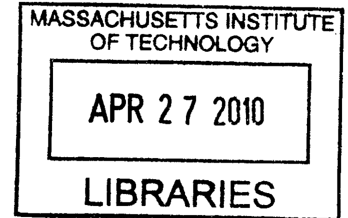
Interaction between the AAA+ protease ClpXP and the adaptor protein SspB

by

Tahmeena Chowdhury

B.A. Biochemistry
Smith College, 2004


MASSACHUSETTS INSTITUTE OF TECHNOLOGY
APR 27 2010
LIBRARIES

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

ARCHIVES

JUNE 2010

© Tahmeena Chowdhury. All Rights Reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author: ______________________________
Department of Biology
May 21, 2010

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

Interaction between the AAA+ protease ClpXP and the adaptor protein SspB

by

Tahmeena Chowdhury

Submitted to the Department of Biology on May 21, 2010 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy at the Massachusetts Institute of Technology

ABSTRACT

Proteolysis plays a vital role in cellular processes including regulatory pathways and protein quality control in prokaryotes and eukaryotes. ATP-dependent protein degradation is mediated by multimeric protease complexes, each consisting of a AAA+ ATPase and a peptidase component. Substrate selection by the proteases is a highly regulated process to ensure minimal errant protein degradation. Substrates are usually recognized by proteases through degradation tags or degrons. Accessory proteins called adaptors can also modulate substrate selection by proteases. These adaptors have the potential to affect substrate specificity as well as expand the repertoire of substrates that can be degraded by proteases. Understanding how proteases interact with a wide range of adaptors and substrates can provide valuable insight into the complex process of substrate selection.

In this thesis, I have investigated the interaction between the AAA+ protease ClpXP and the adaptor protein SspB. The highly conserved N-terminal domain of the unfoldase ClpX interacts with SspB and other specific adaptor proteins and substrates. However, these binding partners do not use one simple sequence motif to mediate the protein-protein interaction. This diversity in protein-binding was further demonstrated by the cross-species ClpX-SspB interactions in *Caulobacter crescentus* and *Escherichia coli*. Despite little sequence homology, *C. crescentus* SspB (^{Cc}SspBα) and *E. coli* SspB (^{Ec}SspB) are able to interact with ClpX from either species. We analyzed these interactions to understand how the N-terminal domain of ClpX is able to recognize diverse adaptors and substrates while still retaining specificity. We identified the region important for interaction of ^{Cc}SspBα with ClpX. Mutagenesis studies of the C-terminal region of the adaptor were conducted and the variants were tested for their ability to functionally interact with the ClpXP protease. Using these data and the results of peptide-binding experiments, we identified residues within the C-terminal region of ^{Cc}SspBα that are important for tethering to ClpX. We also conducted functional and peptide-binding studies on the ^{Ec}SspB ClpX-binding (XB) region. Interestingly, the two XB regions are very different in both length and sequence. However, despite this dissimilarity, competition studies argue that the two XB peptides bind to identical or overlapping sites of the ClpX N domains of *C. crescentus* and *E. coli*. This cross-species interaction between SspB and ClpX highlights how the ClpX N domain provides a versatile platform for binding a variety of adaptors and substrates.

We also performed a proteomic-screen to investigate the effect of SspB on ClpXP substrate profile in *E. coli*. The preliminary data has provided a list of candidate SspB-interacting substrates, further analyses of which will contribute to the understanding of the biological impact of SspB on substrate selection by ClpXP.

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

ACKNOWLEDGMENTS

I want to give a huge thanks to my advisor Tania Baker for her patience, guidance, and enthusiasm. Her contribution to my scientific growth is unparalleled. I am very, very fortunate that I got the opportunity to be a graduate student in her lab. I would also like to thank my co-advisor Bob Sauer for great advice and technical support. I would like to extend a big thank you to members of my thesis committee: Tom Rajbhandary, Graham Walker, and Michael Eck for their advice.

All the past and present Baker lab members are to be greatly thanked! Everyone has always been incredibly helpful and ready to answer all my questions, no matter how busy they are. Their scientific expertise and sense of humor are a fantastic combination and I am very lucky that I got to enjoy both! A special thanks to Saskia Neher, Mihaela Pruteanu, Kevin Wang, and Igor Levchenko for a lot of technical help when I first started at the Baker lab and, of course, for their support since then!

I was fortunate enough to work in a room with Peter Chien and Jennifer Hou for most of my graduate career. I have learned an incredible amount from them and I cannot imagine what graduate school would have been like without the two of them. A huge thanks to them for all the science and laughs and utter randomness and, of course, for putting up with me!

I would like to thank my friends (especially Lea Medeiros) from college and graduate school for all their support and friendship over the years.

I am eternally grateful to my whole, very global family for all their support and love. Despite being in many different time zones, they have always been supportive and available at pretty much any time. A special thanks to my family in the US for providing such a wonderful home away from home!

Lastly, a very special thanks to my parents and my sister Ameena for all their love, encouragement, kindness, support, and, most of all, patience! Their tremendous affection across thousands of miles over so many years has played an invaluable part in my personal and professional development and I will forever be grateful for being a part of the Chowdhury family.

TABLE OF CONTENTS

TITLE PAGE 1

ABSTRACT 3

ACKNOWLEDGMENTS 5

TABLE OF CONTENTS 7

LIST OF FIGURES 10

LIST OF TABLES 12

CHAPTER ONE: Introduction 13

PART I: BACKGROUND AND SIGNIFICANCE OF PROTEOLYSIS 14

PART II: SUBSTRATE SELECTION BY THE PROTEOLYTIC MACHINERY 15

A. DEGRADATION TAGS (DEGRONS) 15

1. Prokaryotes 15

i. Protein Quality Control: ssrA tag 15

ii. Intrinsic Degradation Signals 17

iii. Latent Degradation Signals 17

2. Eukaryotes 18

i. Ubiquitinated Proteins 18

ii. Non-ubiquitinated Proteins 19

B. PROTEOLYTIC MACHINERY 20

1. Bacterial ATP-dependent Proteases 20

2. Eukaryotic 26S Proteasome 23

C. ADAPTOR PROTEINS 24

1. Prokaryotes 24

i. SspB 25

ii. RssB (SprE) 26

iii. UmuD 27

iv. ClpS 27

2. Eukaryotes 28

PART III: The CLPXP PROTEASE AND THE ADAPTOR SSPB 29

A. CLPXP PROTEASE 29

1. Structural Features of ClpX: Focus on Pore 30

2. ClpX Pore: Substrate-Binding 32

3. ClpX N domain: Substrate- and Adaptor-Binding 35

B. SSPB ADAPTOR PROTEIN 37

1. Structure and Function 37

2. Substrate Selection: ssrA-tag versus NRseA 39

3. Impact of Substrate-Binding Flexibility 40

PART IV: OUTLINE OF THESIS 41

CHAPTER TWO: Versatile modes of peptide recognition by the ClpX N domain mediate alternative adaptor-binding specificities in different bacterial species 42

ABSTRACT 43

INTRODUCTION 43

RESULTS 45

Phylogenetic comparisons suggest use of different adaptor-tethering contacts 46

Adaptor delivery of cognate substrates to ^{Cc}ClpXP and ^{Ec}ClpXP 48

Residues involved in ^{Cc}SspBα tethering to ClpXP 50

^{Cc}SspBα XB mutations decrease N-domain affinity 55

Effects of ^{Ec}SspB XB mutations on substrate delivery and N-domain affinity 57

Different XB sequences compete for binding the N domain 59

DISCUSSION 61

MATERIALS AND METHODS 64

ACKNOWLEDGMENTS 67

CHAPTER THREE: Substrate Selection by the ClpXP Protease: Future Directions 68

THE SSPB ADAPTOR PROTEIN 69

THE N-TERMINAL DOMAIN OF THE AAA+ UNFOLDASE CLPX 71

APPENDIX I: Proteomic Screen to Elucidate the Effect of SspB on ClpX Substrate Profile in *E. coli* 73

INTRODUCTION 74

RESULTS AND DISCUSSION 78

Trapping experiment 78

Validation of results 82

Caveats 84

MATERIALS AND METHODS 85

APPENDIX II: *In Vitro* Degradation of the ribosomal protein S7 by ClpXP 88

INTRODUCTION 89

RESULTS AND DISCUSSION 91

MATERIALS AND METHODS 94

REFERENCES 96

LIST OF FIGURES

Figure 1.1. The different steps in ssrA-tagging in prokaryotes 16

Figure 1.2. Ubiquitin-tagging of proteins in eukaryotes 19

Figure 1.3. The 26S Proteasome 22

Figure 1.4. Functions of AAA+ Proteins 24

Figure 1.5. Substrate degradation by ClpXP facilitated by SspB 26

Figure 1.6. ClpX Recognition Motifs 30

Figure 1.7. Structure of the hexameric *E. coli* ClpX 31

Figure 1.8. Model of substrate-binding by ClpX pore in the presence and absence of SspB 34

Figure 1.9. *E. coli* ClpX N domain bound to SspB ClpX-binding (XB) peptide 36

Figure 1.10. Structure of SspB with RseA and ssrA peptides 39

Figure 1.11. RseA and ssrA peptides bind overlapping substrate-binding site (SBD) of *E. coli* SspB 40

Figure 2.1. XB conservation and cross-species interaction of SspB and ClpX from α- and γ/β-proteobacteria 47

Figure 2.S1. Comparison of ClpX N-terminal domains of *E. coli* and *C. crescentus* 48

Figure 2.2. *C. crescentus* and *E. coli* ClpXP interact with both ^{Cc}SspBα and ^{Ec}SspB 49

Figure 2.3. ^{Cc}SspBα interacts with ClpXP via its C-terminal region 51

Figure 2.4. Substrate-delivery activity of ^{Cc}SspBα variants with substitution mutations 52

Figure 2.5. Degradation of GFP-^{Cc}DAS by ^{Cc}ClpXP 54

Figure 2.6. Secondary structure in the ^{Cc}SspBα C-terminal region is not critical 55

Figure 2.7. Binding interaction between ^{Cc}ClpX N domain and ^{Cc}SspBα 56

Figure 2.8. Functional and binding interactions between ^{Ec}ClpX N domain and ^{Ec}SspB 58

Figure 2.9. Competition between ^{Cc}SspB XB and ^{Ec}SspB XB motifs for binding ^{Cc}ClpX and ^{Ec}ClpX 60

Figure A1.1. Substrates associated with ClpXP^{trap} in *sspB*$^{+}$ and *sspB*$^{-}$ strains 76

Figure A1.2. Schematic depiction of the different steps in analysis by SILAC 78

Figure A1.3. *In vitro* degradation of YgfZ by ClpXP 84

Figure A2.1. *In vitro* degradation of the N-terminally His-tagged ribosomal protein S7 93

LIST OF TABLES

Table 2.1. K_D values of ^{Cc}SspBα C-terminal peptides binding to *C. crescentus* ClpX N domain (dimer) 57

Table A1.1. "Trapped" ClpX substrates with and without SspB 79

CHAPTER ONE

Introduction

Part I. Background and Significance of Proteolysis.

Proteolysis is important for protein quality control and several regulatory mechanisms in the cells of both prokaryotes and eukaryotes. Given that proteins are involved in virtually all biological pathways in the cell, it is extremely important that they are maintained in their proper structures and at optimal levels. Regulation of protein levels and, consequently, functions can occur at many stages of protein synthesis, such as during transcription, translation, or even by modifications made post-translationally. However, when a fast, switch-like response is required, regulated proteolysis is a particularly efficient way to quickly alter protein levels.

Cells can be exposed to a multitude of environmental stresses from their surroundings, including changes in temperature, pH, or salt concentration. These stress factors can have detrimental effects on protein folding thereby affecting protein activity. In addition to the loss of protein function, unfolding also results in the exposure of normally hidden hydrophobic residues, leading to undesired protein-protein interactions and, in extreme cases, aggregation. A characteristic type of aggregation of usually soluble proteins can be observed in several neurodegenerative diseases such as Parkinson's, Huntington's, and Alzheimer's (Forman et al., 2004).

Formation of protein aggregates can be prevented by removal of unfolded or misfolded proteins. One method of dealing with these aberrant proteins is by using molecular chaperones, which bind exposed hydrophobic patches and help proteins refold into correct conformations. The other, more drastic way of removing misfolded proteins is by proteolysis. The defective proteins are targeted to the proteolytic machinery in the cell where they are broken down into smaller peptides and the amino acids recycled.

In addition to protein quality control, proteolysis is also used as a regulatory mechanism for various cellular processes, such as the cell cycle, gene transcription, and the immune response (Tanaka and Chiba, 1998). In eukaryotes, entry and progress through mitosis is driven by a number of kinases and phosphatases. The balance between these two groups of enzymes regulates the different stages during the cycle. Although phosphorylation is reversible, the irreversible degradation of a kinase or phosphatase provides directionality to the cell cycle. The timing of degradation of the various cell cycle components, such as cyclins and Aurora kinases, is very important to ensure that the stages of mitosis are able to proceed in an orderly manner (Pines, 2006). Proteolysis is also a key process regulating the tumor suppressor p53, which is

usually degraded and kept at low levels in the cell. However, under stress conditions, p53 is stabilized and activated to turn on expression of genes involved in cell cycle arrest, apoptosis, and DNA repair (Marine and Lozano, 2010).

The aim of this work is to understand how the proteolytic machinery chooses specific proteins for destruction. The next part of this chapter (part II) will give an overview of the cell's proteolytic system and its intricate modes of substrate selection. The topic of substrate recognition has been divided into three main sections: degradation signals, features of the proteolytic machinery, and adaptor proteins. Both eukaryotic and prokaryotic systems are discussed with an emphasis on bacterial proteases.

Part II. Substrate Selection by the Proteolytic Machinery

A. Degradation tags (Degrons).

A principle method of targeting proteins for destruction is the use of degradation tags or degrons on substrates. These tags enable proteases to recognize proteins marked for degradation. Degrons may be short sequences encoded in the primary amino acid sequence of proteins (e.g. the ssrA-tag in prokaryotes, see below). This section will focus on examples of some well-characterized degrons in prokaryotes and eukaryotes.

1. *Prokaryotes.*

i. Protein Quality Control and the ssrA tag degron

The ssrA-tagging system is an elegant method used to remove aberrant polypeptides in prokaryotes. The main player in the ssrA-tagging system is ssrA (small stable RNA A), a stable approximately 360-nucleotide RNA molecule that exhibits both mRNA- and tRNA-like properties. The highly conserved ssrA gene is responsible for the elimination of defective polypeptides in all eubacteria (Karzai et al., 2000). These polypeptides can arise from translation of cleaved or prematurely terminated mRNA molecules. In this situation, the cell is faced with stalled ribosomes and aberrant polypeptides. In *E. coli*, ssrA, charged with Ala by alanyl-tRNA synthetase, binds the stalled ribosomes and behaves like a tRNA in adding Ala to nascent polypeptides. The ribosome then switches to translating the mRNA-like sequence of

ssrA to add 10 more residues to the polypeptides, resulting in tagging with the peptide sequence "AANDENYALAA" at the C-termini (Fig. 1.1). This 10-residue-coding-sequence is followed by a normal termination codon, thereby freeing the ribosomes for more rounds of protein synthesis and releasing the polypeptides with ssrA tags. Sequence determinants within the ssrA tag are recognized by proteases, which bind and subsequently degrade the tagged polypeptides (Gottesman et al., 1998; Keiler et al., 1996).

Despite variation in the actual sequence, certain characteristics are shared by most identified ssrA tags in the different organisms. The tag usually consists of polar residues followed by a hydrophobic pentapeptide. In addition, all of them contain the C-terminal Ala. Multiple proteases recognize the ssrA tag albeit the sequences recognized by specific proteases are not always identical. It is thought that the sequence of the tag is optimized in organisms depending on which protease is primarily responsible for ssrA-tagged substrate degradation (Gur and Sauer, 2008; Karzai et al., 2000).

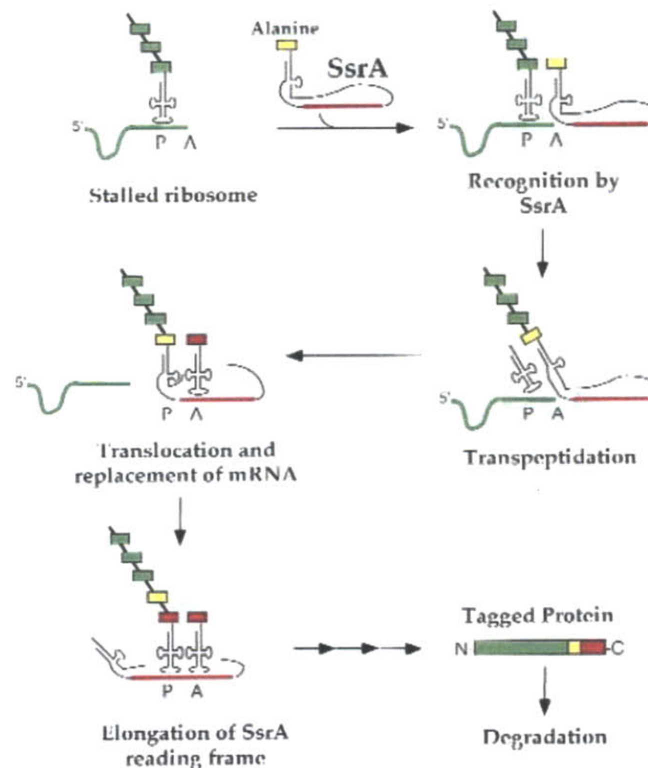


Figure 1.1. The different steps in ssrA-tagging in prokaryotes. If a ribosome stalls on a defective mRNA, the ssrA molecule is recruited to the translation site, leading to the addition of a C-terminal degradation tag on the polypeptide (Figure taken from (Karzai et al., 2000)).

ii. Intrinsic Degradation Signals

As mentioned earlier, many functional proteins are degraded by proteases as a mode of controlling various biological pathways. Many of the proteins that are regulated by proteolysis contain intrinsic degradation signals in their primary sequence. Usually these degrons are located near the N- or C-termini of proteins, which may be because those regions are easier for proteases to access. The DNA-binding protein Dps in *E. coli* has been shown to be regulated by proteolysis (Stephani et al., 2003). Dps binds non-specifically to DNA to form stable nucleoprotein complexes known as biocrystals and confers protection in conditions such as oxidative stress, UV and gamma irradiation, thermal stress, metal toxicity, etc. (Almiron et al., 1992; Nair and Finkel, 2004). During stationary phase or under nutrient limiting conditions, Dps is expressed at high levels whereas in exponential phase, it is recognized through its N-terminal residues by proteases and rapidly degraded. There are various other substrates, which are also recognized through N-terminal regions, such as the λ-bacteriophage DNA replication initiator protein λO and the stationary-phase RNA polymerase sigma factor σ^S (Flynn et al., 2003; Gonciarz-Swiatek et al., 1999; Weichart et al., 2003). Interestingly, the stability of a protein can also be attributed to a single residue at the N-termini of proteins. According to the N-end rule, which is present in both prokaryotes and eukaryotes, specific residues can act as degrons when located at the very N-termini of proteins. In *E. coli*, an N-terminal Leu, Phe, Trp, or Tyr residue targets proteins to intracellular proteases for destruction (Varshavsky, 1996).

Many substrates can also have degrons at their C-termini, which are recognized by proteases. An example is the ssrA-tag described above, which is appended on the C-termini of polypeptides. SulA protein, an inhibitor of cell division, also has a C-terminal degron. SulA levels are kept low by proteases, one of which is thought to recognize the C-terminal His residue of SulA (Gottesman, 2003).

iii. Latent Degradation Signals

There are several interesting proteins that are targeted for proteolysis via latent or hidden degrons. To be degraded, such proteins have to be first processed, which exposes a previously hidden degron resulting in subsequent degradation. The transcriptional repressor LexA is a well-

characterized example. Under normal conditions, LexA has an inhibitory effect on expression of genes involved in DNA-damage response. However, upon DNA damage, exposed single-stranded DNA activates a protein called RecA, which then induces self-cleavage of LexA to separate its N-terminal DNA-binding and C-terminal dimerization domains. The C-terminal region of the newly formed N-terminal domain is VAA-COO^-, which is very similar to the C-terminal region of the ssrA tag (LAA-COO^-). Proteases are able to recognize this degron and degrade the N-terminal domain of LexA. Consequently, the inhibitory effect of LexA is removed, enabling induction of genes involved in the DNA-damage response (Neher et al., 2003; Pruteanu and Baker, 2009).

Latent degrons may also become exposed upon changes in conformation or oligomeric states of substrates, as observed in the case of the ribosomal protein L10. Although studies have shown that L10 is degraded both *in vivo* and *in vitro*, it is stable when bound to the L7/L12 subunit and assembled into the 50S ribosome. Interestingly, like the ssrA degron, L10 also contains two Ala residues at its C-terminal region, which is thought to be "hidden" when in complex with L7/L12. Upon complex disassembly, the C-terminal degron of L10 becomes exposed and is thus recognized by the proteolytic machinery resulting in its rapid degradation (Flynn, 2004; Petersen, 1990).

2. *Eukaryotes*

i. Ubiquitinated Proteins

In eukaryotes, substrates targeted for degradation are usually tagged with multiple copies of a small protein called ubiquitin. The post-translational covalent attachment of ubiquitin to substrates is facilitated by three enzymes: the first activating enzyme E1 uses ATP-hydrolysis to form a thiolester bond with the C-terminal glycine of ubiquitin, which is then transferred to a second ubiquitin-carrier enzyme E2. In the next step, a ubiquitin-ligase E3 catalyzes the transfer of ubiquitin to the substrate resulting in formation of an amide bond between the C-terminus of ubiquitin and the amino-group of lysine residues in the substrate. This event is repeated to form a poly-ubiquitin chain on the substrate with bonds forming between C-terminus of one ubiquitin subunit and K^{48} of the next subunit. The poly-ubiquitinated substrate is recognized by the proteolytic machinery (discussed below in Part III) and degraded. The ubiquitin subunits are

removed prior to substrate degradation and can therefore be recycled (Fig. 1.2) (Hershko and Ciechanover, 1998).

The substrate specificity of E3 ligases is key to the selection of proteins for degradation. There are various families of ligases in eukaryotes. These enzymes recognize specific substrates through recognition signals and transfers ubiquitin subunits onto them, from either E2 ligases or from E3 ligases themselves, to form the poly-ubiquitin chains. Many substrates that are ubiquitinated contain PEST elements, which are regions rich in Pro, Glu, Ser, and Thr residues. In many cases, residues within PEST elements are phosphorylated and thus targeted by E3 ligases. Examples of phosphorylated substrates include yeast G1 cyclins Cln3 and Cln2. Substrates can also be recognized by E3 ligases through their N-termini. N-end rule substrates in eukaryotes contain basic or bulky hydrophobic residues at their N-termini, which are recognized by the ligases resulting in subsequent ubiquitin-transfer and proteolysis (Hershko and Ciechanover, 1998).

