid support in multiple orientations via its nine lysines and N-terminus, thus not systematically eliminating a possible protein binding interaction through steric occlusion. Conversely, the advantage of C-terminal linkage to a column is that the point of linkage is known. The C-terminal His-tag did not appear to inhibit the folding and function of ClpS-H_6 in inhibiting GFP-ssrA degradation by ClpAP (Fig. A.1).

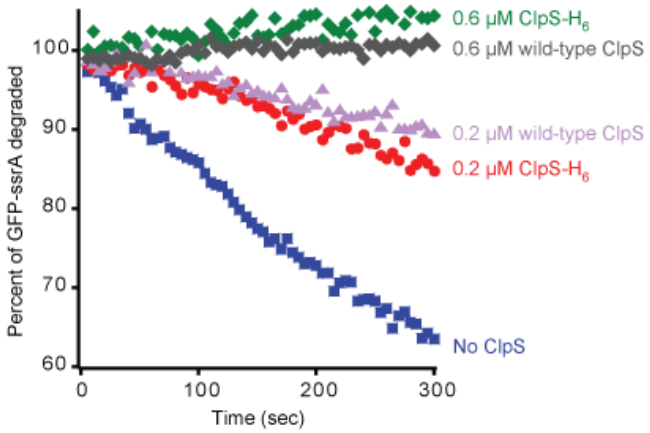


Figure A.1. Testing of ClpS-H_6.
The degradation of GFP-ssrA by ClpAP was monitored with wild-type ClpS or ClpS-H_6. At the two adaptor concentrations tested, the His-tagged ClpS behaves similarly to wild-type ClpS.

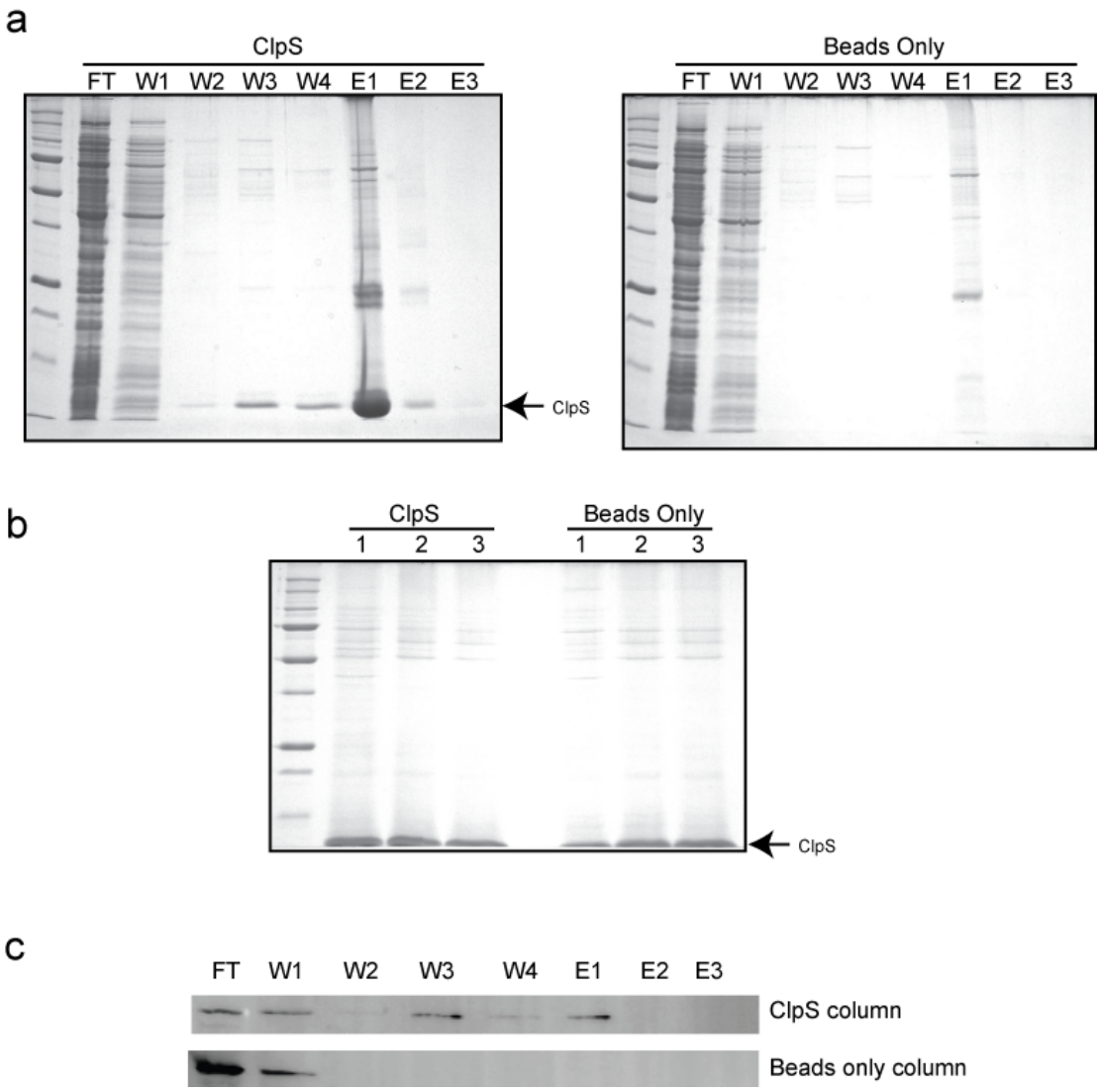


Figure A.2. Initial ClpS co-precipitation experiments.
(a) Coomassie-stained polyacrylamide gels showing isolation of putative ClpS-binding proteins from ClpS-H_6 column. FT = flow through; W1 and W2 = washes with 10 mM imidazole buffer; W3 and W4 = washes with 20 mM imidazole buffer; E1, E2, and E3 = elution fractions with 500 mM imidazole buffer. ClpS-H_6 partially emerged in the washes and was eluted from the matrix along with co-purifying proteins from the ClpS column. (b) Coomassie-stained polyacrylamide gels showing isolation of putative ClpS-binding proteins from ClpS-Affigel column. Free ClpS was used to elute bound proteins, and only the three elution fractions from the ClpS and beads-only column are shown. (c) Anti-ClpA western showing the emergence of ClpA from the ClpS-H_6 or beads-only column in a pull-down experiment. See (a) for lane labels.