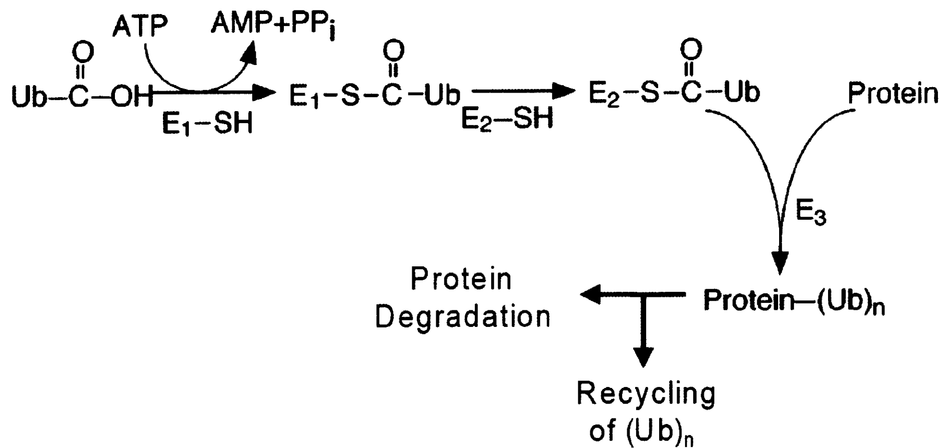


Figure 1.2. Ubiquitin-tagging of proteins in eukaryotes. Ubiquitin molecules, activated by ATP hydrolysis, are eventually transferred onto proteins that are consequently degraded (Figure (modified) taken from (Hershko and Ciechanover, 1998)).

ii. <u>Non-ubiquitinated Proteins</u>

Proteins without attached ubiquitin chains can also be degraded, adding another layer of complexity to substrate selection. L-ornithine decarboxylase (ODC) was the first identified non-ubiquitinated proteolytic substrate. This enzyme catalyzes the first step in synthesis of

polyamines, high levels of which can cause tumorigenesis. Degradation of ODC is one of the main methods of regulating polyamine synthesis. However, ODC is not ubiquitinated but is targeted for proteolysis through its unstructured C-terminal region. The last 19 C-terminal residues appear to be important in length but not sequence whereas two upstream residues C^{441} and A^{442} (numbering corresponds to the 461-residue mouse ODC) have been shown to be important for ODC recognition for proteolysis (Jariel-Encontre et al., 2008). More details about ODC recognition will be discussed in a later section on adaptors.

Since the identification of ODC, several other proteins (e.g. c-Fos, p53, and Rb), have been shown to be degraded without ubiquitination. However, because these proteins can also be degraded in an ubiquitin-dependent manner, it is currently unclear if their ubiquitin-independent degradation is physiologically relevant (Jariel-Encontre et al., 2008; Schrader et al., 2009).

There is strong evidence that "aged" or oxidized proteins can be degraded by the eukaryotic proteolytic machinery without being ubiquitin-tagged. As proteins "age", they undergo spontaneous modifications, such as deamidation of asparaginyl residues and isomerization of aspartyl residues, leading to inactivation. An example of an "aged" protein is calmodulin (CaM), which functions as a calcium sensor in eukaryotes. CaM has a very long half-life of up to 25 hours in mammalian cells and, during its lifetime, gets modified in its calcium-binding region. Calcium-binding is thought to make the CaM structure rigid and stable, thereby preventing its degradation. Upon modifications due to aging, CaM cannot bind calcium to form this stable conformation. It is postulated that the destabilized CaM is recognized by the proteolytic machinery in a ubiquitin-independent way and subsequently degraded. Oxidized proteins have also been shown to be degraded without being ubiquitinated although details of the mechanism are currently unknown (Jariel-Encontre et al., 2008).

B. Proteolytic Machinery

To understand the details of how proteases select substrates, it is important to look at both sides of the substrate-protease interaction. As discussed in the previous section, substrates are recognized by the proteolytic machinery in the cell via degrons. The next step is to explore how the proteases interact with these degrons of substrates. This section will discuss interesting characteristics of a few proteolytic machineries in prokaryotes and eukaryotes, which allow them to recognize degrons and degrade substrates.

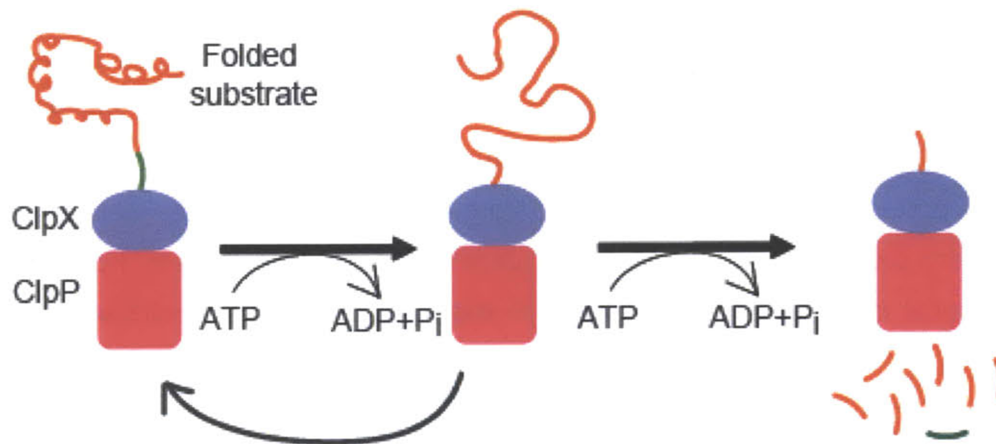
1. *Bacterial ATP-dependent proteases*

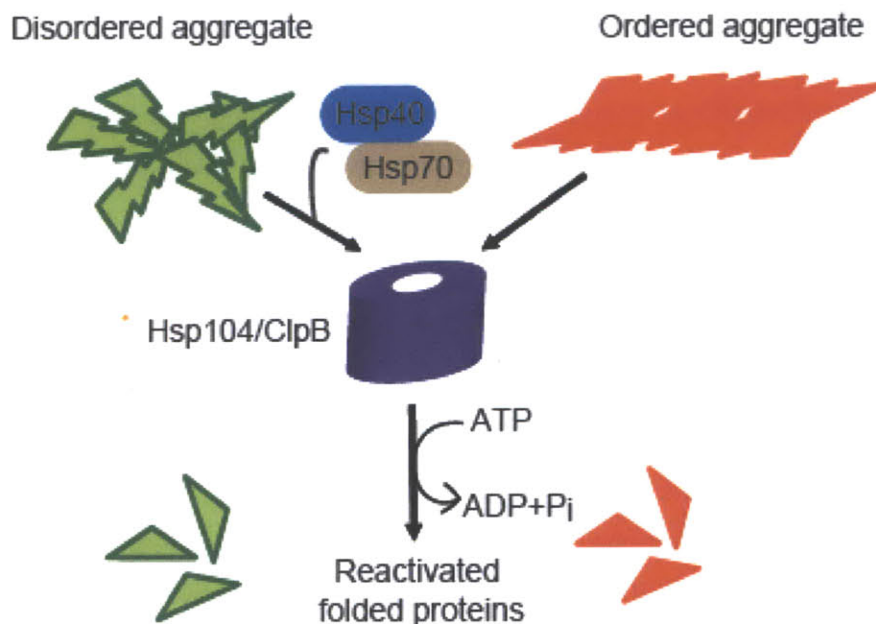
In bacteria, there are multiple classes of energy-dependent proteases that are involved in proteolysis. These proteases, including ClpXP, ClpAP, HslUV, Lon, and FtsH, have orthologs in mitochondria and chloroplasts of eukaryotes. Their compartmentalized set-up is architecturally very similar to that of the 26S proteasome (discussed below). There are two distinct subunits or domains in these proteases: the AAA+ (ATPases associated with various cellular activities) subunit and the proteolytic subunit. In *E. coli*, the ClpXP, ClpAP, and HslUV proteases are composed of an ATPase subunit (ClpX, ClpA, HslU) and a peptidase subunit (ClpP, HslV). The peptidase ClpP can associate with the ATPases ClpX or ClpA to form the active proteases ClpXP and ClpAP respectively. Similarly, the HslUV protease consists of the ATPase HslU (ClpY) and the peptidase HslV (ClpQ). In contrast, the proteases FtsH and Lon have both ATPase and proteolytic components as separate domains on a single subunit (Gottesman, 2003). Different proteases may be essential depending on the organism. For instance, only FtsH is essential in *E. coli* whereas ClpXP is essential in the α-proteobacteria *Caulobacter crescentus*.

Members of the AAA+ superfamily are present in all kingdoms in varying numbers with most eukaryotes containing about 50-80 members (Hanson and Whiteheart, 2005). They are characterized by an ATPase domain containing special structural motifs including the Walker-A and –B motifs, which are sequence elements important for nucleotide-binding and hydrolysis by ATPases. AAA+ proteins are usually oligomeric, most forming hexamers with a central pore. They undergo conformational changes upon nucleotide-binding and –hydrolysis to unfold and thread substrates through the pore. They perform a wide and diverse range of functions in the cell by associating with various other proteins (Fig. 1.3). These functions include protein unfolding for proteolysis and disassembly of protein aggregates and complexes, an example being the AAA+ protein N-ethylmaleimide-sensitive factor (NSF)-mediated disassembly of SNARE complexes formed during membrane fusion (Hanson and Whiteheart, 2005).

The ATPase subunit or domain is responsible for selecting, unfolding, and translocating substrates into the barrel-like proteolytic subunit where the peptidase sites are sequestered. Once in contact with the proteolytic subunit, the substrates are degraded to form ~10-to-15-residue peptides (Gottesman, 2003). Although ClpP is the proteolytic component of both ClpXP and ClpAP proteases, the two proteases have distinct substrate preferences attributed to the

associated ATPase. However, ClpX and ClpA also share at least some common substrates, a characteristic shared by many ATPases. For instance, both ClpXP and ClpAP are capable of degrading ssrA-tagged substrates although each recognizes distinct parts of the tag. Functional redundancy because of this overlap in substrate choice can explain why specific proteases are not essential in all organisms.





c. Protein-complex disassembly

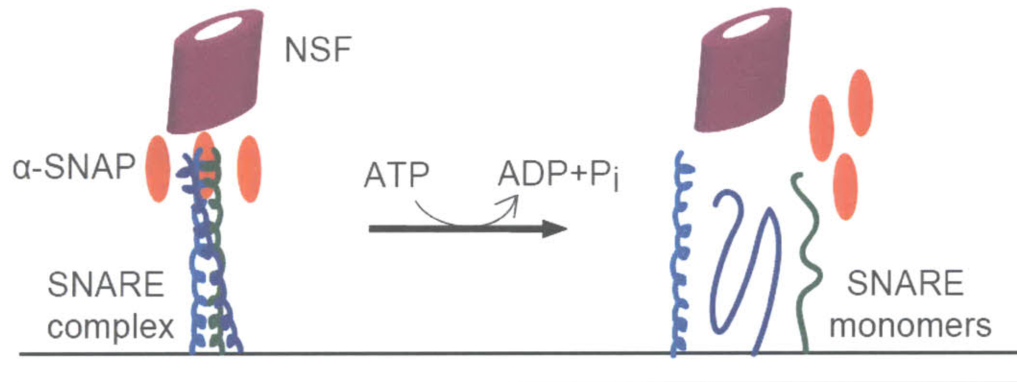


Figure 1.3. Functions of AAA+ Proteins. (a) is showing the steps in ClpXP-mediated degradation of a folded substrate containing a degron (shown in green) in bacteria. ATP hydrolysis is required for the unfolding and translocation of the substrates. (b) is showing the disaggregation of proteins (formed due to stress) by the Hsp104 (ClpB) ATPase, which is thought to cooperate with the Hsp70 system for the disaggregation of disordered aggregates. (c) is showing the role of the ATPase NSF, together with the accessory α-SNAP protein, in disassembling SNARE complexes that form during membrane fusion (Figure (modified) from (Hanson and Whiteheart, 2005)).

2. *Eukaryotic 26S Proteasome*

The 26S proteasome in eukaryotes is responsible for ATP-dependent degradation of most cellular proteins. It consists of two sub-complexes: the 20S threonine-peptidase component and the 19S regulatory component, which stack to form a barrel-like structure. The core complex is made up of 4 stacked rings, each composed of 7 protein subunits. The two outer rings are known as the α-rings and the two inner ones, the β-rings. The active sites on the β-rings, which are sequestered in the center of the stack, have trypsin-, chymotrypsin-, and caspase-like proteolytic activities (Fig. 1.4) (Navon and Ciechanover, 2009).

The 20S particle can associate with one or two 19S regulatory complexes. The 19S component is composed of at least 19 subunits, 9 of which form the lid and the remaining 10, the base. There are 6 ATPases in the base, which associates with the 20S particle to form the active 26S proteasome. Substrates are selected via their associated poly-ubiquitin chains by subunits of the 19S particle, such as the ATPase subunit Rpt5 and the non-ATPase subunits Rpn10 and Rpn13 (Fig. 1.4). After selection, these substrates are unfolded and the ubiquitin chains are

removed. The unfolded polypeptides are then transmitted into the stacked rings of the 20S particle where proteolytic cleavage occurs (Navon and Ciechanover, 2009; Wolf and Hilt, 2004).

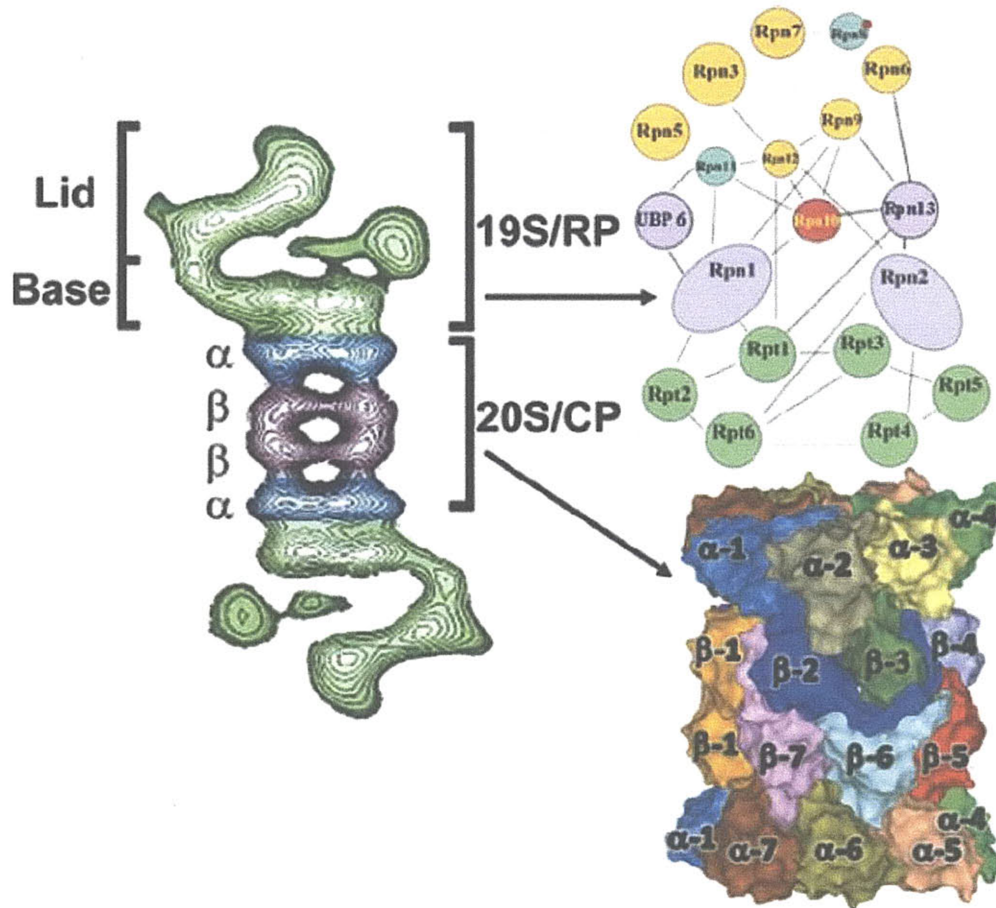


Figure 1.4. The 26S Proteasome. The 26S proteasome consists of the 20S core particle (CP) and the 19S regulatory particle (RP). The 20S CP is formed by four 7-subunit rings: two α- and two β-rings with the active sites located within the internal chamber on the β-subunits. The 19S RP, which can be further divided into the lid and base, is made up of several subunits including 6 ATPases shown as green circles (Figure taken from (Navon and Ciechanover, 2009)).

C. Adaptor Proteins

In addition to directly recognizing degrons on substrates, intracellular proteases can also use accessory proteins called adaptors to modulate substrate choice. Adaptors have different

modes of function with some acting as delivery proteins that bring substrates to proteases and others directly affecting protease activity, e.g. by inducing conformational changes. These accessory proteins have the potential to expand the repertoire of substrates for proteases as well as provide a level of proteolytic regulation.

1. *Prokaryotes*

Prokaryotic adaptors are usually specific to a given protease and can affect not only degradation of specific substrates but also the global composition of the cell's proteome. Multiple adaptors in *Bacillus subtilis* and *E. coli* have been well-studied in the past few years. One of the first prokaryotic adaptor proteins to be identified was the MecA adaptor in *B.subtilis*, which interacts with the AAA+ protease ClpCP. Interestingly, all activities of ClpCP appear to be adaptor-dependent (Kirstein et al., 2009). Because the work in this thesis focuses on the proteolytic system in *E. coli*, a general overview of the four identified *E. coli* adaptors SspB, RssB, UmuD, and ClpS is given below.

i. SspB

The best-characterized *E. coli* adaptor protein is SspB, a brief overview of which is given here. It is one of the main players in this thesis and will therefore be discussed in greater detail in the next section (see Part III). SspB functions as a canonical adaptor by binding substrates and delivering them to the ClpXP protease. It tethers specific substrates to ClpX and enhances their degradation rates (Fig. 1.5). Substrates containing the ssrA-tag are SspB-modulated, the adaptor and protease recognizing distinct portions of the tag (Dougan et al., 2003; Levchenko et al., 2000; Wah et al., 2003). For instance, in the *E. coli* ssrA tag (AANDENYALAA[11]-COO$^-$), SspB recognizes the first four and the seventh residues (highlighted in cyan) whereas ClpX recognizes the last three residues and the carboxyl group (underlined). Interestingly, ClpA recognizes the residues 1, 2, and 8-10 of the ssrA-tag (highlighted in yellow): AANDEN**Y**ALAA[11] (Flynn et al., 2001). Because SspB and ClpA recognize overlapping residues in the ssrA-tag, SspB inhibits degradation of ssrA-tagged substrates by the ClpAP protease. Although physiological relevance remains unclear, it is possible that this inhibition allows ClpAP to focus on degrading other substrates of higher priority, leaving ClpXP to degrade ssrA-tagged polypeptides, which are consistently generated in the cell.

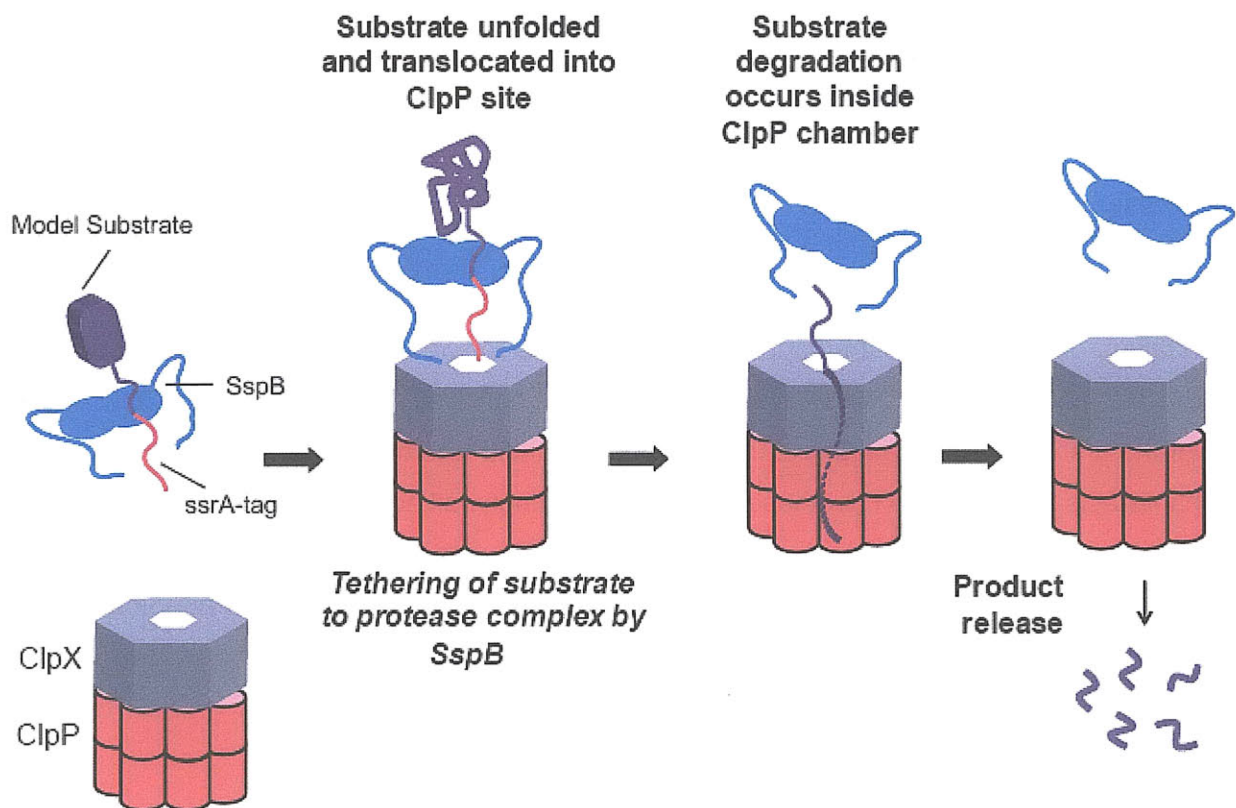


Figure 1.5. Substrate degradation by ClpXP facilitated by SspB. SspB interacts with ssrA-tagged substrates and tethers them to ClpXP. ClpX unfolds and translocates substrates into the active site chamber of ClpP, leading to proteolysis. The tethering of substrate to ClpX causes an increase in local concentration of the substrate, thereby resulting in an enhanced rate of degradation.

ii. RssB (SprE)

The ClpXP-specific adaptor RssB or SprE is involved in degradation of the stationary phase sigma factor σ^S in *E. coli*. RpoS or σ^S is the master regulator of general stress response in bacteria and is expressed under conditions such as high temperature, nutrient limitation, or osmotic stress. It is tightly regulated at the levels of transcription, translation, stability, and activity (Kirstein et al., 2009). ClpXP-mediated σ^S degradation is dependent on RssB (SprE), which is a response regulator phosphorylated by acetyl phosphate and/or the histidine sensor kinase ArcB (Bouche et al., 1998; Mika and Hengge, 2005; Muffler et al., 1996; Pratt and Silhavy, 1996; Zhou and Gottesman, 1998; Zhou et al., 2001). Phosphorylated RssB binds σ^S and enables it to be degraded by ClpXP by exposing an otherwise hidden degron on the N-

terminal region of σ^S (Klauck et al., 2001; Studemann et al., 2003). Unlike ssrA-tagged proteins, which can be degraded in the absence of SspB, degradation of σ^S does not occur in the absence of RssB. This dependence on the adaptor allows more stringent regulation on σ^S proteolysis, which is not surprising given the wide effect of the sigma factor on the cell's transcriptome.

iii. UmuD

The UmuD adaptor is also specific to ClpXP and plays a key role in the SOS response to DNA-damage. The SOS response induces expression of genes involved in DNA repair and replication. Because these proteins can themselves cause damage if expressed inappropriately or allowed to accumulate, the cell uses proteolysis to maintain them at optimal levels. The Umu proteins fall into this category as they are involved in error-prone DNA synthesis, allowing DNA replication despite lesions in the template. Upon DNA-damage, the UmuD protein is processed to form a shorter version UmuD', which homodimerizes and interacts with UmuC to form the translesion DNA polymerase V (Pruteanu and Baker, 2009). ClpXP-modulated degradation of UmuD' regulates the levels of this potentially-harmful enzyme. The unprocessed UmuD can heterodimerize with UmuD' and act as a delivery protein, binding ClpX and enabling the protease to degrade UmuD' thereby lowering levels of the error-prone polymerase (Frank et al., 1996; Gonzalez et al., 2000; Neher et al., 2003).

iv. ClpS

In contrast to RssB, UmuD, and SspB, ClpS is a ClpAP-specific adaptor protein. Interestingly, ClpS has both positive and negative effects on degradation, depending on the substrate. It inhibits degradation of ssrA-tagged substrates by the ClpAP protease, at the same time facilitating ClpAP-mediated degradation of a group of substrates known as the N-end rule substrates (Dougan et al., 2002; Erbse et al., 2006; Hou et al., 2008; Schmidt et al., 2009; Wang et al., 2007). According to the N-end rule, the N-terminal residue of a protein dictates its stability. Specific N-terminal residues, such as residues L, F, Y, or W in *E. coli,* when present at the very N-terminus act as degradation signals, resulting in short half-lives of these proteins (Mogk et al., 2007). By having differential effects on ClpAP-substrates, ClpS heavily influences substrate selectivity by the protease. Interestingly, ClpS interacts with the N domain of ClpA,

which has also been implicated in substrate specificity of the ClpAP protease in a manner similar to the ClpX N domain described later (Dougan et al., 2002; Erbse et al., 2008; Lo et al., 2001).

2. *Eukaryotes*

There are also adaptor proteins in eukaryotes, which target substrates to the proteasome for degradation. Three adaptors ("ubiquitin receptors") Rad23, Dsk2, and Ddi1 in *Sacchromyces cerevisiae* have been shown to bind ubiquitinated proteins and the proteasome simultaneously to deliver substrates to the proteasome. Moreover, genetic evidence indicates that there are probably additional adaptor proteins that have not yet been identified (Schrader et al., 2009). Interestingly, these adaptor proteins can also interact with ubiquitin ligases suggesting the possibility that adaptors can deliver substrates from the site of ubiquitination to the proteasome. For instance, the proteasome and the ligase Ufd2 compete to interact with Rad23, thereby indicating that Rad23 may shuttle ubiquitinated substrates from the ligase to the proteasome. Similarly, Ddi1 interacts with the ligase Ufo1 and facilitates degradation of the Ufo1-substrate HO endonuclease (Elsasser and Finley, 2005).

Degradation of the non-ubiquitinated substrate ODC is facilitated by the antizyme-1 protein (AZ1). Binding to AZ1 causes exposure of the ODC C-terminal degron and functions as an adaptor in enhancing ODC-proteasome interaction. Interestingly, AZ1 has also been shown to affect degradation of additional non-ubiquitinated substrates such as the cell cycle regulatory proteins cyclin D1 and Aurora-A kinase (Jariel-Encontre et al., 2008).

Although traditionally ubiquitin has been considered a degron itself, recent studies have indicated that the eukaryotic degron is actually composed of two parts, one of them being the attached ubiquitin chains. To be degraded, the ubiquitinated substrate has to also contain an unstructured region, which forms the second part of the degron. Thus, the substrates are delivered to the proteasome via their poly-ubiquitin tags where the unstructured region acts as a degradation initiation site. In this scenario, ubiquitin can be thought of as an adaptor protein, tethering the substrate to the proteasome, which can now engage the unstructured initiation site and subsequently degrade the substrate (Schrader et al., 2009). The unstructured region can be located at either termini or even in the middle of the substrate. In fact, the initiation site may also be flanked by folded domains on either side of substrates. However, there appears to be certain

requirements, e.g. specific length or distance from the ubiquitin chains, for an unstructured region to be an initiation site as not all unstructured regions can behave like degrons. This characteristic enables the proteasome to selectively choose specific subunits of complexes to degrade, leaving the other subunits intact. An example is the cell cycle regulator protein complex Sic1-Cdk-cyclin where only the ubiquitinated Sic1 is degraded despite both Sic1 and cyclin containing unstructured regions (Schrader et al., 2009).

Part III. The ClpXP Protease and the Adaptor SspB

E. coli ClpXP protease has been well studied over the past few years. Although insights have been gained on the protease, questions still remain about how ClpXP selects substrates and adaptors and the work in this thesis will focus on trying to provide new insights into these critical issues. The next section will provide more detailed analyses of the mode of substrate selection by ClpXP and SspB as well as how these two key elements of the bacterial proteolytic machinery communicate with each other.

A. ClpXP Protease

One of the best-studied AAA+ proteases is the ClpXP protease in bacteria. Interestingly, the hexameric ATPase ClpX itself is able to function as a molecular chaperone in the absence of ClpP, preventing protein aggregation (e.g. λO) and disaggregating already-formed protein complexes such as the tetrameric MuA transposase (Abdelhakim et al., 2008; Levchenko et al., 1995; Wawrzynow et al., 1995). However, it needs to associate with the tetradecameric serine peptidase ClpP to form an active protease (Gottesman et al., 1993; Grimaud et al., 1998; Wojtkowiak et al., 1993). The active sites of ClpP are sequestered within a barrel-like structure formed by two stacked 7-mers. Access to the active site is regulated by ClpX, which forms a ring and associates with ClpP through conserved loops. ClpX recognizes substrates and uses rounds of ATP hydrolysis to unfold and translocate them into the ClpP pore for degradation. ClpX is responsible for recognizing specific substrates because, once a polypeptide is in proximity to the ClpP active sites, it is hydrolyzed essentially regardless of sequence (Wojtkowiak et al., 1993).

ClpXP is responsible for the degradation of hundreds of substrates via degrons in *E. coli*. Interestingly, the ClpX degradation tags vary greatly in sequence (Fig. 1.6). A proteomic study by Flynn et al. (2003) identified five general classes of degrons recognized by ClpX: two classes of C-terminal tags (C-motifs 1 and 2) and three classes of N-terminal tags (N-motifs 1, 2, and 3). Although both have C-terminal degrons, the ssrA tag has a C-motif 1 (AA-COO$^-$) whereas MuA has a C-motif 2 (RRKKAI-COO$^-$). The N-terminal region of Dps (NH_3^+-STAKL), discussed earlier, falls into the N-motif 1 class and is distinct from the other classes of N-motifs (classes 2 and 3). By having a range of degrons as opposed to a single consensus sequence, ClpX is able interact with a large spectrum of substrates, providing flexibility to the ClpXP protease in substrate choice.

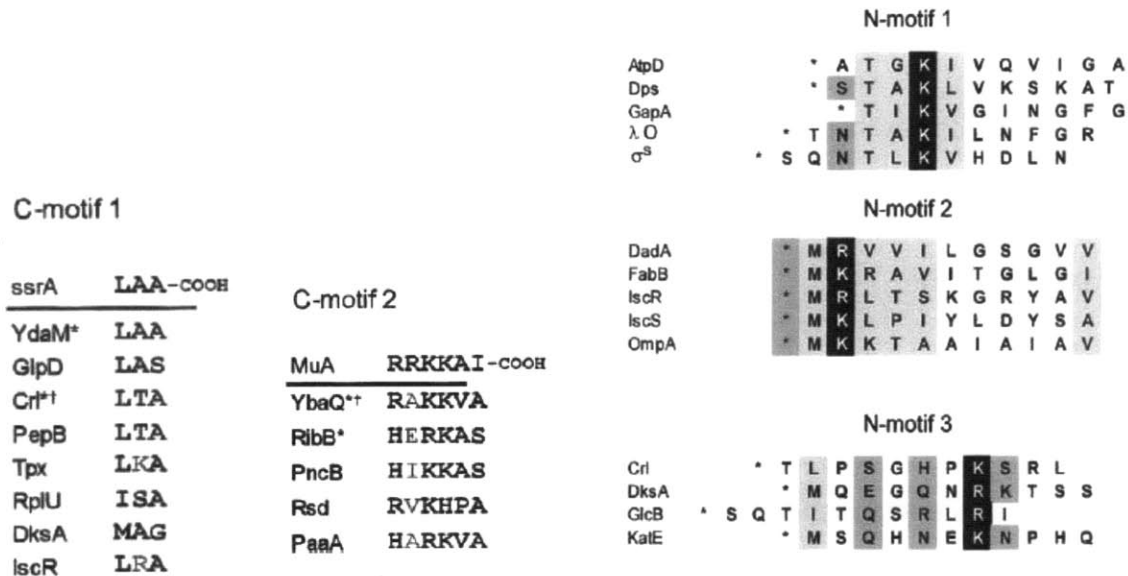


Figure 1.6. ClpX Recognition Motifs. There are two C-terminal degrons, one resembling the ssrA-tag and the other one, the tag of MuA transposase. There are also three distinct N-terminal motifs that can function as degrons (Figure taken from (Flynn et al., 2003)).

1. *Structural Features of ClpX: Focus on Pore*

This section will focus on the structure of the ClpX pore formed by the hexameric ring. ClpX is composed of an N-terminal domain (discussed in detail later) followed by large and small AAA+ domains. Although the monomers are identical in sequence and structure, the orientation of the two AAA+ domains vary in the subunits such that the monomers can be classified into two

groups: type 1 and type 2. Four of the subunits belong to type 1 class and are arranged to efficiently bind nucleotides whereas the other two type 2 subunits are in conformations that prevent nucleotide-binding. Thus, at any given time, only four subunits of the hexamer are competent for nucleotide-binding (Fig. 1.7) (Glynn et al., 2009; Hersch et al., 2005). The interaction of the small AAA+ domain of one subunit with the large AAA+ domain of the adjacent subunit forms the major interface of subunit-interaction and is thought to be structurally static. Nucleotide-binding to only the type 1 subunits changes the orientations of the linkers between the small and large domains within each subunit. This combination of static "inter-subunit" and dynamic "intra-subunit" interactions results in subunit-staggering and an overall asymmetric conformation of the hexamer. Therefore, hydrolysis of ATP in the type 1 subunits may be able to cause conformational changes that are propagated throughout the hexamer and have an effect on bound substrate (Glynn et al., 2009).

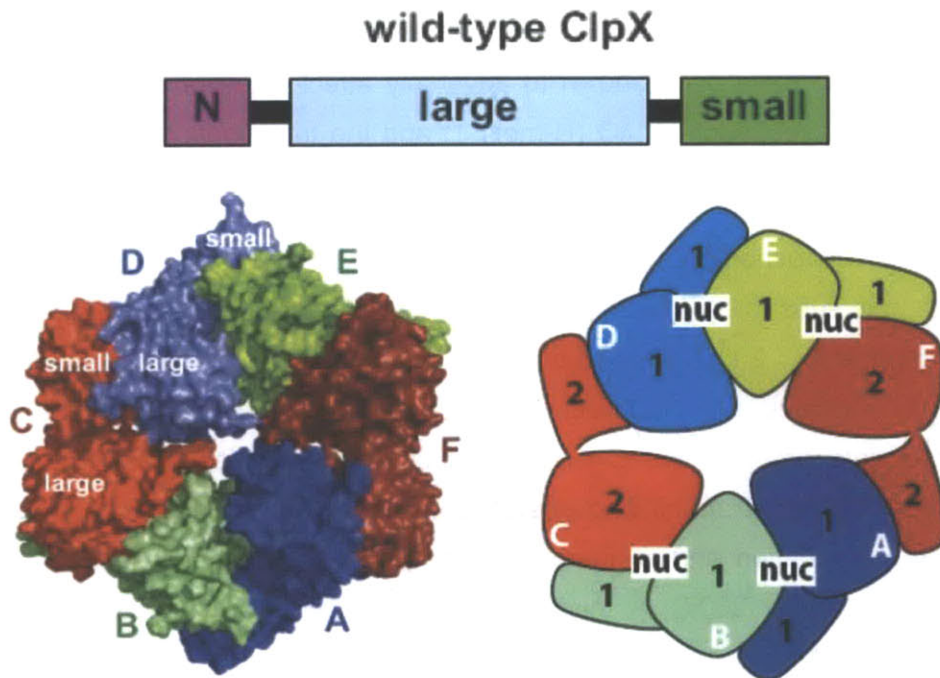


Figure 1.7. Structure of the hexameric *E. coli* ClpX (monomers labeled A-F). ClpX consists of an N domain followed by large and small AAA+ domains. The variant crystallized did not contain the N domain (1-61 residues). The large and small domains are labeled on the hexamer shown on the left. Type 1 and 2 subunits are labeled on the hexamer shown on the right. The type 1 subunits conformations allow nucleotide (nuc)-binding whereas the type 2 subunits are unable to bind nucleotide (Figure taken from (Glynn et al., 2009)).

AAA+ proteins that associate with ClpP (such as ClpX, ClpA, and ClpC) have a conserved [LIV]-G-[F/L] peptide in the C-terminal region of the ATPase domain (Kim et al., 2001). Organisms lacking ClpP, such as *Saccharomyces cerevisiae* and *Mycoplasma pneumoniae*, do not have any Clp ATPase containing this peptide. This tripeptide in *E. coli* ClpX has the sequence IGF and, if substitutions are made at these positions, the variant ClpX proteins are able to function as ATPases but are unable to interact with ClpP (Kim et al., 2001). Studies indicate that the IGF tripeptide lies in on an exposed loop and all 6 of the IGF loops in the hexamer are necessary for ClpP-association and protease activity (Kim et al., 2001; Martin et al., 2007; Singh et al., 2001; Wang et al., 1997).

2. *ClpX Pore: Substrate-Binding*

Studies indicate that the there is not just a single mechanism through which ClpX selects different classes of substrates. A highly conserved GYVG motif (pore 1 loop) in the ClpX pore has been implicated in the recognition and engagement of substrates with C-motif 1. Whereas mutations in this region have been shown to adversely affect selection of C-motif 1 proteins, substrates in other classes, such as λO (N-motif 1) and the transcriptional regulator DksA (N-motif 2), are not affected (Siddiqui et al., 2004).

Interestingly, human mitochondrial ClpX does contain this pore GYVG loop but is unable to degrade ssrA-tagged substrates, indicating the presence of multiple ssrA-tag binding regions in *E. coli* ClpX (Kang et al., 2002). One such binding site is a second positively-charged loop called the RKH loop located at the opening of the ClpX pore, which evidence indicates interacts with the α-carboxylate group at the C-terminus of ssrA-tagged substrates. Mutations in this region have differential substrate-binding effects, causing a dramatic decrease in rate of ssrA-tagged substrate degradation and an increase in degradation of other classes of substrates, such as λO (Farrell et al., 2007). A third loop that appears to be involved in ssrA-tagged substrate selection is a highly conserved region located at the bottom of the central pore of ClpX known as the pore 2 loop (Fig. 1.8). Mutations in this loop interfere with binding and degradation of ssrA-tagged substrates as well as interaction with ClpP (Glynn et al., 2009; Martin et al., 2007). The importance of the RKH and pore 2 loops in degradation of ssrA-tagged substrates was further reiterated by an elegant pore-loop swap experiment by Martin et al. (2008a). Unlike the GYVG loop, the RKH and pore 2 loops of *E. coli* ClpX are not conserved in human mitochondrial ClpX. In this study by Martin et al., the RKH and pore 2 loops from *E. coli* ClpX

were swapped onto the human mitochondrial ClpX. In contrast to wild-type human mitochondrial ClpXP, this variant was able to degrade substrates tagged with the *E. coli* ssrA peptide almost as efficiently as the *E. coli* ClpXP, indicating the importance of the RKH and pore 2 loops in ssrA-tagged substrate recognition.

The pore loops of ClpX are vital for the protease activity of the ClpXP complex. They are involved in substrate recognition, binding, unfolding, translocation as well as interaction with the ClpP peptidase (Farrell et al., 2007; Martin et al., 2007, 2008a, b; Siddiqui et al., 2004). Recent studies have shed more light into how these loops influence ClpX-binding of different substrates. As mentioned earlier, all three loops are implicated in binding ssrA-tagged substrates despite being spaced apart along the ClpX pore. Structural and biochemical studies provide a consistent model for the mode of action of these loops in substrate-binding (Fig. 1.8). The RKH loop, located at the opening of the axial pore, is positioned to act as a specificity filter in engaging the negatively-charged C-terminus of C-motif 1 (e.g. ssrA-tag) substrates via transient electrostatic interactions. Once close to the pore, the substrates are further engaged by the pore 1 and pore 2 loops. The staggered arrangement of the ClpX hexamer causes the pore 1 loops of some subunits to be near the pore 2 loops of others, thereby allowing a substrate to simultaneously interact with both pore 1 and 2 loops (Glynn et al., 2009; Martin et al., 2008a; Siddiqui et al., 2004). This two-step substrate-binding model explains why the RKH and the pore 2 loops are both important for binding to a short C-motif 1 (C-terminal LAA-COO^- in ssrA tag) despite being located at the opposite ends of the ClpX hexamer (Farrell et al., 2007; Martin et al., 2008a).

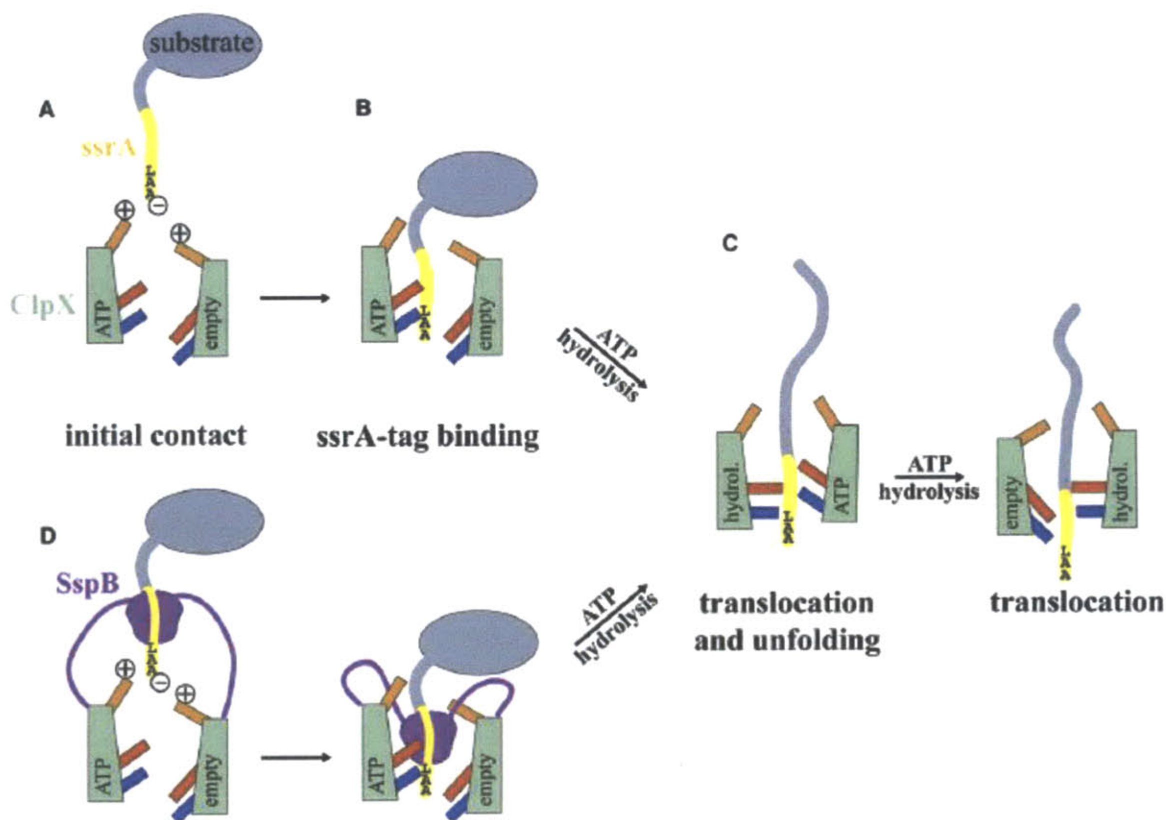


Figure 1.8. Model of substrate-binding by ClpX pore in the presence and absence of SspB. The three loops: RKH, GYVG (pore 1 loop), and pore 2 loops are shown in gold, red, and blue respectively. In (A), the C-motif 1 (e.g. LAA-COO$^-$) is recruited via electrostatic interactions with the RKH loop followed by interactions with pore 1 and pore 2 loops, which pull the substrate further into the pore (shown in (B)). In (C), conformational changes in the staggered subunits, due to ATP hydrolysis, result in unfolding and translocating of substrates. As shown in (D), SspB can also deliver substrates to ClpX and, because of loop staggering, have access to deeper parts of the pore (Figure taken from (Martin et al., 2008a)).