In both experiments, *E. coli* cellular lysates were applied to a solid support matrix bound with ClpS followed by washing and elution. Elution samples were precipitated with trichloroacetic acid prior to loading on the gel (Fig. A.2a,b). Protein indeed eluted from the ClpS columns, but the beads-only column also yielded eluted proteins. These contaminating proteins in the beads-only column persisted in the washing steps, and a more stringent washing might reduce these levels.

We considered the possibility that extensive washing might remove potential binding proteins that are transiently bound to ClpS. Therefore, we monitored the binding of ClpA as a control between the ClpS column and a beads-only column. Figure A.2c shows the presence of ClpA in wash and elution fractions from the ClpS and beads-only columns. As expected, most of the ClpA is present in the lysate flow-through and first wash fraction in the beads-only column. In the ClpS column, ClpA is bound to the column and co-eluted in the first elution fraction. However, a substantial amount of ClpA is removed from the column in the flow-though as well as the wash fractions. Therefore, instead of more stringent washing, we decided to identify the proteins eluted in the beads-only column and label them as contaminants.

Elution fractions from both the ClpS column and the beads-only were run on SDS-PAGE, and the gel slices contained eluted protein were excised for trypsinization and mass spectrometric analysis. Many proteins that eluted from all columns were ribosomal proteins. Even though ribosomal proteins may bind specifically to ClpS, we decided to disregard these proteins as contaminants due to their sheer abundance in the beads-only elutions. Not unexpectedly, proteins with consecutive histidines were eluted from both the ClpS and beads-only Ni-NTA columns. These histidine-rich proteins were also dismissed as contaminants.

The proteins AphA, ClpA, MreB, and MutM were identified in the ClpS column and not the beads-only column for both matrices. All of these proteins were identified in the SDS-PAGE gel slice that approximately corresponds to their full molecular weight, indicating that these proteins were not substantially cleaved prior to being recognized for binding by ClpS. As described in Chapter Four, identification of multiple unique peptides by mass spectrometry increases the confidence that the identification is not erroneous, keeping into consideration that proteins of higher mass will yield more

peptides. ClpA was identified by 12 unique peptides in a ClpS-H_6 column and 10 unique peptides in the covalently-linked ClpS column. AphA was identified by 1 and 3 peptides, MreB was identified by 2 peptides each, and MutM was identified by 1 unique peptide in the Ni-NTA and Affigel columns, respectively. With many peptides identified, ClpA serves as a good positive control for the ClpS column, showing that a binding partner can recognize ClpS tethered to a solid support. Conversely, AphA, MreB, and MutM were identified by few peptides. Because the vast majority of the proteins identified by mass spectrometry were either identified in the beads-only column or were not seen in a second mass spectrometry experiment, we sought a more specific way to isolate ClpS-binding proteins before protein-specific validation.

Due to the role of ClpS as an adaptor, we assumed that most proteins that may potentially bind to ClpS, with the exception of ClpA, are transiently bound. An isolated N-degron bound to *E. coli* ClpS with a K_D of ~6-7 μM (Chapter Three). It would therefore be difficult to specifically identify these proteins amongst potential contaminants under wash and purification conditions that did not perturb true binding interactions. Therefore, a cross-linking approach was used with the intention of covalently linking interacting proteins to ClpS followed by a more extensive purification of ClpS. Although we were interested in any new interacting protein, we focused specifically on seeking novel N-end-rule substrates. We used non-site-specific cross-linkers in conjunction with ClpS-H_6. To ensure that the isolated proteins are N-degron substrates, we duplicated each experiment with ClpS-H_6 D35G D35A ("GA"), a mutant that does not bind N-degrons (Erbse et al. 2006). Proteins pulled-down by wild-type ClpS and not the GA mutants would be considered for further validation.

A control experiment was carried out testing the ability of ClpS-H_6 and ClpS-H_6 GA to pull down a purified N-end-rule substrate added to cellular lysates with several cross-linkers. YLFVR-GFP was enriched by wild-type ClpS but not by the GA mutant (Fig. A.3a). Furthermore, both ClpS and GFP were detected at higher molecular weights, indicating that ClpS and GFP cross-linked to form higher order complexes. This control shows that ClpS-H_6 attached to Ni-NTA beads can bind an N-end-rule protein and that the GA mutant prevents this interaction.

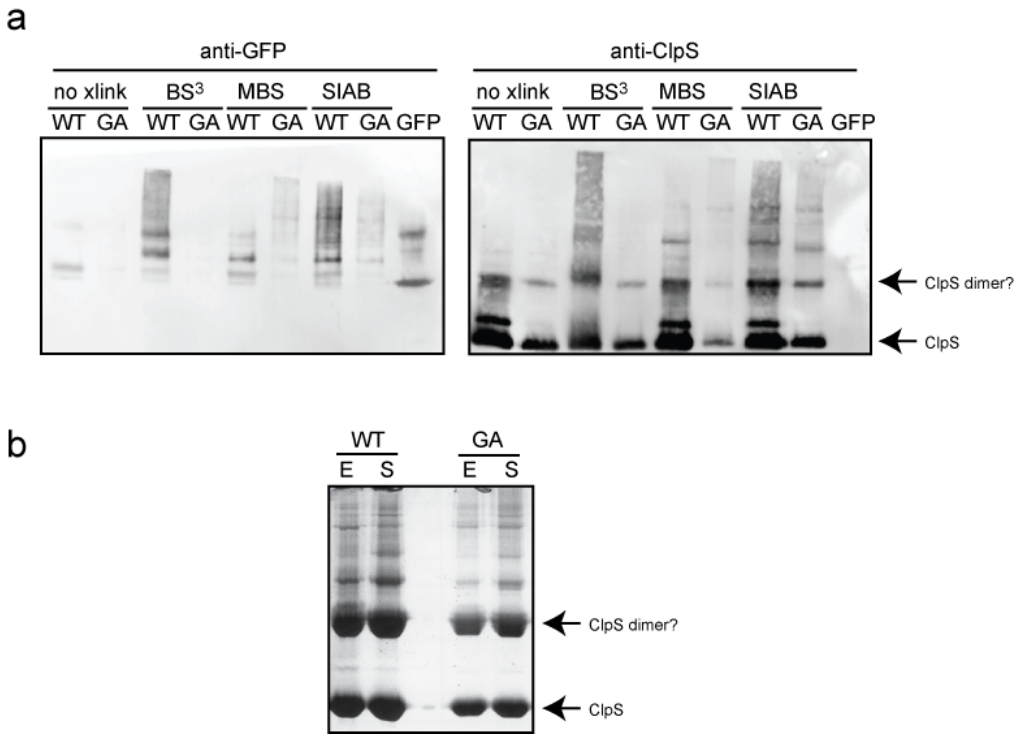


Figure A.3. ClpS co-precipitation using cross-linkers.
(a) Control experiment showing the preferential binding of YLFVR-GFP to wild-type ClpS and not ClpS GA. WT = wild-type ClpS-H_6, GA = ClpS-H_6 D35G D36A. YLFVR-GFP was cross-linked to wild-type ClpS and formed higher molecular weight complexes, as determined by the YA_4-GPF-only lane on the right-hand side of the blots. YA_4-GPF did not cross-react with the anti-ClpS antibody. Most of the ClpS did not cross-link with YLFVR-GFP as shown by migration of a monomeric ClpS band. Insufficient boiling prior to electrophoresis may have resulted in ClpS dimer formation on the gel. (b) Elution fractions from a sulfo-SIAB cross-linking experiment using ClpS-H_6 and ClpS-H_6 GA expressed from a plasmid *in vivo*. E = exponential phase cell lysates, S = stationary phase cell lysates. Insufficient boiling prior to electrophoresis may have resulted in ClpS dimer formation on the gel. Peptides from this gel were extracted and analyzed by mass spectrometry.

Unclear differences between wild-type ClpS and ClpS GA column elutions were seen in several experiments attempting to pull-down potential binding partners from cellular lysate. We therefore, decided to express ClpS-H_6 *in vivo* to maximize binding. Cross-linker was added to the cell lysates prior to isolation of ClpS-H_6 and bound proteins, and ClpS was purified under denaturing conditions to remove as many contaminants as possible. The differences between wild-type ClpS and ClpS GA are still unclear, but we surmised that potential binding partners might be few and undetectable by eye (Fig. A.3b). We analyzed some of the eluted proteins by mass spectrometry and

generated a list of potential proteins to test. Some proteins that were found in wild-type elutions and not GA elutions that were also detected by more than one peptide included: AceA, AhpC, DsbC, GlmU, GsaE/HemL, HldE/RfaE, IscR/YfhP, MprA, NagD, PtsI, and YdaO. We cloned these genes into a vector encoding an N-terminal His_6 and a C-terminal FLAG tag. If these genes were processed *in vivo* by a cleavage reaction, each end of the protein would still bear an epitope tag for detection. We aimed to conduct ClpS far-western experiments to determine if ClpS could bind to the proteins of interest, and if so, if ClpS could bind a proteolytically processed form that might bear an N-degron. These experiments were put on hold due to technical problems and should be addressed in the future.

Because these studies have not been completed, there is still ample possibility of detecting and identifying ClpS binding proteins. For instance, a more stringent protocol for isolating ClpS-H_6 cross-linked to proteins would assist in the specific identification of N-end-rule substrates. The cross-linking concentration and time can be adjusted to reduce the cross-linking of contaminants to ClpS, and the wash procedure can be made harsher to lower the level of contaminant binding.

Validation of potential binding partners identified by mass spectrometry is more complex. ClpS has been shown to bind to N-degrons by far-western blots (Erbse et al. 2006), but the overexpression of our candidate proteins might not permit sufficient processing *in vivo* for N-degron exposure. Solution-based assays may also be required to test for binding. Additional challenges to validation may include low cellular abundance of N-degrons for detection and the lack of knowledge of potential upstream processing steps.

MATERIALS AND METHODS

Proteins

The *clpS* gene was amplified and cloned into a pET23b overexpression vector (Novagen). Protein was overexpressed with 0.4 mM IPTG for 2 hours after which the cells were harvested. The cell lysate was cleared with 35% saturated ammonium sulfate followed by resuspension of the spun pellet into S1 buffer (50 mM sodium phosphate pH 8, 300 mM NaCl, 5 mM imidazole, 10% glycerol). The sample was then purified over

Ni-NTA resin (Qiagen), washed, and eluted in high imidazole buffer (S1 with 500 mM imidazole).