Although all three pore loops of ClpX are required for efficient degradation of specific substrates, only the pore 1 (GYVG) loop is most highly conserved in AAA+ proteases. This conservation is explained by the key role played by the pore 1 loop in substrate unfolding and translocation (Martin et al., 2008b; Siddiqui et al., 2004; Wang et al., 1997). Mutational studies in *E. coli* ClpX indicate that the GYVG loop may be playing an important role in gripping, unfolding, and translocating substrates by the application of force induced by ATP-driven conformational changes in type 1 ClpX subunits (Glynn et al., 2009; Martin et al., 2008b). Nearby type 2 subunits, which are unable to bind ATP, use their GYVG loops to grip the substrates, preventing slipping and assisting substrate translocation. The staggering and

movement of the loops in different subunits assists unfolding of substrates by pulling them further into the pore (Glynn et al., 2009; Martin et al., 2008b). By virtue of being involved in the core function of unfolding and translocation of substrates, it is understandable that the GYVG loop is highly conserved in ClpX (and more widely among AAA+ protein unfoldases). On the other hand, the other loops play roles in substrate-specificity of the proteases, which can vary in different organisms, thereby explaining why they are not as conserved across species.

3. *ClpX N domain: Substrate- and Adaptor-Binding*

As mentioned earlier, *E. coli* ClpX is composed of an N-terminal domain (N domain) followed by small and large AAA+ domains. This N domain is highly conserved and present in almost all ClpX orthologs. However, a variant of ClpX without the N domain is able to form a functional protease with ClpP, degrading ssrA-tagged substrates efficiently. In contrast, the variant is unable to support degradation of certain other substrates such as λO and MuA or interact with adaptor proteins such as SspB (Abdelhakim et al., 2008; Chien et al., 2007b; Dougan et al., 2003; Singh et al., 2001; Thibault et al., 2006; Wojtyra et al., 2003).

The N domain, comprised of the first 60 residues of ClpX, contains multiple conserved Cys residues, which bind zinc. Four of the five Cys residues (C^{14}, C^{17}, C^{36}, C^{39}) form a C4-type zinc finger to coordinate one zinc per N domain subunit whereas the fifth residue, C^{43}, does not appear to play a role. The zinc is required for N domain stability, which is a very stable dimer both in isolation and in the assembled ClpX hexamer where it is thought to function as a trimer of dimers (Donaldson et al., 2003; Wojtyra et al., 2003). Although zinc finger-containing proteins are traditionally considered DNA-binding proteins, these metal-binding motifs have been implicated in modulating protein-protein interactions (Donaldson et al., 2003; Mackay et al., 1998). NMR and crystal structures of *E. coli* ClpX N domain show that each monomer consists of an N-terminal β-hairpin followed by a C-terminal helix (Fig. 1.9). The dimer interface is formed by hydrophobic interactions between the helices of the monomers. The secondary structure of the monomer is similar to the fold in the treble clef zinc finger family member GATA-1 with which the N domain surprisingly shares very little sequence homology. However, the overall structure of the N domain dimer is unrelated to that of any known protein (Donaldson et al., 2003; Park et al., 2007).

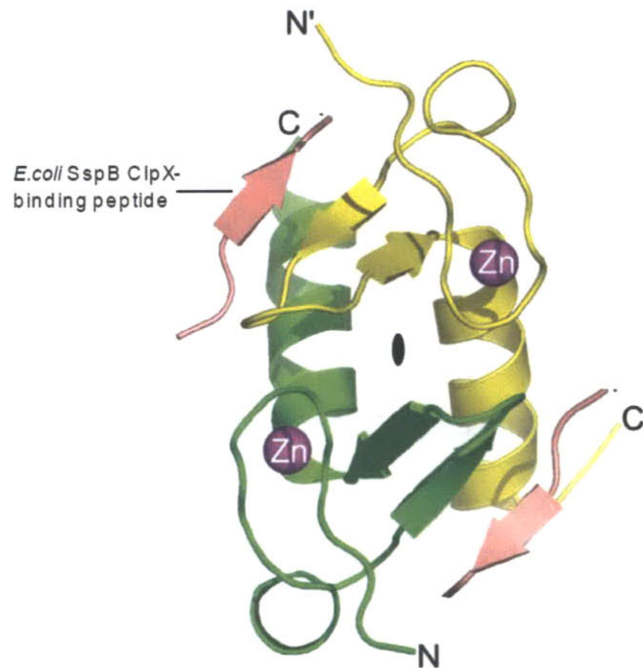


Figure 1.9. *E. coli* ClpX N domain bound to SspB ClpX-binding (XB) peptide. The N domain monomers are shown in green and yellow with labeled N- and C-termini. The XB peptide is shown in salmon and the bound zinc atoms are shown in purple. Each N domain monomer forms a β-hairpin followed by a C-terminal helix, which forms the dimer interface (Figure taken from (Park et al., 2007)).

Although the ClpX N domain clearly plays a role in substrate and adaptor selection by the ClpXP protease, how it recognizes these proteins is not well-understood. The N domain-binding proteins identified so far do not appear to share similar sequences or secondary structures. However, several of them have been shown to compete with one another for binding the N domain, indicating possible overlapping binding sites (see chapter 2). This uncommon flexibility in sequence-recognition provides ClpXP with the important ability to degrade a wide range of proteins. Interestingly, the AAA protein p97, which is a key player in the ubiquitin-proteasome system in eukaryotes, also interacts with diverse groups of proteins and this versatility enables it to participate in a variety of biological processes (Yeung et al., 2008).

B. SspB Adaptor Protein

Initially, only γ- and β-proteobacteria were thought to contain orthologs of the SspB adaptor protein despite strong conservation of the ssrA-tagging system and ClpXP. However, more

recent studies identifying SspB in the α-proteobacteria *Caulobacter crescentus* indicate that SspB is more wide-spread amongst species than previously appreciated (Chien et al., 2007b; Lessner et al., 2007). The challenge in identifying SspB adaptors in other species may be attributed to the sequence diversity amongst SspB orthologs. Indeed, the adaptors in γ- and α-proteobacteria share very little sequence identity (Chien et al., 2007a; Chien et al., 2007b; Lessner et al., 2007). Surprisingly, however, there is considerable similarity between the overall structures and functions of *E. coli* and *C. crescentus* SspB adaptors (Chien et al., 2007a; Chien et al., 2007b; Lessner et al., 2007; Levchenko et al., 2005; Levchenko et al., 2003; Song and Eck, 2003). More details about the interesting characteristics of SspB are discussed in the next sections.

1. Structure and Function

SspB was discovered by its stimulatory effect on ClpXP-mediated degradation of ssrA-tagged substrates (Levchenko et al., 2000). It binds specific substrates and delivers them to ClpXP for degradation (Fig. 1.5). This bound substrate has been shown to be directly "handed off" to the protease without needing to dissociate from the adaptor (Bolon et al., 2004a). By binding substrate and ClpX simultaneously, SspB increases the local concentration of substrate available to the protease, resulting in the enhancement of degradation rate. This increase in rate is caused by a dramatic (10-20-fold) decrease of the Michaelis-Menten constant (K_m) of degradation.

E. coli SspB is a 165 amino acid protein consisting of a conserved, folded substrate-binding domain or SBD (residues 1-117) followed by an unconserved and unstructured linker region and a short conserved C-terminal tail region (residues 155-165). The unstructured C-terminal-most peptide (^{161}LRVVK^{165}) is responsible for interacting with the N-terminal domain of ClpX (discussed above) and is termed the XB (ClpX-binding) region. Native SspB is in the form of a very stable dimer and can, thus, bind two substrate molecules simultaneously. The dimer is formed by hydrophobic interactions between the N-terminal α-helices of the two subunits. The N-terminal helix is followed by a β-sandwich, which is responsible for binding substrates (Fig. 1.10) (Levchenko et al., 2005; Levchenko et al., 2003; Song and Eck, 2003). Substrate-binding is independent of the linker and XB regions as the isolated SBD (residues 1-117) with or without the linker (residues 117-155) and tail (residues 155-165) is able to form a dimer and interact

with the ssrA-tag. However, isolated SBD (± linker/linker and tail) cannot bind ClpX and is therefore unable to increase substrate-degradation rate.

In addition to delivering ssrA-tagged substrates, *E. coli* SspB also enhances the degradation rate of a second ClpXP substrate: the N-terminal domain of the anti-sigma factor RseA (NRseA) (Flynn et al., 2004). RseA is a transmembrane protein with a cytoplasmic N-terminal domain (residues 1-108) and a periplasmic C-terminal domain. The cytoplasmic domain (NRseA) binds the sigma factor E (σ^E) to keep it inactive (De Las Penas et al., 1997; Missiakas et al., 1997). During extracytoplasmic stress, such as accumulation of unfolded proteins in the periplasm, NRseA-σ^E complex is released into the cytoplasm. The "new" C-terminal region of NRseA is actually a C-motif 1 (^{106}VAA-COO$^-$), which is recognized and subsequently degraded by ClpXP, freeing σ^E to turn on downstream stress-response genes. Furthermore, NRseA proteolysis is modulated by SspB, which binds the anti-sigma factor and delivers it to the protease for destruction. Similar to its effect on ssrA-tagged substrate degradation, SspB considerably decreases K_m of NRseA degradation (Flynn et al., 2004). Thus, there are at least two ClpXP substrates whose degradation is directly modulated by SspB.

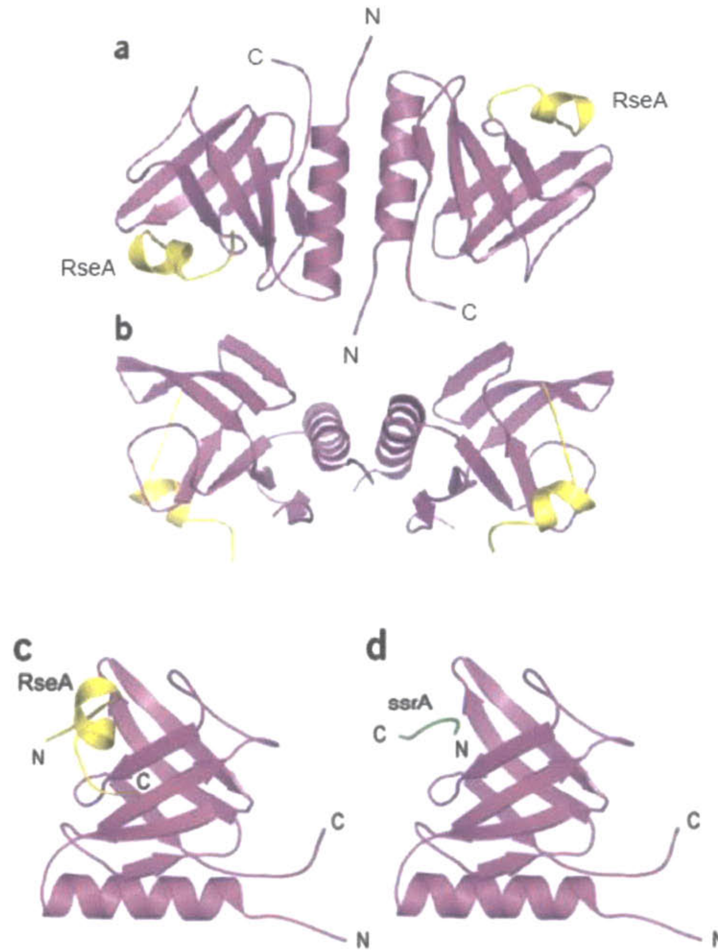


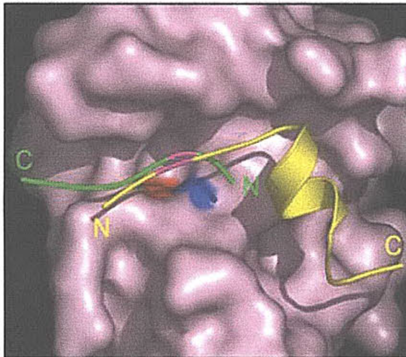
Figure 1.10. Structure of SspB with RseA and ssrA peptides. (a) is showing the stable SspB dimer with each subunit (in purple) consisting of an N-terminal α-helix followed by a β-sandwich. The dimer interface is formed by the hydrophobic interactions between the N-terminal helices. The substrate-binding domain is a hydrophobic groove formed by the β-sandwich and loops in that region. (b) is a depiction of the figure in (a) rotated by 90°. In (c) and (d), the SspB-bound RseA and ssrA peptides are in yellow and green respectively and, as shown, bind the adaptor in opposite orientations (Levchenko et al., 2005; Levchenko et al., 2003; Song and Eck, 2003). (Figure taken from (Levchenko et al., 2005)).

2. *Substrate Selection: ssrA-tag vs. NRseA*

Surprisingly, the SspB-recognition sequence in NRseA is not similar to the one in the ssrA-tag. Mutational analyses of the residues in the ssrA-tag (AANDENYALAA-COO$^-$) show that distinct residues are important for interacting with ClpX and SspB. The first 4 residues (AAND) and the 7th residue (Y) play key roles in SspB-binding whereas the C-terminal "LAA-COO$^-$" is

responsible for binding ClpX (Flynn et al., 2001; Levchenko et al., 2000). In contrast, the SspB-binding region of NRseA is longer with a distinct sequence: ^{77}EAQPAPHQWQKMPFW^{91} (Flynn et al., 2004; Levchenko et al., 2005). Co-crystal structures of SspB with ssrA or NRseA peptides indicate striking differences between the two adaptor-substrate complexes although both peptides bind overlapping sites in SspB (Fig. 1.11). Despite sharing an overlapping binding-site, the two peptides bind the adaptor in opposite orientations (Levchenko et al., 2005; Levchenko et al., 2003; Song and Eck, 2003). Surprisingly, there is actually only one common interaction in the two SspB-peptide complexes: Q^{79} of the NRseA peptide and N^3 of the ssrA peptide form hydrogen bonds with the backbone of N^{54} of SspB. This interaction contributes to the abilities of the two peptides to compete for binding the adaptor (Flynn et al., 2004). Because of the longer SspB-binding region of NRseA, it is not surprising that more contacts are made in the SspB-NRseA complex than in the SspB-ssrA peptide one. In fact, when V^{52} of SspB, which makes a contact with W^{85} of the NRseA peptide, is substituted with isoleucine, the variant SspB is defective in binding only the NRseA but not the ssrA peptide (Levchenko et al., 2005).

SspB interaction with ssrA peptide
(green) and RseA peptide (yellow)



SspB-interacting Regions

ssrA peptide: AANDENY

RseA peptide: EAQPAPHQWQKMPFW

Figure 1.11. RseA and ssrA peptides bind overlapping substrate-binding site (SBD) of *E. coli* SspB. The two peptides (ssrA: green and Rse: yellow) are bound to overlapping sites in the SBD of SspB (shown in purple) albeit in opposite orientations. There is only one conserved interaction in the two peptide-adaptor complexes, shown in pink on the peptides. The SspB interacting regions in the two peptides are shown in blue (on the right) (Levchenko et al., 2005; Levchenko et al., 2003; Song and Eck, 2003). (Figure courtesy of I. Levchenko, MIT).

3. Impact of Substrate-Binding Flexibility

The distinct peptide-binding properties of SspB have interesting biological implications. The lack of a single consensus sequence for recognition by SspB leaves the door open for interaction with ClpXP substrates beyond ssrA-tagged polypeptides and NRseA. Indeed, when Flynn et al. (2004) identified NRseA as an SspB-binding protein by comparing the ClpXP substrate profile in cells with and without SspB, their results showed that there were multiple substrates that were preferentially associated with ClpXP in the presence of SspB. By binding various ClpXP substrates, SspB could provide another layer of proteolytic regulation, especially if these substrates compete for the adaptor. The flexibility of the substrate-binding ability of SspB can also expand the substrate repertoire of ClpXP and perhaps provide degradation of SspB-interacting substrates with a competitive advantage over those that do not bind the adaptor.

Part IV. Outline of Thesis

To gain more insight into the mechanism of substrate selection by AAA+ proteases, the work in this thesis focuses on the specific interaction of the SspB adaptor with ClpXP. Chapter 2 compares the ClpX-SspB interaction in two different bacterial species: *C. crescentus* and *E. coli*. This study contributes to a clearer understanding of how ClpX, specifically its N domain, is able to recognize diverse peptide sequences and retain specificity at the same time. The appendix describes the results of a proteomic screen to investigate the effect of *E. coli* SspB on global substrate choice by ClpXP and discusses a few potential SspB-modulated ClpXP substrates. Chapter 3 is a short discussion on possible future studies stemming from this thesis work.

CHAPTER TWO

Versatile modes of peptide recognition by the ClpX N domain mediate alternative adaptor-binding specificities in different bacterial species.

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

T. C. carried out the experiments. T. C., P.C., R. T. S. and T. A. B. contributed to experimental design. T. C., R. T. S., and T. A. B. wrote the manuscript.

ABSTRACT

ClpXP, a AAA+ protease, plays key roles in protein-quality control and many regulatory processes in bacteria. The N-terminal domain of the ClpX component of ClpXP is involved in recognition of many protein substrates, either directly or by binding the SspB adaptor protein, which delivers specific classes of substrates for degradation. Despite very limited sequence homology between the *C. crescentus* and *E. coli* SspB orthologs, each of these adaptors can deliver substrates to the ClpXP enzyme from the other bacterial species. We show that the ClpX N domain recognizes different sequence determinants in the ClpX-binding (XB) peptides of *C. crescentus* SspBα and *E. coli* SspB. The *C. crescentus* XB determinants span 10 residues and involve interactions with multiple side-chains, whereas the *E. coli* XB determinants span half as many residues with only a few important side-chain contacts. These results demonstrate that the N domain of ClpX functions as a highly versatile platform for peptide recognition, allowing the emergence during evolution of alternative adaptor-binding specificities. Our results also reveal highly conserved residues in the XB peptides of both *C. crescentus* SspBα and *E. coli* SspB that play no detectable role in ClpX-binding or substrate delivery.

INTRODUCTION

Proteolysis of damaged or misfolded proteins by AAA+ proteases is essential for quality control and recycling of amino acids for new protein synthesis. It also plays a regulatory role in numerous cellular processes, including cell-cycle progression and responses to DNA damage (Gottesman, 2003). Because proteolysis occurs in crowded cellular environments with thousands of potential substrates, it is important to understand how the proper proteins are chosen for destruction. For bacterial systems, adaptor proteins and peptide signals (called degradation tags or degrons) in substrates play central roles in determining the specificity of proteolytic recognition. How adaptor proteins and degrons are recognized by AAA+ proteases is an active area of study, but only a handful of these interactions have been characterized in detail.

ClpX and ClpP assemble to form ClpXP, one of the best understood AAA+ proteases. Most biochemical studies have focused on ClpXP from *Escherichia coli*, a member of the γ-proteobacteria, but orthologs from other bacteria and mitochondria appear to have similar structures and mechanisms (Jenal and Fuchs, 1998; Kang et al., 2002; van Dyck et al., 1998). ClpP is a multi-subunit serine peptidase, in which the proteolytic active sites reside within a barrel-shaped structure (Wang et al., 1997). ClpX is a hexameric AAA+ enzyme (ATPases associated with a variety of cellular activities), which recognizes substrates and uses cycles of ATP-powered conformational changes to unfold the native protein and to translocate the denatured polypeptide into the proteolytic chamber of ClpP for degradation.

ClpX typically identifies substrates by binding degrons located near the protein termini. For example, when ribosomes stall during translation, the ssrA tag is appended onto the C-terminus of incomplete polypeptides and subsequently targets these failed translation products to ClpXP and other proteases (Gottesman et al., 1998; Keiler et al., 1996). The 11-residue ssrA tag can be recognized directly by ClpX but is also bound by an adaptor protein, SspB, which aids in delivery of substrates to ClpXP (Levchenko et al., 2000). Indeed, adaptor proteins facilitate ClpXP degradation of numerous substrates (Flynn et al., 2004; Neher et al., 2003; Zhou et al., 2001). Each ClpX subunit consists of a AAA+ domain and a ClpX-family-specific N-terminal domain, which binds zinc via a conserved set of cysteine residues and forms a stable dimer (Donaldson et al., 2003; Wojtyra et al., 2003). ClpX lacking the N domain ($ClpX^{\Delta N}$) can still bind ClpP and power degradation of some substrates (Singh et al., 2001; Wojtyra et al., 2003), establishing that the N domain is not required for the basic enzymatic functions of ClpX. However, $ClpX^{\Delta N}$ fails to recognize certain substrates and does not support degradation mediated by many adaptors (Abdelhakim et al., 2008; Chien et al., 2007a; Dougan et al., 2003; Neher et al., 2003; Singh et al., 2001; Thibault et al., 2006; Wojtyra et al., 2003).

SspB consists of a dimeric substrate-binding domain, followed by a flexible linker and a C-terminal peptide that binds the ClpX N domain (Bolon et al., 2004b; Dougan et al., 2003; Levchenko et al., 2003; Levchenko et al., 2000; Park et al., 2007; Song and Eck, 2003; Wah et al., 2003). By binding ClpX and specific substrates simultaneously, SspB increases the local concentration of substrate relative to the protease (Dougan et al., 2003; Flynn et al., 2004; Levchenko et al., 2000; Wah et al., 2002; Wah et al., 2003). As a consequence of this tethering-

mediated avidity increase, SspB enhances the rate of ClpXP degradation at sub-K_M substrate concentrations. SspB orthologs were first identified in the γ- and β-proteobacteria (Levchenko et al., 2000) and were later discovered in α-proteobacteria, including *Caulobacter crescentus* (Chien et al., 2007b; Lessner et al., 2007). The domain organization and structure of all SspB proteins are similar, but those from α-proteobacteria comprise a distinct and more distant subfamily and are therefore called SspBα. For example, the orthologs from *E. coli* (^{Ec}SspB) and *C. crescentus* (^{Cc}SspBα) share only 16% sequence identity. Nevertheless, ^{Cc}SspBα delivers substrates efficiently to *E. coli* ClpXP (^{Ec}ClpXP) (Chien et al., 2007a). The C-terminal residues of ^{Ec}SspB are known to bind to the isolated N domain of ^{Ec}ClpX, (Bolon et al., 2004a; Park et al., 2007) and a co-crystal structure has been solved [Fig. 2.1(A)] (Park et al., 2007). The N domain of ^{Cc}ClpXP and the 5 C-terminal amino acids of ^{Cc}SspBα are also required for adaptor-mediated substrate delivery (Chien et al., 2007b), suggesting a corresponding binding relationship.