ClpS-H_6 D35G D36A was generated from ClpS-H_6 by QuikChange mutagenesis (Stratagene) and purified similarly.

Protein degradation

ClpS-H_6 was tested for ability to degrade GFP-ssrA with 0.1 μM $ClpA_6$, 0.27 μM $ClpP_{14}$, 1 μM GFP-ssrA, 0.2 or 0.6 μM ClpS protein, 4 mM ATP, and an ATP regeneration system (0.05 mg/mL creatine kinase and 2.5 mM creatine phosphate). Loss of fluorescence was measured on a Photon Technology International fluorimeter.

Initial pull-downs

For ClpS-H_6 binding assays, *E. coli* cell cultures grown to either exponential or stationary phase were harvested and lysed. About 0.2 μmol of ClpS-H_6 was bound to 250 μL Ni-NTA beads (Qiagen). The cell lysates were added to the ClpS-H_6 column and a beads-only column, and the samples were washed with S1 buffer containing 10 mM or 20 mM imidazole, and eluted in high imidazole buffer.

For ClpS-Affigel binding assays, 2.4 mg ClpS was bound to washed Affigel-15 beads (Bio-Rad) overnight. The next morning, unreacted ClpS was removed and the remaining reactive groups were quenched with ethanolamine pH. 8. *E. coli* cultures at exponential or stationary phase were harvested after which the spun and filtered lysate was applied to the washed ClpS column. After binding, the column was washed and eluted with excess competitor ClpS.

For both co-precipitation experiments, the elutions were TCA-precipitated, samples run on SDS-PAGE, bands excised, proteins trypsinized, and peptides analyzed by mass spectrometry.

Cross-linking pull-downs

For the YLFVR-GFP cross-linking experiment, 2.5 nmol of ClpS-H_6 or ClpS-H_6 GA was added to 100 μL lysates from exponentially-grown *E. coli* to a total of 200 μL. ~0.1 mg BS^3 (Bis(sulfosuccinimidyl)suberate, homo-bifunctional amine-reactive), sulfo-

MBS (m-Maleimidobenzoyl-N-hydroxysulfosuccinimide ester, heterobifunctional amine- and sulfhydryl-reactive), or sulfo-SIAB (*N*-Sulfosuccinimidyl[4-iodoacetyl]aminobenzoate, heterobifunctional amine- and sulfhydryl-reactive) cross-linker was added to the lysate at 4 ºC for 1 hour after which the sample was added to Ni-NTA resin. The slurry was pipetted into a spin column (Bio-Rad) and washed and eluted by centrifugation. Elutions from all samples were run on SDS-PAGE followed by an anti-ClpS or anti-GFP western to detect the formation of higher molecular weight complexes. Purified YA_4-GFP was loaded onto the gel as a GFP size standard.

For "*in vivo*" cross-linking experiments, ClpS-H_6 and ClpS-H_6 GA were expressed from a pTrc99a plasmid with 0.2 mM IPTG in W3110 *clpA::kan* cells grown to either exponential or stationary phase. Sulfo-SIAB was added to the lysate after which the ClpS proteins were purified by binding to Ni-NTA and washed and eluted under 8 M urea denaturing conditions. After running the samples on SDS-PAGE, gel bands were excised, and the proteins were trypsinized and analyzed by tandem mass spectrometry for identification.

ClpS far westerns

The genes for putative ClpS-binding proteins were amplified and cloned into pTrc99a vectors with an N-terminal His_6 tag and a C-terminal FLAG tag. These plasmids were then transformed into W3110 *clpA::kan* and induced for expression. Cell samples were taken, run on SDS-PAGE, and probed with 50 nM ClpS-H_6 followed by an anti-His or anti-ClpS antibody.

REFERENCES

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

Aizenman, E., H. Engelberg-Kulka and G. Glaser (1996). "An Escherichia coli chromosomal "addiction module" regulated by guanosine [corrected] 3',5'-bispyrophosphate: a model for programmed bacterial cell death." Proc Natl Acad Sci U S A **93**(12): 6059-63.

Almiron, M., A. J. Link, D. Furlong and R. Kolter (1992). "A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli." Genes Dev **6**(12B): 2646-54.

Bachmair, A., D. Finley and A. Varshavsky (1986). "In vivo half-life of a protein is a function of its amino-terminal residue." Science **234**(4773): 179-86.

Bachmann, B. J. (1972). "Pedigrees of some mutant strains of Escherichia coli K-12." Bacteriol Rev **36**(4): 525-57.

Bajorek, M. and M. H. Glickman (2004). "Keepers at the final gates: regulatory complexes and gating of the proteasome channel." Cell Mol Life Sci **61**(13): 1579-88.

Baker, T. A. and R. T. Sauer (2006). "ATP-dependent proteases of bacteria: recognition logic and operating principles." Trends Biochem Sci **31**(12): 647-53.

Ball, C. A., R. Osuna, K. C. Ferguson and R. C. Johnson (1992). "Dramatic changes in Fis levels upon nutrient upshift in Escherichia coli." J Bacteriol **174**(24): 8043-56.

Bantscheff, M., M. Schirle, G. Sweetman, J. Rick and B. Kuster (2007). "Quantitative mass spectrometry in proteomics: a critical review." Anal Bioanal Chem **389**(4): 1017-31.

Beuron, F., M. R. Maurizi, D. M. Belnap, E. Kocsis, F. P. Booy, M. Kessel and A. C. Steven (1998). "At sixes and sevens: characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease." J Struct Biol **123**(3): 248-59.