Here, we probe the fine specificity of the interaction of ^{Cc}SspBα with ^{Cc}ClpX and ^{Ec}ClpX. In both cases, the 10 C-terminal residues of ^{Cc}SspBα comprise the ClpX-binding (XB) region that tethers the adaptor to the N domain. Mutational analyses show that seven side-chains in ^{Cc}SspBα XB contribute to adaptor-enzyme recognition, and all SspBαs have homologous sequences that maintain the chemical character of these residues. Surprisingly, however, the corresponding XB peptide of ^{Ec}SspB is shorter, shares little meaningful homology, and displays a radically different mutational profile with just a few residues playing major roles in recognition. Again, these features appear to be shared by SspB orthologs in other γ- and β-proteobacteria. Nevertheless, we find that ^{Ec}SspB delivers substrates to ^{Cc}ClpXP for degradation. Thus, the N domains of both ^{Ec}ClpX and ^{Cc}ClpX have the ability to recognize two different XB peptides. Apparently, these domains possess distinct peptide-binding specificities and have adopted alternative but non-exclusive modes of adaptor-binding during the evolution of different bacterial lineages. We also find that some highly conserved amino acids in the XB peptides of ^{Cc}SspBα and ^{Ec}SspB play no obvious roles in substrate delivery or ClpX-binding and suggest that these amino acids may help protect the adaptors from degradation during substrate delivery.

RESULTS

Phylogenetic comparisons suggest use of different adaptor-tethering contacts.

A multiple sequence alignment of the C-terminal regions of more than 100 SspBα orthologs revealed a conserved block of residues [Fig. 2.1(B)]. Within this region, ^{Cc}SspBα residues 153, 154, 156, 157, 159, 160, and 161 were most highly conserved. Deletion of a portion of this C-terminal region prevents substrate delivery by ^{Cc}SspBα to ClpXP (Chien et al., 2007b). Sequence conservation near the C-terminus of SspB orthologs from γ- and β-proteobacteria revealed a very different pattern of homology [Fig. 2.1(C)], suggesting that SspBα interacts with ClpX in a fashion distinct from their γ and β counterparts.

Previous studies showed that the ^{Ec}ClpX N domain binds the ^{Ec}SspB XB peptide and is important for efficient substrate delivery by ^{Cc}SspBα (Bolon et al., 2004a; Chien et al., 2007a; Park et al., 2007), suggesting that the ^{Cc}SspBα XB region directly binds the N domains of ^{Ec}ClpX and ^{Cc}ClpX. To test this idea, a peptide consisting of the C-terminal decapeptide of ^{Cc}SspBα preceded by a fluorescent dye and tyrosine was synthesized for binding studies monitored by fluorescence anisotropy. As shown in Figure 2.1(D), this ^{Cc}SspBα peptide was bound with similar affinity (K_D ~25 μM) by the purified ^{Cc}ClpX and ^{Ec}ClpX N domains. Thus, the N domains of both ClpX enzymes, which share ~60% sequence identity (Fig. 2.S1 (supplemental)), have the ability to recognize very different XB-peptide sequences.

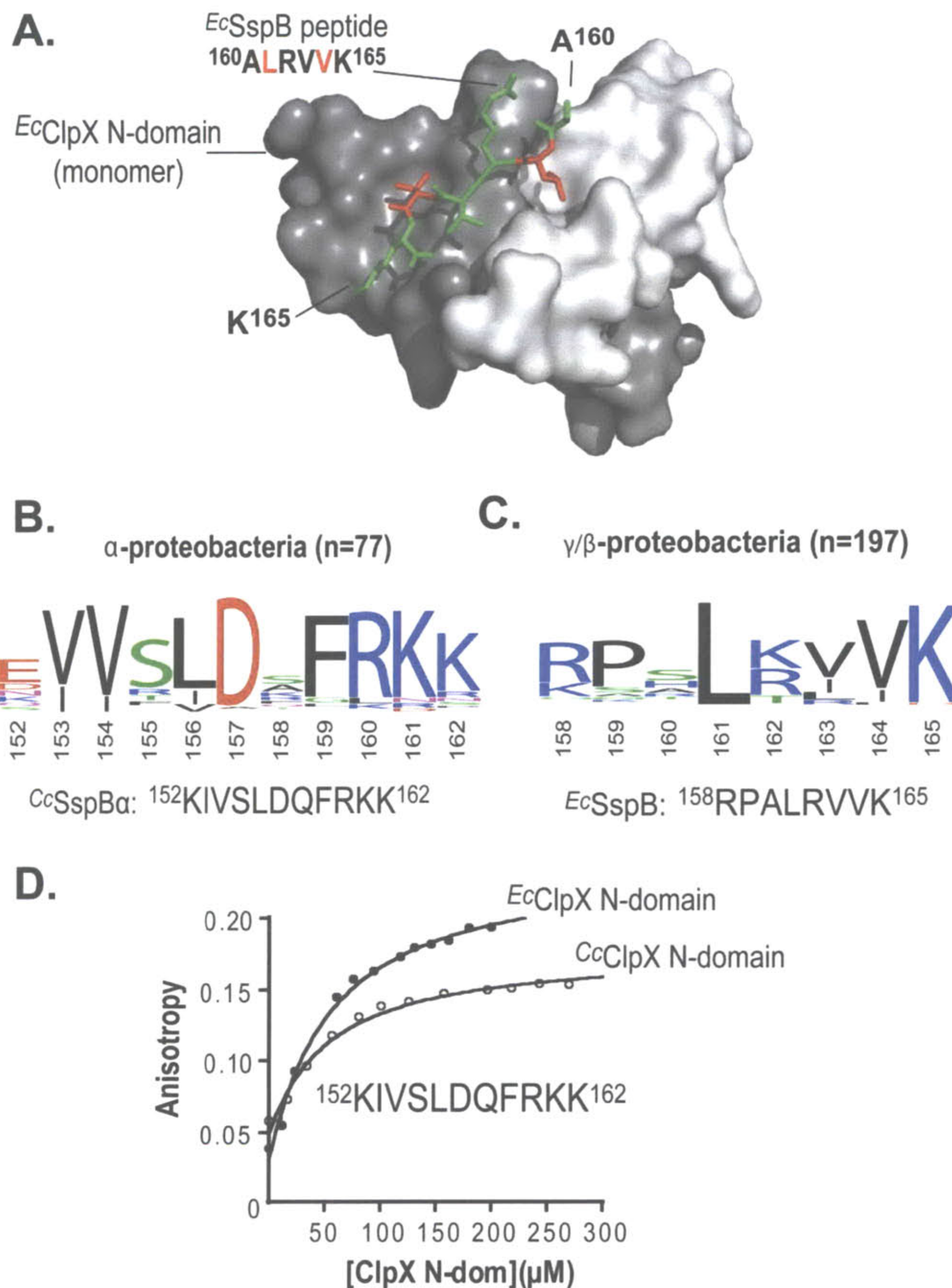


Figure 2.1. XB conservation and cross-species interaction of SspB and ClpX from α- and γ/β-proteobacteria.

A. The structure of the *E. coli* ClpX N domain dimer bound to the C-terminal peptide of *E. coli* SspB shows the hydrophobic pockets of the N domain monomers occupied by the L^{161} and V^{164} residues of the peptide (Park et al., 2007). The N domain monomers are shown in dark and pale gray and the peptide in green with L^{161} and V^{164} highlighted in red.
B. Weblogo (Crooks et al., 2004) depiction of sequence conservation within the C-terminal regions of 77 SspB proteins from α-proteobacteria. Alignments were performed using

Jalview (Clamp et al., 2004). The C-terminal region of the α-proteobacteria *C. crescentus* SspB is also depicted.

C. Sequence conservation in the C-terminal regions of 197 SspB proteins from γ/β-proteobacteria reveals a very different pattern than observed in panel A. The C-terminal region of the γ -proteobacteria *E. coli* SspB is also shown.

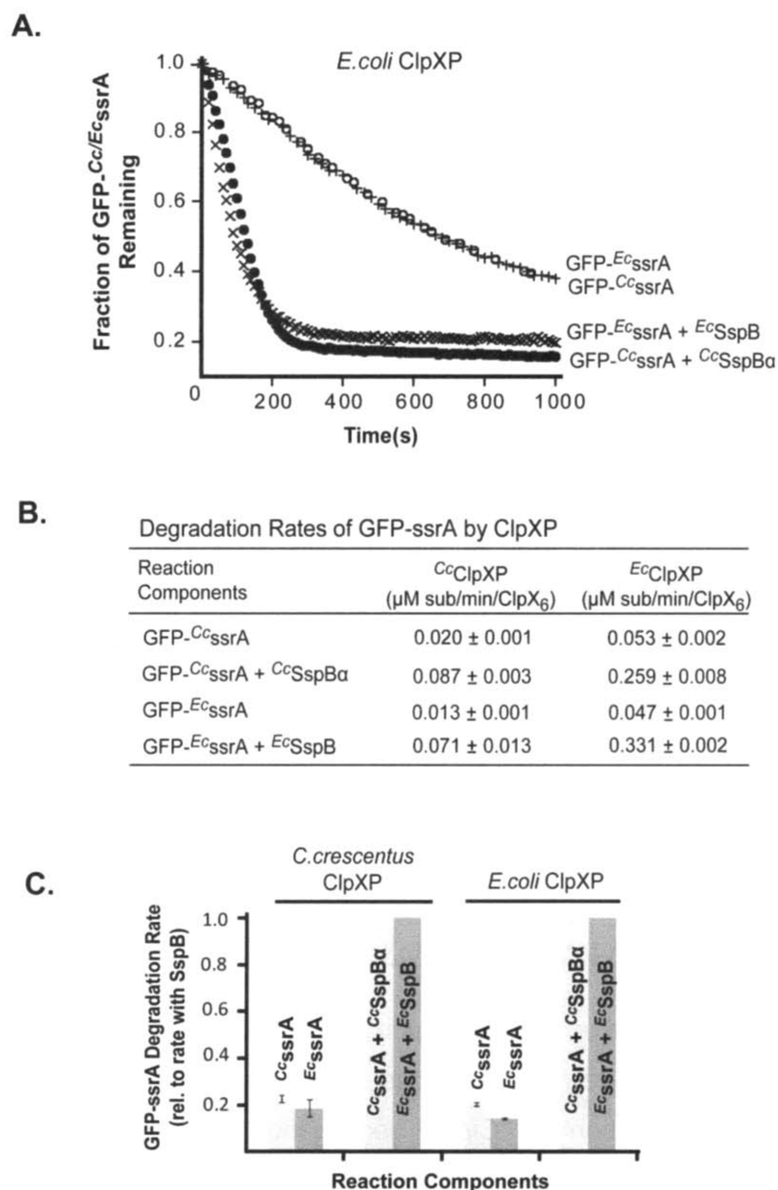
D. Binding of the N domains from *C. crescentus* (K_D 25 μM) or *E. coli* (K_D 25 μM) ClpX to a fluorescein-labeled peptide (60 nM) corresponding to the XB region of ^{Cc}SspBα.

```
                 1        10        20        30        40        50        60
       E. coli   MTDKRKDGSGKLLYCSFCGKSQHEVRKLIAGPSVYICDECVDLCNDIIREEIKEVAPHRER
C. crescentus    MTKAASGDTKSTLYCSFCGKSQHEVRKLIAGPTVFICDECVELCMDIIREEHK-IAFVKSK
```

Figure 2.S1. Comparison of ClpX N-terminal domains of *E. coli* and *C. crescentus*. Identical residues are highlighted in lilac. Alignment was performed using Jalview (Clamp M, Cuff J, Searl SM, Burton JG (2004) The Jalview Java alignment editor, Bioinformatics 20(3):426-7).

Adaptor delivery of cognate substrates to ^{Cc}ClpXP and ^{Ec}ClpXP.

For studies of adaptor stimulation of degradation, we used green fluorescent protein (GFP) bearing either a *C. crescentus* ssrA tag (AANDNFAEEFAVAA; GFP-^{Cc}ssrA), which binds well to ^{Cc}SspBα, or an *E. coli* ssrA tag (AANDENYALAA; GFP-^{Ec}ssrA), which binds well to ^{Ec}SspB (Chien et al., 2007a; Wah et al., 2003). As anticipated (Chien et al., 2007a; Chien et al., 2007b), ^{Cc}SspBα stimulated degradation of GFP-^{Cc}ssrA by the ^{Cc}ClpXP protease and by the ^{Ec}ClpXP enzyme (Fig. 2.2). Importantly, ^{Ec}SspB also stimulated degradation of GFP-^{Ec}ssrA by both ^{Cc}ClpXP and ^{Ec}ClpXP [Fig. 2.2(B,C)]. Thus, despite minimal XB-sequence homology, the adaptors from *C. crescentus* and *E. coli* were both able to stimulate degradation of cognate substrates by the ClpXP enzyme from the other species.



Degradation Rates of GFP-ssrA by ClpXP

Reaction Components	^{Cc}ClpXP (μM sub/min/ClpX$_6$)	^{Ec}ClpXP (μM sub/min/ClpX$_6$)
GFP-^{Cc}ssrA	0.020 ± 0.001	0.053 ± 0.002
GFP-^{Cc}ssrA + ^{Cc}SspBα	0.087 ± 0.003	0.259 ± 0.008
GFP-^{Ec}ssrA	0.013 ± 0.001	0.047 ± 0.001
GFP-^{Ec}ssrA + ^{Ec}SspB	0.071 ± 0.013	0.331 ± 0.002

Figure 2.2. *C. crescentus* and *E. coli* ClpXP interact with both ^{Cc}SspBα and ^{Ec}SspB.

A. ^{Cc}SspB and ^{Ec}SspB (1.2 μM) both enhanced degradation of their cognate GFP-ssrA by ^{Ec}ClpXP as monitored by decreases in fluorescence. ClpXP concentration in each case was 0.1 μM and the substrate concentration was 0.1 μM. This low substrate concentration (sub-K_M) was used to help ensure that degradation was adaptor-stimulated.

B. Both ^{Cc}ClpXP and ^{Ec}ClpXP were able to degrade GFP containing either the ^{Cc}ssrA or the ^{Ec}ssrA tags. Protein concentrations used were as in (A). The rate of substrate degradation was enhanced by the cognate adaptor SspB. Under the purification conditions used in this study, ^{Cc}ClpX was less active compared to ^{Ec}ClpX.

C. Normalized degradation rates of GFP-^{Cc}ssrA (pale gray bars) or GFP-^{Ec}ssrA (dark gray bars) by ^{Cc}ClpXP and ^{Ec}ClpXP in the presence or absence of ^{Cc}SspBα or ^{Ec}SspB; in this case, the

adaptor species (*Cc* vs. *Ec*) matched that of the ssrA tag sequence on the substrate. Protein concentrations were as described in (A).

We also constructed a chimera, consisting of the substrate-binding domain of ^{Ec}SspB followed by the ^{Cc}SspBα C-terminal linker and XB region. This chimeric adaptor enhanced ^{Ec}ClpXP degradation of GFP-^{Ec}ssrA [Fig. 2.3(A)], establishing that tethering interactions mediated by ^{Cc}SspBα XB can replace the interactions normally mediated by ^{Ec}SspB XB.

To address the importance of residues near the C-terminus of ^{Cc}SspBα, we constructed truncated variants and assayed their adaptor activity. The last two lysine residues (K^{161}K^{162}) could be deleted without a major effect on delivery, whereas deletion of additional upstream residues eliminated activity [Figure 2.3(B)]. However, the substrate-binding domain of ^{Ec}SspB followed by the C-terminal residues ^{158}QFRKK^{162} of ^{Cc}SspBα was inactive as an adaptor (data not shown), establishing that the ^{158}QFR sequence may be necessary but is not sufficient for adaptor function.

<u>Residues involved in ^{Cc}SspBα tethering to ClpX.</u>

To determine which residues in the XB region of ^{Cc}SspBα are important for ClpX-binding, we individually mutated the 10 C-terminal residues to alanine and purified these variants. In one set of assays, we tested stimulation of ^{Cc}ClpXP degradation of 0.1 μM GFP-^{Cc}ssrA (Fig. 2.4). This substrate concentration is below K_M (~1 μM) for unassisted ^{Cc}ClpXP degradation, allowing adaptor-mediated stimulation to be observed. Alanine substitutions at I^{153}, V^{154}, L^{156}, R^{160}, and K^{162} caused the largest defects in substrate delivery [Fig. 2.4(A)]. Milder effects were observed for substitutions at the other positions, with mutations at S^{155} and D^{157} having essentially no effect on delivery activity. The alanine mutations in ^{Cc}SspBα had generally similar effects on adaptor-mediated degradation by ^{Ec}ClpXP. However, V154A appeared to be more active with the ^{Ec}ClpXP protease than with ^{Cc}ClpXP [Figs. 2.4(A) and (B)], suggesting that this residue plays a somewhat different role in binding the two enzymes.

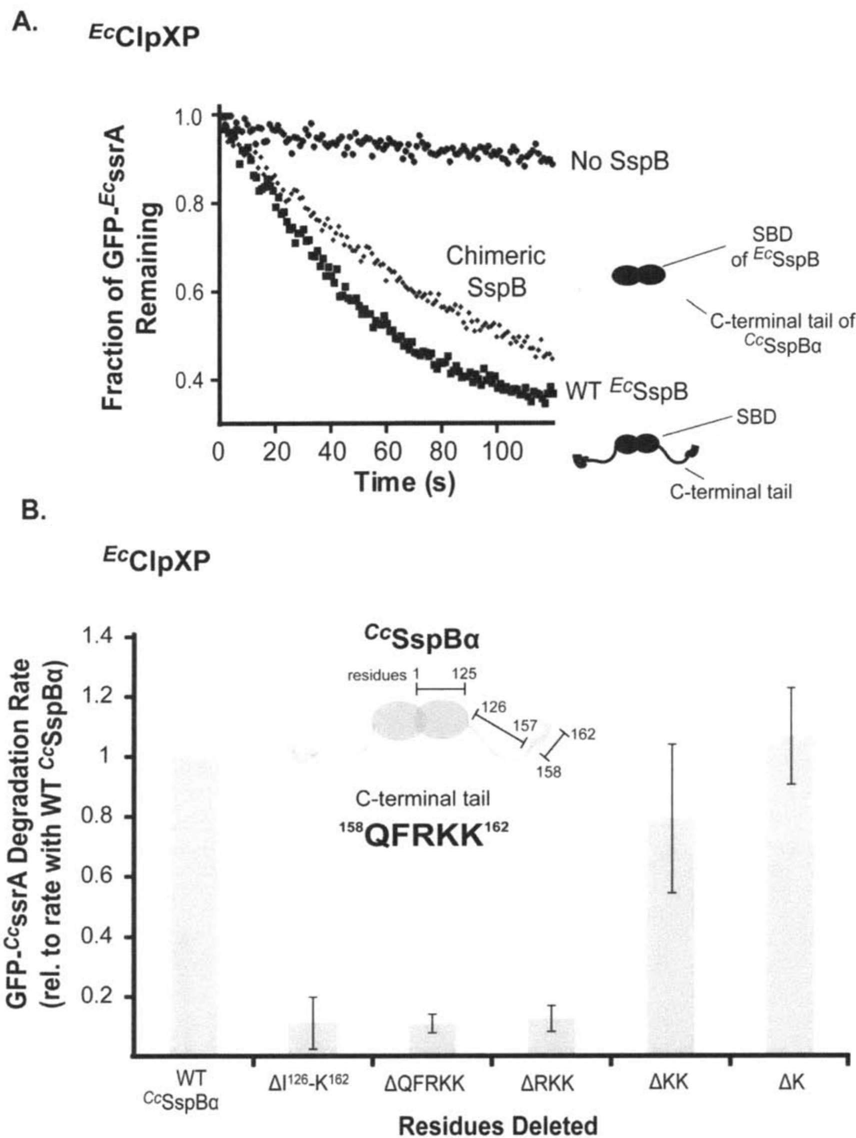


Figure 2.3. ^{Cc}SspBα interacts with ClpXP via its C-terminal region.

A. A chimeric SspB (0.3 μM), obtained by substituting the C-terminal region of ^{Ec}SspB with that of ^{Cc}SspBα, enhanced the degradation of GFP-^{Ec}ssrA (0.1μM) by ^{Ec}ClpXP (0.05 μM).
B. Different segments of the C-terminal region of ^{Cc}SspBα were removed and the variant adaptors tested for their ability to enhance degradation of GFP-^{Ec}ssrA (0.1 μM) by ^{Ec}ClpXP (0.05 μM).

Most alanine substitutions did not reduce activity to the level of unassisted ClpXP degradation, suggesting that these mutations weaken but do not eliminate tethering. To test if more dramatic

mutations had larger effects, we also constructed and purified variants in which I^{153}, V^{154}, L^{156}, and R^{160} were changed to aspartic acid. Except for R160D, these substitutions decreased ^{Cc}ClpXP degradation to the level of the no-SspB control when assayed with ^{Cc}ClpXP [Fig. 2.4(C)]. The aspartate substitutions were also more severe than the alanine substitutions in ^{Ec}ClpXP degradation assays [Fig. 2.4(C)].

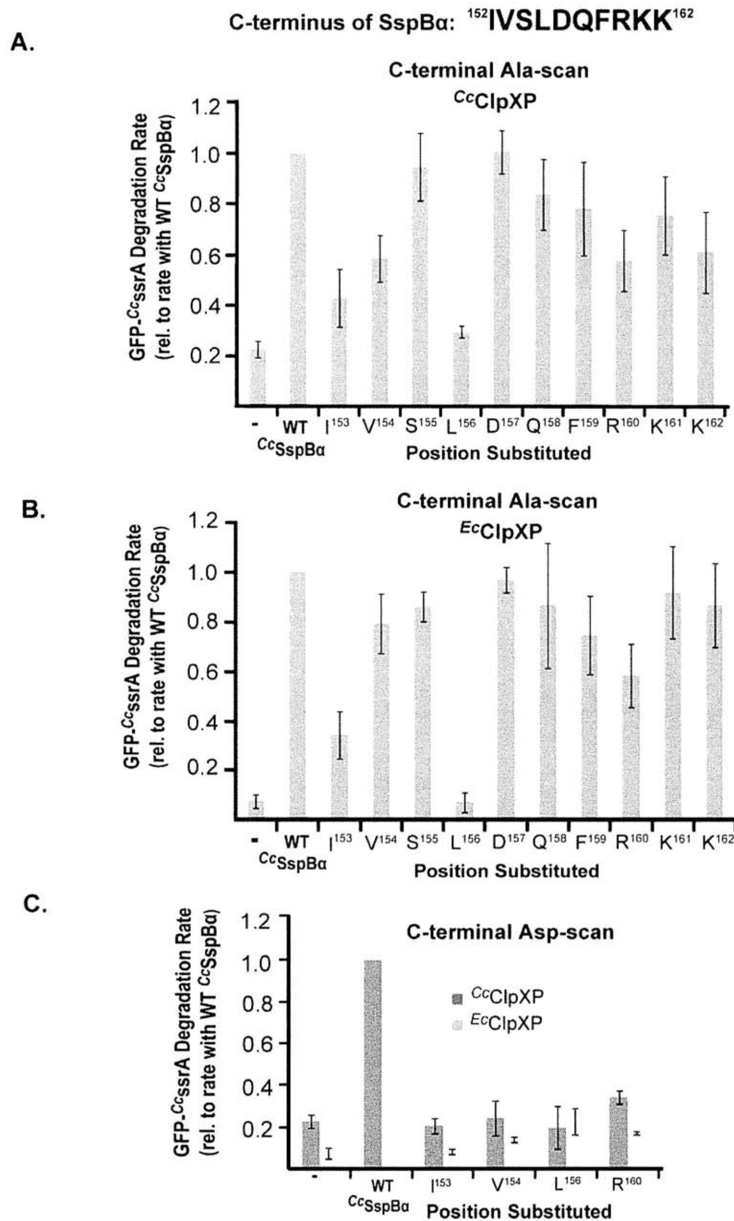


Figure 2.4. Substrate-delivery activity of ^{Cc}SspBα variants with substitution mutations.