Bolon, D. N., D. A. Wah, G. L. Hersch, T. A. Baker and R. T. Sauer (2004). "Bivalent tethering of SspB to ClpXP is required for efficient substrate delivery: a protein-design study." Mol Cell **13**(3): 443-9.

Bougdour, A., C. Cunning, P. J. Baptiste, T. Elliott and S. Gottesman (2008). "Multiple pathways for regulation of sigmaS (RpoS) stability in Escherichia coli via the action of multiple anti-adaptors." Mol Microbiol **68**(2): 298-313.

Bougdour, A., S. Wickner and S. Gottesman (2006). "Modulating RssB activity: IraP, a novel regulator of sigma(S) stability in Escherichia coli." Genes Dev **20**(7): 884-97.

Burton, B. M. and T. A. Baker (2003). "Mu transpososome architecture ensures that unfolding by ClpX or proteolysis by ClpXP remodels but does not destroy the complex." Chem Biol **10**(5): 463-72.

Chaba, R., I. L. Grigorova, J. M. Flynn, T. A. Baker and C. A. Gross (2007). "Design principles of the proteolytic cascade governing the sigmaE-mediated envelope stress response in Escherichia coli: keys to graded, buffered, and rapid signal transduction." Genes Dev **21**(1): 124-36.

Chang, A. C. and S. N. Cohen (1978). "Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid." J Bacteriol **134**(3): 1141-56.

Clamp, M., J. Cuff, S. M. Searle and G. J. Barton (2004). "The Jalview Java alignment editor." Bioinformatics **20**(3): 426-7.

Cowan, P. M. and S. McGavin (1955). "Structure of Poly-L-Proline " Nature **176**: 501-503.

Cranz-Mileva, S., F. Imkamp, K. Kolygo, Z. Maglica, W. Kress and E. Weber-Ban (2008). "The flexible attachment of the N-domains to the ClpA ring body allows their use on demand." J Mol Biol **378**(2): 412-24.

Damerau, K. and A. C. St John (1993). "Role of Clp protease subunits in degradation of carbon starvation proteins in Escherichia coli." J Bacteriol **175**(1): 53-63.

Dougan, D. A., B. G. Reid, A. L. Horwich and B. Bukau (2002). "ClpS, a substrate modulator of the ClpAP machine." Mol Cell **9**(3): 673-83.

Dougan, D. A., E. Weber-Ban and B. Bukau (2003). "Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX." Mol Cell **12**(2): 373-80.

Engelberg-Kulka, H., R. Hazan and S. Amitai (2005). "mazEF: a chromosomal toxin-antitoxin module that triggers programmed cell death in bacteria." J Cell Sci **118**(Pt 19): 4327-32.

Erbse, A., R. Schmidt, T. Bornemann, J. Schneider-Mergener, A. Mogk, R. Zahn, D. A. Dougan and B. Bukau (2006). "ClpS is an essential component of the N-end rule pathway in Escherichia coli." Nature **439**(7077): 753-6.

Erbse, A. H., J. N. Wagner, K. N. Truscott, S. K. Spall, J. Kirstein, K. Zeth, K. Turgay, A. Mogk, B. Bukau and D. A. Dougan (2008). "Conserved residues in the N-domain of the AAA+ chaperone ClpA regulate substrate recognition and unfolding." Febs J **275**(7): 1400-10.

Farrell, C. M., T. A. Baker and R. T. Sauer (2007). "Altered specificity of a AAA+ protease." Mol Cell **25**(1): 161-6.

Farrell, C. M., A. D. Grossman and R. T. Sauer (2005). "Cytoplasmic degradation of ssrA-tagged proteins." Mol Microbiol **57**(6): 1750-61.

Flynn, J. M., I. Levchenko, R. T. Sauer and T. A. Baker (2004). "Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation." Genes Dev **18**(18): 2292-301.

Flynn, J. M., I. Levchenko, M. Seidel, S. H. Wickner, R. T. Sauer and T. A. Baker (2001). "Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis." Proc Natl Acad Sci U S A **98**(19): 10584-9.

Flynn, J. M., S. B. Neher, Y. I. Kim, R. T. Sauer and T. A. Baker (2003). "Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals." Mol Cell **11**(3): 671-83.

Fox, G. and J. W. Brown (1961). "Protein degradation in Escherichia coli in the logarithmic phase of growth." Biochim Biophys Acta **46**: 387-9.

Glotzer, M., A. W. Murray and M. W. Kirschner (1991). "Cyclin is degraded by the ubiquitin pathway." Nature **349**(6305): 132-8.

Gottesman, S. (2003). "Proteolysis in bacterial regulatory circuits." Annu Rev Cell Dev Biol **19**: 565-87.
Gottesman, S., W. P. Clark and M. R. Maurizi (1990). "The ATP-dependent Clp protease of Escherichia coli. Sequence of clpA and identification of a Clp-specific substrate." J Biol Chem **265**(14): 7886-93.
Gottesman, S. and M. R. Maurizi (1992). "Regulation by proteolysis: energy-dependent proteases and their targets." Microbiol Rev **56**(4): 592-621.
Gottesman, S., E. Roche, Y. Zhou and R. T. Sauer (1998). "The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system." Genes Dev **12**(9): 1338-47.
Guo, F., L. Esser, S. K. Singh, M. R. Maurizi and D. Xia (2002a). "Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA." J Biol Chem **277**(48): 46753-62.
Guo, F., M. R. Maurizi, L. Esser and D. Xia (2002b). "Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease." J Biol Chem **277**(48): 46743-52.
Guo, H. and J. R. Ecker (2003). "Plant responses to ethylene gas are mediated by SCF(EBF1/EBF2)-dependent proteolysis of EIN3 transcription factor." Cell **115**(6): 667-77.
Halvorson, H. (1958a). "Intracellular protein and nucleic acid turnover in resting yeast cells." Biochim Biophys Acta **27**(2): 255-66.
Halvorson, H. (1958b). "Studies on protein and nucleic acid turnover in growing cultures of yeast." Biochim Biophys Acta **27**(2): 267-76.
Hanson, P. I. and S. W. Whiteheart (2005). "AAA+ proteins: have engine, will work." Nat Rev Mol Cell Biol **6**(7): 519-29.
Heessen, S., M. G. Masucci and N. P. Dantuma (2005). "The UBA2 domain functions as an intrinsic stabilization signal that protects Rad23 from proteasomal degradation." Mol Cell **18**(2): 225-35.
Hershko, A. and A. Ciechanover (1998). "The ubiquitin system." Annu Rev Biochem **67**: 425-79.
Hinnerwisch, J., W. A. Fenton, K. J. Furtak, G. W. Farr and A. L. Horwich (2005a). "Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation." Cell **121**(7): 1029-41.
Hinnerwisch, J., B. G. Reid, W. A. Fenton and A. L. Horwich (2005b). "Roles of the N-domains of the ClpA unfoldase in binding substrate proteins and in stable complex formation with the ClpP protease." J Biol Chem **280**(49): 40838-44.
Hirel, P. H., M. J. Schmitter, P. Dessen, G. Fayat and S. Blanquet (1989). "Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the side-chain length of the penultimate amino acid." Proc Natl Acad Sci U S A **86**(21): 8247-51.
Hoskins, J. R., K. Yanagihara, K. Mizuuchi and S. Wickner (2002). "ClpAP and ClpXP degrade proteins with tags located in the interior of the primary sequence." Proc Natl Acad Sci U S A **99**(17): 11037-42.
Hou, J. Y., R. T. Sauer and T. A. Baker (2008). "Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP." Nat Struct Mol Biol **15**(3): 288-94.

Huggins, M. L. (1943). "The Structure of Fibrous Proteins." Chem Rev **32**(2): 195-218.
Hwang, B. J., W. J. Park, C. H. Chung and A. L. Goldberg (1987). "Escherichia coli contains a soluble ATP-dependent protease (Ti) distinct from protease La." Proc Natl Acad Sci U S A **84**(16): 5550-4.
Inobe, T. and A. Matouschek (2008). "Protein targeting to ATP-dependent proteases." Curr Opin Struct Biol **18**(1): 43-51.
Ishihama, Y., T. Schmidt, J. Rappsilber, M. Mann, F. U. Hartl, M. J. Kerner and D. Frishman (2008). "Protein abundance profiling of the Escherichia coli cytosol." BMC Genomics **9**: 102.
Ishikawa, T., M. R. Maurizi and A. C. Steven (2004). "The N-terminal substrate-binding domain of ClpA unfoldase is highly mobile and extends axially from the distal surface of ClpAP protease." J Struct Biol **146**(1-2): 180-8.
Jenal, U. and T. Fuchs (1998). "An essential protease involved in bacterial cell-cycle control." Embo J **17**(19): 5658-69.
Jennings, L. D., J. Bohon, M. R. Chance and S. Licht (2008). "The ClpP N-terminus coordinates substrate access with protease active site reactivity." Biochemistry **47**(42): 11031-40.
Karzai, A. W., E. D. Roche and R. T. Sauer (2000). "The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue." Nat Struct Biol **7**(6): 449-55.
Katayama-Fujimura, Y., S. Gottesman and M. R. Maurizi (1987). "A multiple-component, ATP-dependent protease from Escherichia coli." J Biol Chem **262**(10): 4477-85.
Katayama, Y., S. Gottesman, J. Pumphrey, S. Rudikoff, W. P. Clark and M. R. Maurizi (1988). "The two-component, ATP-dependent Clp protease of Escherichia coli. Purification, cloning, and mutational analysis of the ATP-binding component." J Biol Chem **263**(29): 15226-36.
Keiler, K. C., P. R. Waller and R. T. Sauer (1996). "Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA." Science **271**(5251): 990-3.
Kenniston, J. A., T. A. Baker, J. M. Fernandez and R. T. Sauer (2003). "Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine." Cell **114**(4): 511-20.
Kim, Y. I., R. E. Burton, B. M. Burton, R. T. Sauer and T. A. Baker (2000). "Dynamics of substrate denaturation and translocation by the ClpXP degradation machine." Mol Cell **5**(4): 639-48.
Kim, Y. I., I. Levchenko, K. Fraczkowska, R. V. Woodruff, R. T. Sauer and T. A. Baker (2001). "Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase." Nat Struct Biol **8**(3): 230-3.
Kirstein, J., T. Schlothauer, D. A. Dougan, H. Lilie, G. Tischendorf, A. Mogk, B. Bukau and K. Turgay (2006). "Adaptor protein controlled oligomerization activates the AAA+ protein ClpC." Embo J **25**(7): 1481-91.
Klauck, E., M. Lingnau and R. Hengge-Aronis (2001). "Role of the response regulator RssB in sigma recognition and initiation of sigma proteolysis in Escherichia coli." Mol Microbiol **40**(6): 1381-90.

Levchenko, I., R. A. Grant, J. M. Flynn, R. T. Sauer and T. A. Baker (2005). "Versatile modes of peptide recognition by the AAA+ adaptor protein SspB." Nat Struct Mol Biol **12**(6): 520-5.