A. Variants of ^{Cc}SspBα with Ala substitutions in the C-terminal residues were assayed for their delivery activity by determining how well they enhanced degradation of GFP-^{Cc}ssrA (0.1 μM) by ^{Cc}ClpXP (0.2 μM). All rates were normalized to the degradation rate of GFP-^{Cc}ssrA in the presence of WT ^{Cc}SspBα. Wild-type ^{Cc}SspBα or variants were present at 0.3 μM.
B. Analysis is the same as in (A), except that the protease was ^{Ec}ClpXP (at 0.05 μM). Similar results were observed, however the V154A variant appeared to be more active with ^{Ec}ClpXP than with ^{Cc}ClpXP.
C. Aspartate-substitutions were made at four positions (I^{153}, V^{154}, L^{156}, and R^{160}) in ^{Cc}SspBα and the variants (0.3 μM) tested for adaptor function with ^{Cc}ClpXP (0.2 μM) and ^{Ec}ClpXP (0.05 μM). The rates were normalized to the degradation rate of GFP-^{Cc}ssrA in the presence of WT ^{Cc}SspBα.

Assays of ^{Cc}SspBα-mediated stimulation of GFP-^{Cc}ssrA degradation have limited dynamic range because this substrate is degraded reasonably well (K_M ~1-2 μM) by ClpXP alone [Fig. 2.5(B)]. To address this concern, we changed the C-terminal residues of this substrate from VAA to DAS (GFP-^{Cc}DAS) [Fig. 2.5(A)]. This substitution weakens ^{Ec}ClpX recognition of the ssrA tag and increases the adaptor-dependence of degradation (McGinness et al., 2006). When we assayed ^{Cc}SspBα stimulation of ^{Cc}ClpXP degradation of GFP-^{Cc}DAS, the I153A, V154A, L156A, R160A, and K162A mutations caused substantial reductions in the stimulated degradation rate whereas the D157A, F159A, and K161A substitutions had only modest effects [Fig. 2.5(C)].

Substitution of alanine for S^{155} or D^{157} in ^{Cc}SspBα did not have a large effect on substrate delivery. It seemed possible, however, that proline substitutions at these positions might interfere with ClpX-binding by disrupting conformations (for example, an α-helix or β-strand) of the entire XB peptide and thus interfering with contacts made by residues flanking these positions. However, proline substitutions at either position caused only minor reductions in the ability of these ^{Cc}SspBα variants to stimulate ^{Cc}ClpXP degradation (Fig. 2.6). It appears, therefore, that the XB peptide of ^{Cc}SspBα binds in a conformation compatible with the restrictions of the backbone dihedral angle that would be enforced by proline at these positions.

A. ^{Cc}ssrA tag: AANDNFAEEFA**VAA** → ^{Cc}DAS: AANDNFAEEFA**DAS**

^{Ec}ssrA tag: AANDENYA**LAA** → ^{Ec}DAS(+4): AANDENYA**SENYDAS**

B.

C. Ala-scan: ^{153}IVSLDQFRKK^{162}

Figure 2.5. Degradation of GFP-^{Cc}DAS by ^{Cc}ClpXP.

A. The ssrA tag in *C. crescentus* (^{Cc}ssrA) consists of 14 amino acids, the last three (VAA) resembling the terminal "LAA" residues of the 11-residue *E. coli* ssrA tag (^{Ec}ssrA). Substituting the last three residues of either tag with the sequence "DAS" made the tag a weaker degron, and thus proteolysis of tagged protein more adaptor-dependent. For ^{Ec}ssrA, the tag was also elongated (to remove an SspB-ClpX clash) to generate the DAS+4 tag.
B. Cartoon depicting ssrA-DAS recognition. The efficiency of substrate degradation depends on the protein-protein interactions occurring in the ternary complex formed between the substrate (GFP-ssrA), the adaptor (SspB), and ClpX (*left panel*). When the tag is modified

such that there is only a weak substrate-ClpX interaction (e.g. the DAS tags where used), adaptor•ClpX interactions dictate the efficiency of degradation (*right panel*).

C. The ^{Cc}SspBα C-terminal residues were individually changed to alanine and assayed for their ability to enhance degradation of GFP-^{Cc}DAS (0.1 μM) by ^{Cc}ClpXP (0.2 μM). ^{Cc}SspBα variants were present at 0.3 μM. (*Top*) Degradation traces for representative ^{Cc}SspBα variants. (*Bottom*) Summary of the alanine-scan results of the C-terminal region of ^{Cc}SspBα. The rates were normalized to the degradation rate of GFP-^{Cc}DAS (0.1 μM) in the presence of WT ^{Cc}SspBα (0.3 μM).

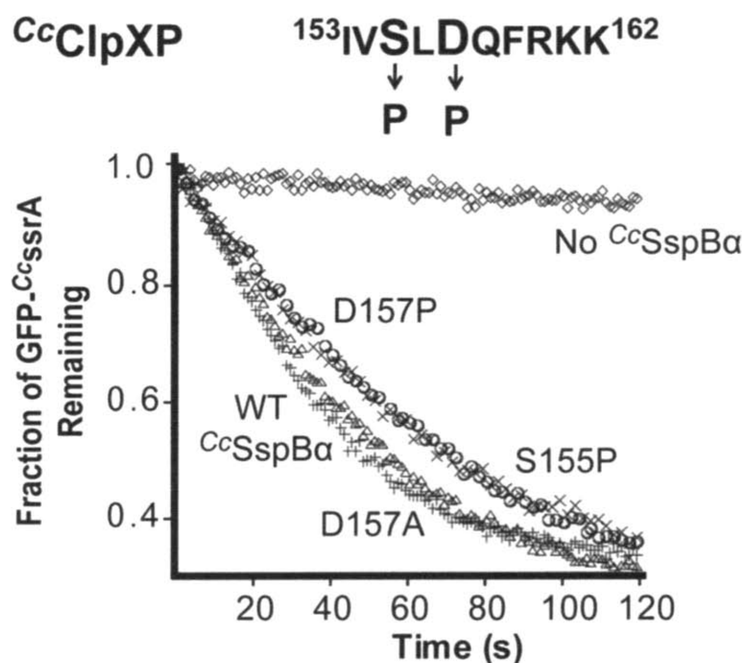


Figure 2.6. Secondary structure in the ^{Cc}SspBα C-terminal region is not critical.
S^{155} or D^{157} (shown in bold in sequence, *Top*) in the C-terminal region of ^{Cc}SspBα were substituted with proline, which disrupts secondary structure, and the proteins were assayed for their ability to deliver GFP-^{Cc}ssrA (0.1 μM) to ^{Cc}ClpXP (0.2 μM). Neither substitution inhibited adaptor function.

<u>^{Cc}SspBα XB mutations decrease N-domain affinity.</u>

We anticipated that the alanine substitutions in the XB region of ^{Cc}SspBα would reduce affinity for the ^{Cc}ClpXP N domain. To test this idea directly, we synthesized fluorescent XB-peptide variants and assayed binding. Alanine substitutions for I^{153}, V^{154}, L^{156}, F^{159}, R^{160}, and K^{162} decreased affinity to varying extents (Fig. 2.7, Table 2.1). Substitutions at the remaining

positions had very small effects. These results largely mirror the defects in substrate delivery for the corresponding substitutions in full-length ^{Cc}SspBα.

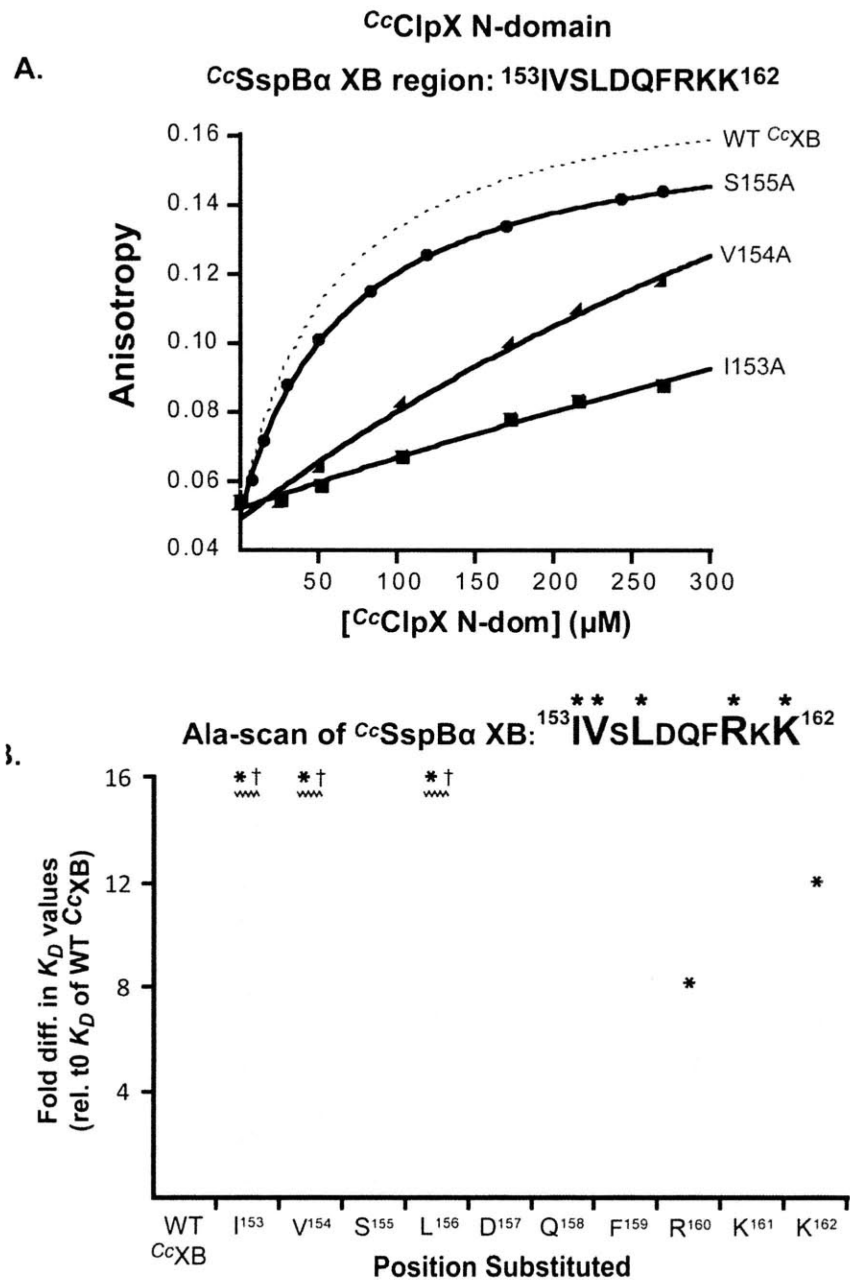


Figure 2.7. Binding interaction between ^{Cc}ClpX N domain and ^{Cc}SspBα.

A. An alanine-scan of a peptide consisting of the C-terminal residues of ^{Cc}SspBα was done and the peptides were tested for their ability to bind the ^{Cc}ClpX N domain. The wild-type sequence is shown in dashed gray [see Fig. 1(C)]. Peptide concentration was 60 nM. K_D values of all the ^{Cc}SspBα C-terminal peptide variants are shown in Table 2.1.

B. Comparison of the change in K_D value caused by the Ala-substitution (relative to the value of WT ^{Cc}SspBα XB). Residues I^{153}, V^{154}, L^{156}, R^{160}, and K^{162} (marked with *) were most important for binding ^{Cc}ClpX N domain. The K_D values of Ala-substitutions at I^{153}, V^{154}, and L^{156} (marked with †) were > 16-fold higher than that of WT adaptor and could not be depicted within the scale of the y-axis.

Table 2.1. K_D values of ^{Cc}SspBα C-terminal peptides binding to *C. crescentus* ClpX N domain (dimer)

Peptide	Sequence	K_D (μM)
WT	YKIVSLDQFRKK	26.0 ± 4.8
I153A	YK**A**VSLDQFRKK	> 400
V154A	YKI**A**SLDQFRKK	> 400
S155A	YKIV**A**LDQFRKK	32.1 ± 2.3
L156A	YKIVS**A**DQFRKK	> 400
D157A	YKIVSL**A**QFRKK	50.9 ± 12
Q158A	YKIVSLD**A**FRKK	23.1 ± 3.9
F159A	YKIVSLDQ**A**RKK	66.9 ± 13
R160A	YKIVSLDQF**A**KK	200 ± 82
K161A	YKIVSLDQFR**A**K	39.3 ± 5.4
K162A	YKIVSLDQFRK**A**	303 ± 100

Effects of ^{Ec}SspB XB mutations on substrate delivery and N-domain affinity.

As shown in Figure 2.1 (B,C), phylogenetic comparisons reveal very different patterns of sequence conservation for the XB peptides of SspB orthologs from the γ- and β-proteobacteria as opposed to those from α-proteobacteria. To probe the functional importance of residues in the ^{Ec}SspB XB peptide, we purified alanine-substituted variants and assayed their ability to

deliver GFP-^{Ec}DAS to ^{Ec}ClpXP. The largest defects were observed for the L161A and V164A variants [Fig. 2.8(A)]. For comparison, this figure also shows the relative abilities of the same variants to deliver a normal ssrA-tagged substrate, a similar but less sensitive assay (Wah et al., 2003).

To determine the effects of the alanine substitutions on the affinity of the ^{Ec}SspB XB peptide for the ^{Ec}ClpX N domain, we carried out peptide-binding experiments [Fig. 2.8(B)]. The L161A substitution made N-domain binding too weak to measure, the V164A substitution decreased binding ~7-fold, whereas smaller effects were detected for the remaining substitutions. These results agree well with the functional studies. Importantly, they establish that the N domain of ClpX recognizes ^{Cc}SspBα XB peptides in a substantially different manner than the XB peptides from γ- and β-proteobacterial SspBs.

Different XB sequences compete for binding the N domain.

Because ^{Cc}XB and ^{Ec}XB sequences have such distinct features, we sought to determine if they bound distinct sites in the N domains of ^{Cc}ClpX and ^{Ec}ClpX. Therefore, degradation of GFP-^{Ec}DAS by either ^{Cc}ClpXP or ^{Ec}ClpXP was performed in the presence of ^{Ec}SspB with or without high concentrations of ^{Cc}XB peptide. As shown in Fig. 2.9 (A), this peptide inhibited degradation by both proteases, indicating that it competes with ^{Ec}SspB for interaction with these enzymes. Furthermore, GFP-^{Ec}ssrA degradation, which is less adaptor-dependent, was also inhibited by the ^{Cc}XB peptide [Fig. 2.9(B)]. Importantly, ^{Cc}XB only inhibited reactions in which substrate delivery was promoted by an adaptor (^{Ec}SspB). Thus, these results indicate that the *C. crescentus* and *E. coli* XB peptides bind the same or overlapping sites on the N domain of ClpX.

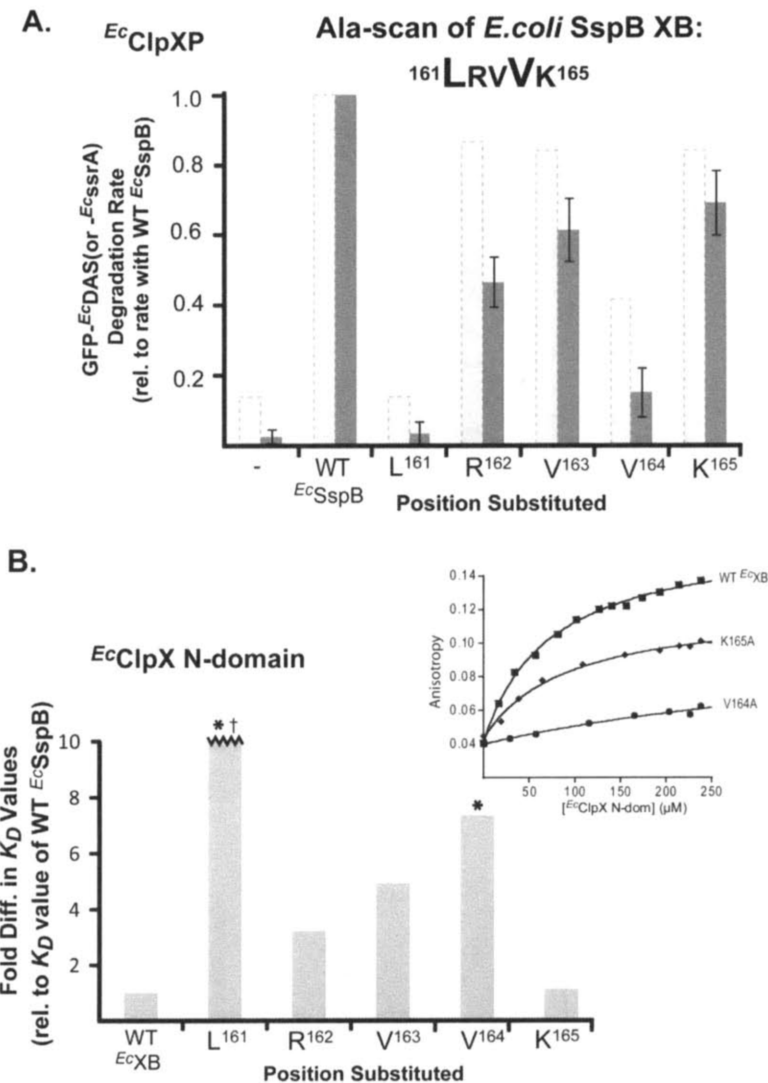


Figure 2.8. Functional and binding interactions between ^{Ec}ClpX N domain and ^{Ec}SspB.

A. The 5 C-terminal residues of ^{Ec}SspB were changed individually to Ala and the variants tested for function by monitoring degradation of GFP- ^{Ec}DAS (dark gray bars). Reactions contained 0.05 μM ^{Ec}ClpXP, 0.3 μM ^{Ec}SspB variant and 0.1 μM substrate. Rates were normalized to the rate of degradation of GFP-^{Ec}DAS in the presence of WT ^{Ec}SspB. The L161A and V164A variants were the most defective. A similar result was observed by Wah et al. (2003) who used GFP- ^{Ec}ssrA as a substrate to test the function of these Ala-variants (shown in pale gray/dashed bars). As expected, the activities of all variants were higher for degradation of WT GFP-ssrA than the DAS variant.
B. Peptides (60 nM) corresponding to an alanine-scan of the ^{Ec}SspB C-terminal region were tested for ^{Ec}ClpX N-domain-binding. The K_D values were determined by fluorescence anisotropy (inset). Residues L^{161} and V^{164} (marked with asterisk) were the most important for the interaction. The K_D value of L161A (marked with †) was > 10-fold higher than that of WT adaptor and could not be depicted within the scale of the y-axis.

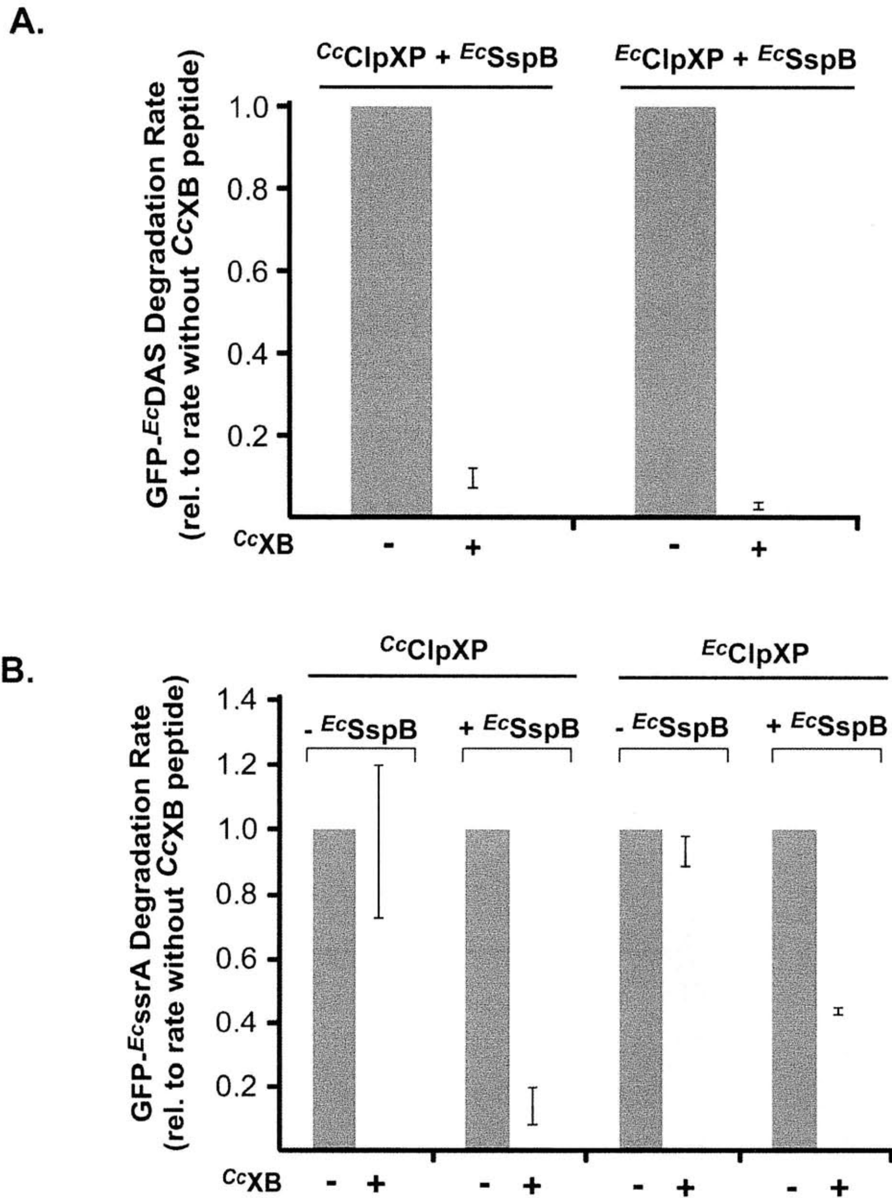


Figure 2.9. Competition between ^{Cc}SspB XB and ^{Ec}SspB XB motifs for binding ^{Cc}ClpX and ^{Ec}ClpX.

A. GFP-^{Ec}DAS (0.1 μM) degradation reactions were set-up using either ^{Ec}ClpXP (0.05 μM) or ^{Cc}ClpXP (0.2 μM) as the protease. When a high concentration of ^{Cc}XB peptide (200 μM) was added, GFP-^{Ec}DAS (0.1 μM) degradation was inhibited. Degradation rates were normalized to that of GFP-^{Ec}DAS in the presence of WT ^{Ec}SspB without any competitor peptide.
B. ClpXP-mediated degradation of GFP- ^{Ec}ssrA (0.1 μM) in the presence of ^{Ec}SspB (0.3 μM) was inhibited when 200 μM ^{Cc}XB (YKIVSLDQFRKK) peptide was added to the reaction. This result was observed with ClpXP from both *C. crescentus* and *E. coli*. In the absence of

^{Ec}SspB, inhibition was not observed, indicating that peptide-inhibition was due to disruption of the ^{Ec}SspB-ClpX interaction rather than a direct effect on the enzyme.

DISCUSSION

As previously shown for *E. coli* SspB, the C-terminal residues of *C. crescentus* SspBα mediate its interaction with ClpX. Surprisingly, however, phylogenetic conservation and studies of mutant adaptors and XB peptides indicate that ClpX recognizes ^{Cc}SspBα and ^{Ec}SspB in very different ways. For instance, the XB regions of orthologs from α-proteobacteria are longer and appear to make many more side-chain contacts with ClpX than the XB regions of orthologs from γ- and β-proteobacteria. Nevertheless, the N-terminal domain of ClpX from either *E. coli* or *C. crescentus* is able to bind both XB peptides. As a consequence, both ^{Cc}SspBα and ^{Ec}SspB can deliver cognate substrates to the ClpXP proteases from *C. crescentus* or *E. coli.* These results establish that the N domains of ClpX from both species have at least two distinct peptide-binding specificities. One of these modes of binding appears to be exclusively used for SspBα recognition in the α-proteobacteria, whereas the other seems to be employed for SspB tethering in all γ- and β-proteobacteria.

Mutant studies presented here and previously (Wah et al., 2003) indicate that just two residues in the XB peptide of ^{Ec}SspB, L^{161} and V^{164} (LRVVK^{165}), play major roles in ClpX recognition, with the leucine side-chain being most important. Because the XB peptide of ^{Cc}SspBα (IVSLDQFRKK^{162}) contains a leucine separated by two residues from a phenylalanine, it might be argued that these peptides bind the N domain of ClpX in generally similar ways. However, several results are difficult to reconcile with this model. First, the LxxF sequence in the ^{Cc}SspBα peptide is still present in the ΔRKK chimeric variant, which fails to deliver substrates [see Fig. 2.3(B)], and in the ΔRKK ^{Cc}SspBα XB peptide, which fails to bind the N domain of ^{Cc}ClpX (data not shown). Second, the crystal structure of a complex of ^{Ec}ClpX N domain with ^{Ec}XB peptide shows that the residue immediately following L^{161} (L**R**VVK) adopts dihedral angles that would be inaccessible to proline (Park et al., 2007). In contrast, our results show that the ^{Cc}SspBα XB variant IVSL**P**QFRKK is active in delivering substrates to ClpXP (see Fig. 2.6). Third, the side-chain of V^{164} in the ^{Ec}SspB XB peptide packs into a hydrophobic pocket in the ^{Ec}ClpX N domain that is too small to accommodate a phenylalanine side-chain (Park et al., 2007). It seems most likely, therefore, that these peptides bind the N domain in fundamentally different fashions.

However, competition experiments suggest that both XB peptides bind at the same or overlapping sites in the ClpX N domain (see Fig. 2.9). Thus, the same general peptide-binding pocket may have an unusual amount of flexibility in potential modes of binding specificity.

One minor anomaly in analysis of the ^{Cc}SspBα XB region concerns differences between experiments using deletions and alanine substitutions. Specifically, we found that one (K^{162}) or two ($K^{161}K^{162}$) C-terminal residues could be deleted from the ^{Cc}SspBα XB peptide without causing significant defects in substrate delivery (see Fig. 2.3). We also synthesized ΔKK and ΔRKK XB peptides and assayed binding to the N domain of ^{Cc}ClpX. The ΔKK peptide bound the N domain with reduced but substantial affinity, whereas almost no binding was detected for the ΔRKK peptide (data not shown). By contrast, alanine substitutions at positions 161 and 162 reduced substrate-delivery activity in some assays [see Fig. 2.5(C)]. These results could be reconciled if contacts between these lysine side chains and ClpX stabilize the complex, whereas contacts mediated by the peptide backbone of these residues destabilize binding to a roughly comparable extent. Prior studies have also shown that deletion of the C-terminal lysine (K^{165}) of ^{Ec}SspB does not affect substrate delivery (Wah et al., 2003). However, the co-crystal structure shows that this lysine side-chain makes numerous intimate contacts with the ^{Ec}ClpX N domain (Park et al., 2007). Indeed, based on the structure alone, it would be reasonable to suggest that this lysine plays an important role in ClpX-binding, and yet we detected only a marginal decrease in ClpX N-domain affinity when this residue was changed to alanine, indicating that the side-chain contacts are not critical.

If the C-terminal lysine residues of ^{Ec}SspB and ^{Cc}SspBα are not needed for binding ClpX or for delivery of substrates for degradation, then why have these residues been conserved in adaptors and most of their orthologs? We propose that these terminal amino acids might function to help protect SspB from degradation. SspB is a dimer, and both C-terminal tails normally bind N domains in the ClpX hexamer (Bolon et al., 2004a). However, ^{Ec}SspB tethered via a single XB tail also functions as an adaptor (Bolon et al., 2004a; McGinness et al., 2007), which would potentially allow the second XB peptide of a dimer to be engaged by the translocation pore of ClpX, leading to degradation of that subunit. However, replacing the C-terminal residue of the ssrA degron with lysine makes it an exceptionally poor degradation tag for ClpXP (Barkow, 2009). Thus, having lysine at the C-terminus of the XB peptide should minimize inadvertent ClpXP-mediated degradation. Similar considerations may explain the

strong phylogenetic conservation of D^{157} in SspBα orthologs [see Fig. 2.1(A)]. This acidic amino acid is present in more than 95% of all SspBα XB sequences, but we detected no effects of an alanine substitution either in substrate delivery or in ClpX-binding [see Fig. 2.4(A,B), 2.5(C), 2.7(B), Table 2.1]. Numerous experiments have shown that mutation of residues in degrons to aspartic acid also weakens binding to the translocation pore of ClpX (Flynn et al., 2001; Gottesman et al., 1998). Testing these ideas will require further analysis as the ^{Cc}SspBα variants used in this study are not ideally suited for degradation experiments because they carry N-terminal affinity tags.

Why has recognition of the two peptide-binding motifs, exemplified by the XB peptides of ^{Ec}SspB and ^{Cc}SspBα, been retained by ClpX orthologs that no longer need to interact with the other class of adaptor? The obvious possibility is that these binding sites in the ClpX N domain are maintained because they are also used in recognition of other substrates. For example, an $LREI^{12}$ sequence in *E. coli* UmuD helps mediate ClpXP degradation (Gonzalez et al., 2000; Neher et al., 2003) and is a good match to the β/γ- XB consensus motif. Similarly, a peptide from the λO substrate, which binds the N domain of ClpX, contains an $LLAI^{56}$ sequence (Thibault et al., 2006). Moreover, the C-terminal residues of the phage MuA protein ($LDILEQNRRKKAI^{662}$) target it for ClpXP degradation in a partially N-domain-dependent manner and share homology with the ^{Cc}SspBα XB peptide (Abdelhakim et al., 2008; Levchenko et al., 1997; Wojtyra et al., 2003).

Peptide-binding domains (PDZ, WW, SH2, SH3, PTB, FHA, 14-3-3, EVH1, etc.) are used in modular fashions in an enormous number of biological processes to ensure specificity. In virtually all of these cases, each type of domain has a single binding specificity. Thus, it is somewhat unusual that the N domain of ClpX has at least two peptide-binding specificities. Similarly, the SspB adaptor has more than one binding specificity (Chien et al., 2007a; Flynn et al., 2004; Flynn et al., 2001; Levchenko et al., 2005; Levchenko et al., 2003; Song and Eck, 2003). For example, crystal structures show that the peptide-binding groove of ^{Ec}SspB binds to a peptide sequence in the ssrA tag in one way and binds to a non-homologous recognition sequence in the RseA protein in a completely different fashion. Nevertheless, ssrA and RseA peptides compete for SspB because the binding sites for these peptides overlap (Levchenko et al., 2005). Our competition experiments suggest that the binding sites for the β/γ-XB peptides

and α-XB peptides in the ClpX N domain also overlap, although structural experiments will be needed to confirm this surmise.

ClpXP has hundreds of natural substrates, which are recognized via five classes of degrons (Flynn et al., 2003). Moreover, other AAA+ proteases interact with multiple types of peptide signals to identify the correct substrates (Erbse et al., 2006; Gonzalez et al., 1998; Gottesman et al., 1998; Griffith et al., 2004; Gur and Sauer, 2008; Hoskins et al., 2000; Hoskins and Wickner, 2006; Ishii and Amano, 2001; Wang et al., 2007). The AAA+ p97 protein also employs its N domain to interact with disparate sequences in a wide variety of adaptors (Yeung et al., 2008). The peptide-binding versatility exhibited by the ClpX N domain and SspB ensures that ClpXP can recognize many different substrates and adaptors in different ways but with high specificity. This feature allows ClpXP to carry out quality-control surveillance of a large fraction of the proteome and to participate in numerous regulatory circuits without the need for a single type of degron. Moreover, the ability of ClpXP and other AAA+ proteases to recognize multiple classes of degrons permits the recognition of several weak sequence signals to be coupled via avidity effects. These properties of the system free protein substrates to evolve sequence signals that are both compatible with function and only result in degradation under specific circumstances, such as unfolding, complex dissociation, complex assembly, chemical or proteolytic modification (Baker and Sauer, 2006). Competition of different substrates and/or adaptors for distinct but overlapping binding sites provides an additional level of potential regulation of intracellular proteolysis.

MATERIALS AND METHODS

Buffers

PD buffer contained 25 mM HEPES-KOH (pH 7.6), 5 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.032% NP-40, and 10% glycerol. The ATP-regeneration system contained 5 mM ATP, 50 µg/mL creatine kinase, and 5 mM creatine phosphate. Buffers S1, W20, and W500 contained 50 mM sodium phosphate (pH 8.0), 300 mM NaCl, and imidazole at concentrations of 10 mM, 20 mM, or 500 mM, respectively. For purification of the ^{Cc}ClpX N domain, buffer S1 contained 5 mM imidazole and buffers S1 and W20 were supplemented with 10 mM β-mercaptoethanol

(BME) and 10% glycerol. Buffer A contained 20 mM HEPES-KOH (pH 8.0), 150 mM KCl, 10% glycerol, and 10 mM BME. Buffer S contained 20 mM HEPES-KOH (pH 8.0) and 100 mM KCl.

Protein and Peptide Purification

^{Ec}ClpX, ^{Ec}ClpP, ^{Ec}SspB, ^{Cc}ClpX, ^{Cc}ClpP, ^{Cc}SspBα, and GFP proteins bearing the *E. coli* or *C. crescentus* ssrA tags were purified as described (Chien et al., 2007b; Kim et al., 2000; Levchenko et al., 2000; Levchenko et al., 1997; Yakhnin et al., 1998). GFP-^{Ec}DAS(+4) protein was a gift from J.S. Butler (MIT).

The ^{Cc}SspBα variants with an N-terminal His_6 tag were cloned into a pET28b vector under T7-promotor control and transformed into *E. coli* strain BL21(DE3)/pLysS. The N-terminal His_6-tagged ^{Ec}SspB variants, cloned in pET14b vector, were expressed in BL21(DE3) strains (strains provided by laboratory of RT Sauer). Cells were grown at 37 °C to $OD_{600} \approx 0.5$ in Luria-Bertani broth containing 50 μg/mL kanamycin. Protein expression was induced for 2 h by addition of 0.5 mM isopropyl β-D-thiogalactoside. The culture was harvested by centrifugation, re-suspended in 10 mL of buffer S1 per liter of initial cell culture, and 1 μL/mL protease inhibitor cocktail set III (Novagen, Madison, WI) was added. Cells were frozen in liquid nitrogen, stored, thawed, and lysed by incubating with lysozyme. The lysate was treated with benzonase nuclease (Novagen), cleared by centrifugation for 20 min at 30,000 X *g* at 4 °C, and incubated with nickel-NTA agarose beads (Qiagen, Valencia, CA) equilibrated in S1 buffer for 1 h at 4 °C. The beads were collected by centrifugation, re-suspended, and washed sequentially with buffer S1 and buffer W20. Bound protein was eluted in five fractions using buffer W500. Fractions containing SspB variants were identified by SDS-PAGE, buffer-exchanged into buffer S using PD-10 desalting columns (GE Healthcare, Piscataway, NJ), pooled, and the concentration determined by UV absorption at 280 nm (ε = 13,000 M^{-1} cm^{-1}).

GFP-^{Cc}DAS was constructed using a Gateway cloning system as previously published (Skerker et al., 2005) and the protocol described above used to purify the protein (ε = 55,000 M^{-1} cm^{-1}).

^{Cc}ClpX (residues 1-61) and ^{Ec}ClpX (residues 1-64) N domains with cleavable N-terminal His_6 tags were expressed in *E. coli* strains BLR(DE3) (provided by S. Glynn, MIT) and BL21(DE3)/pLysS respectively using the protocol described for expression of ^{Cc}SspBα variants. Harvested cells were re-suspended in 10 mL S1 buffer plus 10 mM BME and 10% glycerol per liter of initial culture and lysed using a French Press (25,000 psi) at 4 °C. The protocol for

purification of ^{Cc}SspBα variants was then followed up to the wash step. After washing with buffer W20 plus 10 mM BME and 10% glycerol, the nickel-NTA beads were re-suspended in wash buffer, recombinant thrombin (Novagen) was added, and the mixture was incubated overnight at 4 °C to cleave the His_6 tag. The nickel-NTA resin was removed by centrifugation, and the supernatant was chromatographed on a Superdex-75 gel filtration column (GE Healthcare) equilibrated in buffer A. Fractions containing the ClpX N domain were identified by SDS-PAGE, pooled, concentrated using Amicon (MWCO 5k) (Millipore, Billerica, MA) tubes, and the protein concentration was determined by UV absorption at 280 nm.

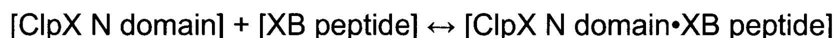
Fluorescein-labeled peptides corresponding to the XB regions of ^{Cc}SspBα (YKIVSLDQFRKK), ^{Ec}SspB (RGGRPALRVVK), and variants containing single alanine substitutions were synthesized by using FMOC techniques on an Apex 396 solid-phase synthesizer.

Protein Degradation Assays

GFP substrates (100 nM) were incubated with ^{Ec}ClpXP (50 nM $^{Ec}ClpX_6$; 100 nM $^{Ec}ClpP_{14}$) or ^{Cc}ClpXP (200 nM $^{Cc}ClpX_6$; 400 nM $^{Cc}ClpP_{14}$) in the presence or absence of adaptor (300 nM monomer) at 30 °C in PD buffer plus an ATP-regeneration system (Flynn et al., 2001). Degradation was monitored by decreased fluorescence (excitation 488 nm; emission 511 nm) using a Photon Technology International fluorimeter (Birmingham, NJ). The rates of reaction were determined by the slopes of linear fits to the decrease in fluorescence within the first 10-30 seconds of reaction. The error bars indicate the standard deviation of three or more independent measurements.

Peptide-binding Assay

Fluorescein-labeled ^{Cc}SspBα XB peptides (60 nM) were incubated with increasing amounts of ^{Cc}ClpX N domain in buffer A at 30 °C, and fluorescence anisotropy was measured using a Photon Technology International fluorimeter (excitation 490 nm; emission 515 nm). The binding of fluorescent ^{Ec}SspB XB peptides to the ^{Ec}ClpX N domain was assayed in the same way. The K_D values from individual experiments were determined by fitting binding data to the quadratic equation determined from the following equilibrium:


[ClpX N domain] + [XB peptide] ↔ [ClpX N domain•XB peptide]

The binding equation used was $y = a + ((b-a)*((d+x+c) - (SQRT((d+x+c)^2 - 4dx)/2d)))$ where y (y-axis) = anisotropy, x (x-axis) = [ClpX N domain], a = anisotropy of free peptide, b = anisotropy when all peptide is bound to the N domain, c= K_D, d = [total peptide].

ACKNOWLEDGMENTS

We thank Igor Levchenko, Jim Butler, Steve Glynn and the Sauer lab for sharing reagents. We also thank Jennifer Hou for comments on the manuscript and the Baker lab for helpful discussions. This work was supported by NIH grants GM49224 and AI16892. P.C. was supported by NIH Grant 5K99GM084157. T.A.B. is an employee of the Howard Hughes Medical Institute.

CHAPTER THREE

Substrate-Selection by the ClpXP Protease: Future Directions

Much progress has been made in understanding the intracellular proteolysis machinery in prokaryotes over the past few years. The balance between degrading a broad spectrum of substrates and retaining a high level of specificity is vital to the proper functioning of the proteolytic machinery. The different strategies used by intracellular proteases to maintain this balance have been the topic of numerous studies. However, significant aspects of the complex process of substrate selection are still unknown. The goal of this work has been to better understand the unusual abilities of two players in this system, the adaptor protein SspB and the N domain of the AAA+ protein ClpX, to interact with a wide range of proteins. This chapter discusses some of the still unanswered questions about the two proteins and how an extension of some of this work can contribute to our understanding of the roles of SspB and ClpX in the cell.

The SspB Adaptor Protein

In *E. coli*, SspB functions as a canonical adaptor by delivering substrates to ClpXP, thereby enhancing degradation rates. Although there have been many studies characterizing its structure and mechanism of action, several details about SspB remain unknown. First, it is still unclear how SspB selects proteins to interact with. The two known groups of SspB-modulated ClpXP substrates, ssrA-tagged polypeptides and NRseA, do not share a common or similar peptide-binding motif. Despite the dissimilarity, both substrates bind SspB in overlapping binding sites (Flynn et al., 2001; Levchenko et al., 2005; Levchenko et al., 2003; Song and Eck, 2003). This lack of a consensus sequence for SspB-binding indicates the versatility of the adaptor-substrate interaction and leaves open the possibility that SspB is able to recognize additional ClpXP substrates. Indeed, there is experimental evidence pointing in that direction. Both previous work and the proteomic screen in this work (see Appendix) have provided a number of candidate substrates, which may be directly interacting with SspB. However, experiments will have to be conducted to validate that these candidates are true SspB-modulated substrates. *In vitro* degradation assays using purified proteins will be useful to test whether they are ClpXP substrates and if they interact with SspB. However, as discussed in chapter 1, several ClpX substrates contain latent degrons, which are exposed only after initial substrate-processing by other proteases. Therefore, lack of *in vitro* degradation does not necessarily imply that a protein is not a *bona fide* ClpXP substrate. Under these circumstances,

monitoring *in vivo* levels of the endogenous proteins over time in strains with and without ClpXP or SspB will likely give more conclusive results.

By identifying ClpXP substrates that interact with SspB (if more direct substrates in fact exist), it will be possible to get a better understanding of how SspB selects substrates. Analyzing properties of SspB-modulated substrates could help identify determinants for SspB-binding as well as provide clues about why these particular substrates associate with adaptors whereas others do not. By extension, we will also get insight into the roles of adaptor proteins in general.

Degradation of several substrates appears to be inhibited by SspB, as they were trapped more in the absence of SspB ((Flynn et al., 2004); Appendix). The trapping experiment using an *sspB* over-expression strain would be a good complement to the proteomic screen in this work. A possibility is that SspB may be directly interacting with substrates to mask degrons and consequently inhibiting degradation. More indirectly, SspB may also be competing with substrates or adaptors to bind ClpXP. Indeed, there are various lines of evidence that SspB can compete with ClpX N domain-binding proteins such as Dps and RssB (Meyer et al., *in prep*). Analyses of substrates that are stabilized in the presence of SspB will also be able to provide important information about novel adaptor proteins.

Despite studies on the mechanism and structure of SspB, surprisingly little is known about regulation and biological functions of SspB in the cell. SspB (stringent starvation protein B) is part of the *ssp* operon, which also codes for an RNA-polymerase associated protein called SspA. SspA expression is induced during stationary phase and under stress conditions such as acid stress or amino acid starvation (Hansen et al., 2005; Williams et al., 1994). Genes involved in acid tolerance are under the regulation of the global regulator H-NS, which inhibits expression under normal growth conditions. SspA down-regulates H-NS levels during acid stress, thereby relieving the repression and enabling the cell to cope with the low surrounding pH (Hansen et al., 2005). Unlike levels of SspA, SspB levels appear unchanged in both exponential and stationary phases (Farrell et al., 2005). However, it is possible that SspB levels change under different stress conditions. For instance, during amino acid starvation conditions, SspB activation would increase the degradation rates of ssrA-tagged polypeptides and free amino acids, which could then be used to synthesize stress-response proteins. Understanding how SspB is regulated will provide insight into the physiological importance of the adaptor.