Levchenko, I., R. A. Grant, D. A. Wah, R. T. Sauer and T. A. Baker (2003). "Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag." Mol Cell **12**(2): 365-72.

Levchenko, I., L. Luo and T. A. Baker (1995). "Disassembly of the Mu transposase tetramer by the ClpX chaperone." Genes Dev **9**(19): 2399-408.

Levchenko, I., M. Seidel, R. T. Sauer and T. A. Baker (2000). "A specificity-enhancing factor for the ClpXP degradation machine." Science **289**(5488): 2354-6.

Lies, M. and M. R. Maurizi (2008). "Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in Escherichia coli." J Biol Chem **283**(34): 22918-29.

Lo, J. H., T. A. Baker and R. T. Sauer (2001). "Characterization of the N-terminal repeat domain of Escherichia coli ClpA-A class I Clp/HSP100 ATPase." Protein Sci **10**(3): 551-9.

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

Maglica, Z., F. Striebel and E. Weber-Ban (2008). "An intrinsic degradation tag on the ClpA C-terminus regulates the balance of ClpAP complexes with different substrate specificity." J Mol Biol **384**(2): 503-11.

Martin, A., T. A. Baker and R. T. Sauer (2008a). "Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates." Mol Cell **29**(4): 441-50.

Martin, A., T. A. Baker and R. T. Sauer (2008b). "Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding." Nat Struct Mol Biol **15**(11): 1147-51.

Mogk, A., D. Dougan, J. Weibezahn, C. Schlieker, K. Turgay and B. Bukau (2004). "Broad yet high substrate specificity: the challenge of AAA+ proteins." J Struct Biol **146**(1-2): 90-8.

Nagashima, K., Y. Kubota, T. Shibata, C. Sakaguchi, H. Shinagawa and T. Hishida (2006). "Degradation of Escherichia coli RecN aggregates by ClpXP protease and its implications for DNA damage tolerance." J Biol Chem **281**(41): 30941-6.

Nair, S. and S. E. Finkel (2004). "Dps protects cells against multiple stresses during stationary phase." J Bacteriol **186**(13): 4192-8.

Neher, S. B., R. T. Sauer and T. A. Baker (2003). "Distinct peptide signals in the UmuD and UmuD' subunits of UmuD/D' mediate tethering and substrate processing by the ClpXP protease." Proc Natl Acad Sci U S A **100**(23): 13219-24.

Neher, S. B., J. Villen, E. C. Oakes, C. E. Bakalarski, R. T. Sauer, S. P. Gygi and T. A. Baker (2006). "Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon." Mol Cell **22**(2): 193-204.

Neidhardt, F. C. and R. Curtiss (1996). Escherichia coli and Salmonella : cellular and molecular biology. Washington, D.C., ASM Press.

Neuwald, A. F., L. Aravind, J. L. Spouge and E. V. Koonin (1999). "AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes." Genome Res **9**(1): 27-43.

Nigg, E. A. (1995). "Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle." Bioessays **17**(6): 471-80.

Okuno, T., K. Yamanaka and T. Ogura (2006). "Characterization of mutants of the Escherichia coli AAA protease, FtsH, carrying a mutation in the central pore region." J Struct Biol **156**(1): 109-14.

Ong, S. E., B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey and M. Mann (2002). "Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics." Mol Cell Proteomics **1**(5): 376-86.

Ortega, J., S. K. Singh, T. Ishikawa, M. R. Maurizi and A. C. Steven (2000). "Visualization of substrate binding and translocation by the ATP-dependent protease, ClpXP." Mol Cell **6**(6): 1515-21.

Osuna, R., D. Lienau, K. T. Hughes and R. C. Johnson (1995). "Sequence, regulation, and functions of fis in Salmonella typhimurium." J Bacteriol **177**(8): 2021-32.

Park, E., Y. M. Rho, O. J. Koh, S. W. Ahn, I. S. Seong, J. J. Song, O. Bang, J. H. Seol, J. Wang, S. H. Eom and C. H. Chung (2005). "Role of the GYVG pore motif of HslU ATPase in protein unfolding and translocation for degradation by HslV peptidase." J Biol Chem **280**(24): 22892-8.

Pickart, C. M. and R. E. Cohen (2004). "Proteasomes and their kin: proteases in the machine age." Nat Rev Mol Cell Biol **5**(3): 177-87.

Piszczek, G., J. Rozycki, S. K. Singh, A. Ginsburg and M. R. Maurizi (2005). "The molecular chaperone, ClpA, has a single high affinity peptide binding site per hexamer." J Biol Chem **280**(13): 12221-30.

Prakash, S., L. Tian, K. S. Ratliff, R. E. Lehotzky and A. Matouschek (2004). "An unstructured initiation site is required for efficient proteasome-mediated degradation." Nat Struct Mol Biol **11**(9): 830-7.

Pruteanu, M., S. B. Neher and T. A. Baker (2007). "Ligand-controlled proteolysis of the Escherichia coli transcriptional regulator ZntR." J Bacteriol **189**(8): 3017-25.

Rao, H., F. Uhlmann, K. Nasmyth and A. Varshavsky (2001). "Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability." Nature **410**(6831): 955-9.

Reid, B. G., W. A. Fenton, A. L. Horwich and E. U. Weber-Ban (2001). "ClpA mediates directional translocation of substrate proteins into the ClpP protease." Proc Natl Acad Sci U S A **98**(7): 3768-72.

Schlothauer, T., A. Mogk, D. A. Dougan, B. Bukau and K. Turgay (2003). "MecA, an adaptor protein necessary for ClpC chaperone activity." Proc Natl Acad Sci U S A **100**(5): 2306-11.

Schmidt, M., J. Hanna, S. Elsasser and D. Finley (2005). "Proteasome-associated proteins: regulation of a proteolytic machine." Biol Chem **386**(8): 725-37.

Schmidt, R., R. Zahn, B. Bukau and A. Mogk (2009). "ClpS is the recognition component for Escherichia coli substrates of the N-end rule degradation pathway." Mol Microbiol.

Seol, J. H., S. H. Baek, M. S. Kang, D. B. Ha and C. H. Chung (1995). "Distinctive roles of the two ATP-binding sites in ClpA, the ATPase component of protease Ti in Escherichia coli." J Biol Chem **270**(14): 8087-92.

Shrader, T. E., J. W. Tobias and A. Varshavsky (1993). "The N-end rule in Escherichia coli: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene aat." J Bacteriol **175**(14): 4364-74.

Siddiqui, S. M., R. T. Sauer and T. A. Baker (2004). "Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates." Genes Dev **18**(4): 369-74.

Singh, S. K. and M. R. Maurizi (1994). "Mutational analysis demonstrates different functional roles for the two ATP-binding sites in ClpAP protease from Escherichia coli." J Biol Chem **269**(47): 29537-45.

Singh, S. K., J. Rozycki, J. Ortega, T. Ishikawa, J. Lo, A. C. Steven and M. R. Maurizi (2001). "Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis." J Biol Chem **276**(31): 29420-9.

Susi, H., S. N. Timasheff and L. Stevens (1967). "Infrared spectra and protein conformations in aqueous solutions. I. The amide I band in H2O and D2O solutions." J Biol Chem **242**(23): 5460-6.

Tian, L., R. A. Holmgren and A. Matouschek (2005). "A conserved processing mechanism regulates the activity of transcription factors Cubitus interruptus and NF-kappaB." Nat Struct Mol Biol **12**(12): 1045-53.

Tobias, J. W., T. E. Shrader, G. Rocap and A. Varshavsky (1991). "The N-end rule in bacteria." Science **254**(5036): 1374-7.

Traub, W. and U. Shmueli (1963). "Structure of Poly-L-Proline I " Nature **198**: 1165-1166.

Vale, R. D. (2000). "AAA proteins. Lords of the ring." J Cell Biol **150**(1): F13-9.

Varshavsky, A. (1995). "The N-end rule." Cold Spring Harb Symp Quant Biol **60**: 461-78.

Varshavsky, A. (1996). "The N-end rule: functions, mysteries, uses." Proc Natl Acad Sci U S A **93**(22): 12142-9.

Wah, D. A., I. Levchenko, T. A. Baker and R. T. Sauer (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**(11): 1237-45.

Wah, D. A., I. Levchenko, G. E. Rieckhof, D. N. Bolon, T. A. Baker and R. T. Sauer (2003). "Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the AAA+ ClpXP protease." Mol Cell **12**(2): 355-63.

Wang, J. (2001). "A corrected quaternary arrangement of the peptidase HslV and atpase HslU in a cocrystal structure." J Struct Biol **134**(1): 15-24.

Wang, K. H., E. S. Oakes, R. T. Sauer and T. A. Baker (2008a). "Tuning the strength of a bacterial N-end rule degradation signal." J Biol Chem **283**(36): 24600-7.

Wang, K. H., G. Roman-Hernandez, R. A. Grant, R. T. Sauer and T. A. Baker (2008b). "The molecular basis of N-end rule recognition." Mol Cell **32**(3): 406-14.

Wang, K. H., R. T. Sauer and T. A. Baker (2007). "ClpS modulates but is not essential for bacterial N-end rule degradation." Genes Dev **21**(4): 403-8.

Weber-Ban, E. U., B. G. Reid, A. D. Miranker and A. L. Horwich (1999). "Global unfolding of a substrate protein by the Hsp100 chaperone ClpA." Nature **401**(6748): 90-3.

Weichart, D., N. Querfurth, M. Dreger and R. Hengge-Aronis (2003). "Global role for ClpP-containing proteases in stationary-phase adaptation of Escherichia coli." J Bacteriol **185**(1): 115-25.

Wojtyra, U. A., G. Thibault, A. Tuite and W. A. Houry (2003). "The N-terminal zinc binding domain of ClpX is a dimerization domain that modulates the chaperone function." J Biol Chem **278**(49): 48981-90.

Woo, K. M., W. J. Chung, D. B. Ha, A. L. Goldberg and C. H. Chung (1989). "Protease Ti from Escherichia coli requires ATP hydrolysis for protein breakdown but not for hydrolysis of small peptides." J Biol Chem **264**(4): 2088-91.

Xia, D., L. Esser, S. K. Singh, F. Guo and M. R. Maurizi (2004). "Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone." J Struct Biol **146**(1-2): 166-79.

Yamada-Inagawa, T., T. Okuno, K. Karata, K. Yamanaka and T. Ogura (2003). "Conserved pore residues in the AAA protease FtsH are important for proteolysis and its coupling to ATP hydrolysis." J Biol Chem **278**(50): 50182-7.

Zeth, K., R. B. Ravelli, K. Paal, S. Cusack, B. Bukau and D. A. Dougan (2002). "Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA." Nat Struct Biol **9**(12): 906-11.

Zhou, Y., S. Gottesman, J. R. Hoskins, M. R. Maurizi and S. Wickner (2001). "The RssB response regulator directly targets sigma(S) for degradation by ClpXP." Genes Dev **15**(5): 627-37.

Zwickl, P., W. Baumeister and A. Steven (2000). "Dis-assembly lines: the proteasome and related ATPase-assisted proteases." Curr Opin Struct Biol **10**(2): 242-50.