Recent work has identified proteins termed anti-adaptors, which can stabilize substrates. Multiple anti-adaptors have been shown to inhibit RssB interaction with the stationary phase

sigma factor σ^S, leading to increased σ^S stability and induction of σ^S-induced genes (Bougdour et al., 2008; Bougdour et al., 2006). It would be interesting to see if there are similar inhibitors specific to SspB. Pulling out SspB-interacting proteins from cell lysate by using an affinity column of SspB-coated beads would be useful for identifying, not only novel ClpX substrates that interact with SspB, but also general SspB-binding proteins including anti-adaptors.

The N-terminal domain of the AAA+ Unfoldase ClpX

Similar to SspB, the peptide recognition by the ClpX N domain is very flexible. As shown in this work (see chapter 2), cross-species interactions between ClpX and SspB in *E. coli* and *C. crescentus,* despite limited sequence identity, demonstrate the variation in length and sequence of N domain-interacting regions. To gain more insight into this flexibility in peptide recognition, co-crystal structures of *E. coli* and *C. crescentus* ClpX N domains with the ^{Cc}SspBα XB identified in this work would be very helpful to learn more details about the interaction. Because ^{Cc}XB and ^{Ec}XB are able to compete for binding the N domain, it is highly likely that they bind overlapping sites on the N domain. Learning more about the detailed interactions in that region will provide more insight into how exactly ClpX N domain selects substrates and adaptors. Co-crystal structures of the N domain with peptides corresponding to other N domain-specific substrates, such as Dps and λO, could also be useful. Because of the sequence and length diversity in N domain-binding regions, details of the peptide-N domain interactions will contribute towards understanding the general mechanism of protein recognition by the ClpX N domain.

Both N domain and the AAA+ domain of ClpX are involved in substrate recognition. They have distinct substrate preference, which is important for the ability of ClpXP to degrade a wide spectrum of substrates. To investigate the contribution of each domain to substrate selection, one approach would be to express the $ClpP^{trap}$ in a strain expressing a variant of ClpX without the N domain. Trapped proteins in a strain expressing the full-length ClpX could be compared to the substrates trapped in a strain expressing $ClpX^{\Delta N}$. Because all the known ClpX adaptors interact with the N domain, this experiment can also provide information about new adaptor proteins. For instance, if a particular substrate is trapped more in the strain expressing full-length ClpX, an adaptor may be involved in the degradation process.

The work in this thesis has been to better understand the process of substrate selection by intracellular proteases. The N domain of the AAA+ unfoldase ClpX has been shown to have a complex mode of interacting with substrates and adaptors of the ClpXP protease. Specificity of the N domain is not surprising, given the importance of ensuring minimal indiscriminate protein degradation. However, interestingly, the N domain is also considerably versatile in peptide recognition, imposing limited restrictions on sequences of interacting-proteins and thereby widening the range of substrates degraded by the protease. Furthermore, the ClpX-specific adaptor protein SspB also has a similar trait of being both specific and flexible in substrate-binding. Several potential SspB-modulated ClpXP substrates have been provided by a proteomic screen (see Appendix), analyses of which may be able to contribute to a greater understanding of the biological role of SspB and perhaps of adaptors in general.

APPENDIX I

Proteomic Screen to Elucidate the Effect of SspB on ClpX Substrate Profile in *E. coli*

This work was done in collaboration with Judit Villén (laboratory of Steven Gygi, Harvard Medical School). J.V. conducted the mass spectrometry and peptide analysis using SILAC.

INTRODUCTION

Proteolysis by intracellular bacterial proteases is important for protein quality control and regulation of many cellular responses. Energy-dependent proteases in bacteria include ClpXP, ClpAP, HslUV, Lon, and FtsH. These proteases consist of two separate components, one ATPase subunit belonging to the AAA+ (ATPases associated with various cellular activities) super-family and one proteolytic subunit (Gottesman, 2003). Stacking of the two ring-shaped subunits forms a barrel-like structure with the active peptidase sites sequestered within the internal chamber. The ATPase component confers substrate-specificity to the proteases, a very important trait as indiscriminate protein degradation can have severe detrimental effects on the cell. After recognizing substrates, the ATPase unfolds and translocates them into the interior of the proteolytic component where the unfolded polypeptides are cleaved into short peptides (Gottesman, 2003).

ClpXP is one of the best-characterized proteases in *E. coli*. The ATPase component ClpX forms a hexameric ring and stacks with the tetradecameric peptidase ClpP, resulting in the active ClpXP protease. ClpP can also associate with another ATPase ClpA to form a different protease ClpAP. Although ClpXP and ClpAP can degrade some of the same substrates, they have distinct substrate-preferences (Gottesman et al., 1993; Grimaud et al., 1998; Katayama et al., 1988; Wojtkowiak et al., 1993).

Proteases are usually able to recognize substrates by degradation signals (degrons) or tags, which are amino acid sequences located mostly near the N- or C-terminal regions of the substrates. Although these tags can often be directly recognized, additional proteins called adaptors are also able to affect substrate-choice. These protease-specific and substrate-specific adaptors often bind both protease and substrates to facilitate degradation. Certain adaptors such as RssB and UmuD are necessary for degradation of substrates (σ^S and UmuD' respectively) whereas others like SspB enhance degradation rates of substrates that can be degraded by ClpXP even in the absence of the adaptor (Baker and Sauer, 2006).

The SspB adaptor protein has been shown to bind two groups of ClpXP-substrates and, by tethering them to the protease, increase the local substrate concentration resulting in a substantial increase in degradation rate. SspB-modulated substrates include the ssrA-tagged polypeptides and the cytoplasmic domain of the anti-sigma factor RseA (NRseA) (Flynn et al., 2004; Levchenko et al., 2000). The ssrA-tag is an 11-residue C-terminal peptide that is appended onto incomplete polypeptides when ribosomes stall during translation. ClpX

recognizes this tag and results in destruction of these defective polypeptides (Karzai et al., 2000). The second SspB-facilitated substrate is a domain of RseA, a transmembrane protein involved in extracytoplasmic stress response. Under non-stress conditions, the N-terminal cytoplasmic domain of RseA (NRseA) remains bound to the sigma factor σ^E to keep it inactive. However, when there is cell envelope stress, such as unfolded proteins in the periplasm, RseA is processed by two membrane-bound proteases to release the NRseA-σ^E complex. ClpXP is now able to recognize the C-terminal residues of NRseA and degrade it, allowing σ^E to turn on stress-response genes. NRseA is also recognized by SspB, which delivers it to the ClpXP protease (Flynn et al., 2004).

Similar to proteases, adaptor proteins recognize specific substrates through recognition tags. Interestingly, the SspB-interacting regions (underlined) of the ssrA-tag (^{1}AANDENYALAA^{11}) and NRseA (^{77}EAQPAPHQWQKMPFW^{91}) share homology in neither length nor sequence (Flynn et al., 2004; Flynn et al., 2001; Levchenko et al., 2005; Levchenko et al., 2003; Song and Eck, 2003). Therefore, it is difficult to define an SspB-binding motif based on these two regions and leaves open the possibility that there are additional yet-unidentified SspB-modulated ClpXP substrates. Indeed, there is evidence that points towards that direction. Flynn et al. (2004) took advantage of an inactive variant of ClpP known as the ClpP^{trap}, which has an Ala-substitution of the active site Ser 97 to investigate the effect of SspB on substrate selection by ClpXP. ClpX can form a complex with ClpP^{trap}, unfolding and translocating substrates into the ClpP^{trap}-pore. However, once inside the pore, the unfolded substrates cannot be degraded by the inactive ClpP^{trap} and remain trapped. By expressing this ClpP^{trap} in cells with and without SspB, Flynn et al. (2004) compared the "trapped" substrates in the two strains by two-dimensional gel electrophoresis and mass spectrometry (Fig A1.1). NRseA was trapped more in the strain expressing SspB and subsequent experiments showed that it directly interacted with the adaptor during proteolysis. Interestingly, there were additional proteins that were differentially trapped in the two strains, implying that SspB may have a global effect on substrate selection by ClpXP and raising the possibility that some of these proteins may directly interact with the adaptor in a manner similar to NRseA and ssrA-tagged substrates.

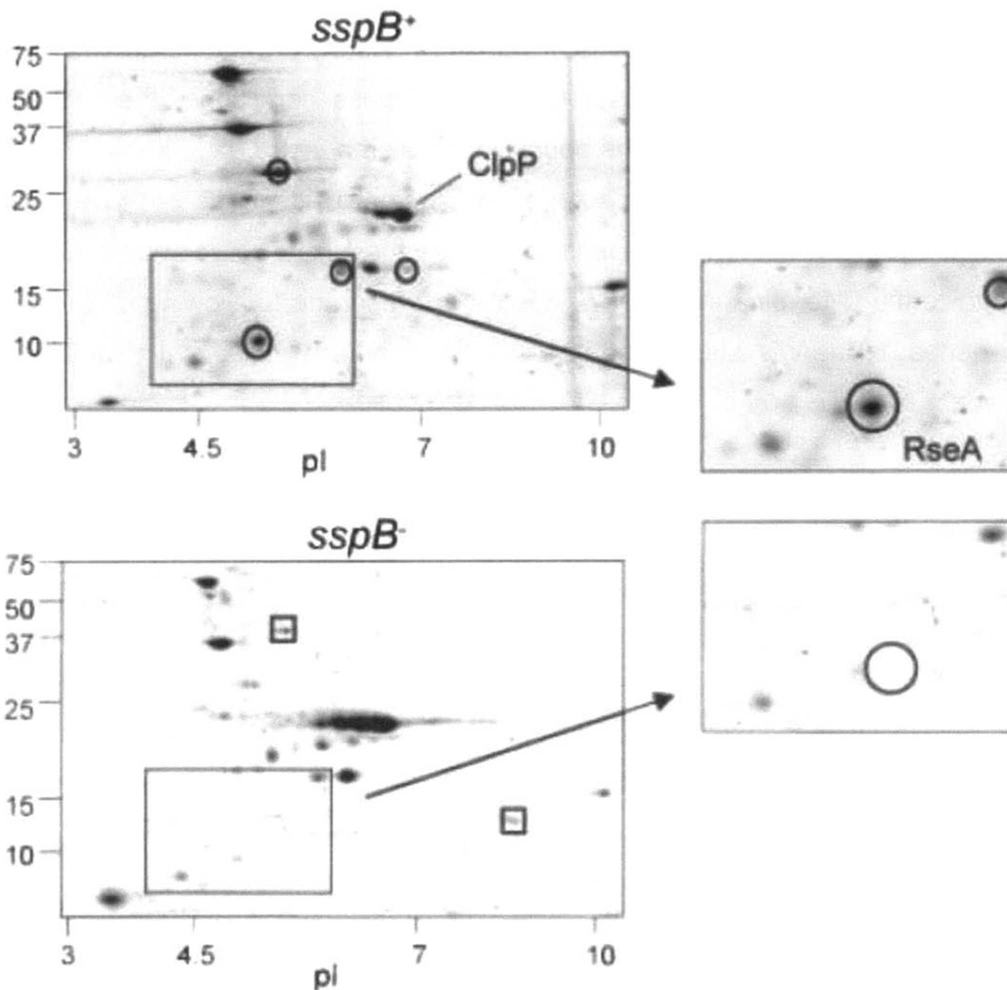


Figure A1.1. Substrates associated with ClpXP^{trap} in *sspB*$^{+}$ and *sspB*$^{-}$ strains (Flynn et al., 2004). These substrates were analyzed by two-dimensional gel electrophoresis to compare the ClpX substrate profile in the two strains. Substrates that were preferentially trapped in the presence of SspB are circled and substrates that were preferentially trapped in the absence of SspB are enclosed by squares (Figure taken from (Flynn et al., 2004)).

Why do certain ClpXP substrates have SspB as an adaptor whereas others do not? One possibility is that SspB is used to co-regulate groups of proteins by modulating their degradation by ClpXP (Flynn et al., 2004). Another option could be that cells use SspB to enhance degradation of substrates of higher priority. Because ClpX has multiple substrates, it is highly likely that there is competition amongst substrates for binding the ATPase. Having SspB as an adaptor can give certain proteins priority over others when competing for ClpX. As a result, SspB-binding proteins might be predicted to be degraded much faster than other substrates

under specific conditions. By profiling ClpXP substrates that are modulated by SspB, it will be possible to explore these hypotheses and gain insight into exact roles of SspB and other adaptors in cells.

This study involves a quantitative analysis of the ClpXP-substrate profiles in the presence and absence of SspB. We used an approach called stable isotope labeling by amino acids in cell culture (SILAC) to identify $ClpXP^{trap}$-substrates in *sspB*$^{+}$ and *sspB*$^{-}$ cells (Ong et al., 2002). This quantitative mass spectrometry technique was successfully used by Neher et al. (2006) to probe the effect of DNA damage on ClpXP substrate profile. The substrate profile of ClpXP in cells grown under regular conditions was compared to the ClpXP substrate profile in cells treated with a DNA-damaging agent. The two sets of cells were labeled with two different isotopes of leucine and $ClpP^{trap}$ (and associated substrates) was purified from an equal mixture of the two samples. An analysis of the SILAC ratio (heavy-to-light peptides) of the trapped substrates allowed a direct comparison between the $ClpXP^{trap}$-substrates in the two sets of cells.

Here, we expressed the inactive $ClpP^{trap}$ in two different strains. The strain expressing *sspB* was grown in defined media containing ^{13}C-Leu whereas the *sspB*$^{-}$ strain was grown in media supplemented with regular (^{12}C-)Leu. The $ClpP^{trap}$ protein from the two strains was purified and associated proteins identified by tandem mass spectrometry. The SILAC ratio was calculated for each of the trapped substrates. A high SILAC ratio [H/L] indicated that the substrate was trapped more in the presence of SspB whereas a low [H/L] ratio indicated the opposite (Fig A1.2). Any potential SspB-modulated ClpXP substrates would be in the first group as they were preferentially recognized by ClpXP in the presence of SspB. Substrates that directly competed with SspB to interact with ClpX, e.g. the DNA-binding stress protein Dps, would be in the second group. Thus, this approach can make it possible not only to find novel SspB-interacting substrates but also identify substrates that may be indirectly affected by the adaptor.

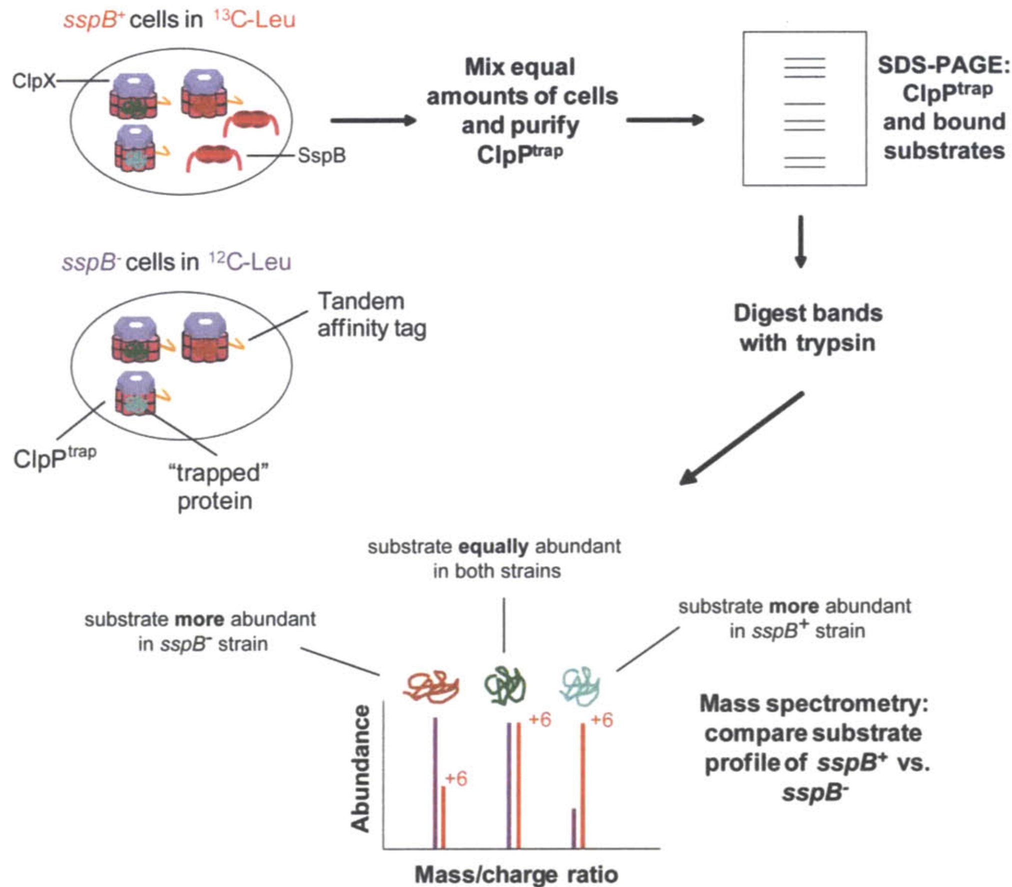


Figure A1.2. Schematic depiction of the different steps in analysis by SILAC. The ClpXP-trapped substrates in the SspB-expressing strain were labeled with a heavy isotope of leucine and the ones in the SspB-knock-out strain were labeled with the regular leucine isotope. The affinity-tagged $ClpP^{trap}$ from both strains was purified and the associated proteins were analyzed by mass spectrometry. The peptides corresponding to the cyan substrate were more abundant in the presence of SspB whereas the ones corresponding to the brown substrate were more abundant in the absence of SspB. Peptides of the green substrate were present in equal amounts in the two strains.

RESULTS AND DISCUSSION

I. Trapping Experiment

The adaptor protein SspB has been shown to have both direct and indirect effects on the degradation of certain ClpX substrates. To analyze these effects, a variant $ClpP^{trap}$ was used to compare the ClpX-substrate profile in cells with and without SspB. For ease of purification, an affinity-tagged version of the $ClpP^{trap}$ was used in this assay (Fig A1.2). Both strains were *clpP*⁻,

clpA⁻, smpB⁻, and *leuB⁻*. Additionally, one strain was *sspB⁻* whereas the other had an intact *sspB* gene. Knocking out the endogenous *clpP* ensured that all *bona fide* ClpXP substrates were associated with the $ClpP^{trap}$ and not degraded by any wild-type ClpP in the cells. ClpA was removed resulting in all ClpP-trapped substrates being ClpX-specific. SmpB is a highly conserved RNA-binding protein, which is necessary for the ssrA-tagging system (Karzai et al., 1999). Because ssrA-tagged proteins constitute a large group of SspB-modulated ClpXP substrates, a knock-out of *smpB* prevented any ssrA-tagged polypeptides from being trapped, thereby allowing other substrates to associate with the $ClpXP^{trap}$ protease. The LeuB protein is involved in leucine biosynthesis and therefore when the gene was knocked out, the cells were unable to synthesize the amino acid. Therefore, all incorporated leucine residues in the cell were taken up from the media, ensuring maximal protein labeling.

After purification of the $ClpP^{trap}$, the associated proteins were analyzed by mass spectrometry. The SILAC ratio [H/L] of peptides corresponding to these associated proteins was used to determine relative amounts trapped in the two strains. Differential labeling of the peptides allowed an internal control making it possible to do this quantitative comparison. Table A1.1 shows a list of all the associated proteins in the two strains and their ratios.

Table A1.1: "Trapped" ClpX Substrates with and without SspB.¥

Gene	NCBI GI	# Unique Peptide‡	SILAC Ratio [H/L]*	Description
gmd	16129993	3	0.03307472	GDP-D-mannose 4,6-dehydratase
rplJ	16131815	7	0.062308842	50S ribosomal subunit protein L10
dps	16128780	10	0.136379646	stress response DNA-binding protein
lpp	16129633	2	0.150200339	murein lipoprotein structural gene
lacZ	16128329	13	0.247449506	β-D-Galactosidase
tufA	16131218	2	0.255487439	duplicate gene for EF-Tu subunit
ompA	16128924	6	0.285588453	outer membrane protein 3a
ydgA	16129572	2	0.295370236	conserved protein
dksA	16128138	6	0.295485139	DNA-binding transcriptional regulator
slyD	16131228	2	0.297401635	FKBP-type peptidyl prolyl cis-trans isomerase
yciW	90111242	5	0.30071026	predicted oxidoreductase
pflB	16128870	8	0.307824472	pyruvate formate lyase
arnA	16130190	13	0.313155106	UDP-GlcA C-4'-decarboxylase
rpoZ	16131520	2	0.366446456	RNA polymerase, ω subunit
crp	16131236	7	0.369557748	DNA-binding transcriptional dual

				regulator
lpdA	16128109	9	0.371386538	dihydrolipoamide dehydrogenase
atpA	16131602	4	0.373309362	alpha subunit membrane-bound ATP synthase
ydjA	16129719	2	0.373547026	predicted oxidoreductase
icdA	16129099	2	0.384845829	isocitrate dehydrogenase
atpD	16131600	10	0.388688056	membrane-bound ATP synthase, β subunit
glgA	16131303	12	0.407811846	glycogen synthase
tsf	16128163	3	0.408440912	elongation factor
glmS	16131597	24	0.411484792	L-glutamine:D-fructose-6-phosphate aminotransferase
hisB	90111373	8	0.415150532	fused histidinol-phosphatase/imidazoleglycerol-phosphate dehydratase
rpsA	16128878	11	0.422182848	30S ribosomal subunit protein S1
yjaE	16131825	2	0.432554221	anti-RNA polymerase sigma 70 factor
crl	16128226	2	0.438954048	sigma factor-binding protein
cysN	16130658	12	0.451743346	sulfate adenylyltransferase subunit 1
htpG	16128457	4	0.452704582	molecular chaperone Hsp90 family
hfq	16131994	2	0.459707848	RNA-binding protein
ydjN	16129683	2	0.460814169	predicted transporter
rpoB	16131817	22	0.465248356	RNA polymerase, β subunit
yggB	16130825	2	0.47002162	mechanosensitive channel
glnD	16128160	6	0.47878349	uridylyltransferase
sthA	90111670	2	0.481921485	pyridine nucleotide transhydrogenase
rplB	16131196	5	0.488606409	50S ribosomal protein L2
rpsL	16131221	2	0.492806312	30S ribosomal subunit protein S12
mreB	90111564	4	0.497706149	actin-like component of cell wall structural complex MreBCD
carB	16128027	27	0.498530681	carbamoyl phosphate synthase large subunit
ftsZ	16128088	13	0.50024535	cell division protein
mukB	16128891	3	0.508638869	ATPase and DNA-binding subunit of chromosome condensin MukBEF
cysK	16130340	3	0.512612387	cysteine synthase A
cysC	16130657	3	0.51500961	adenylylsulfate kinase
prsA	16129170	6	0.523058227	ribose-phosphate pyrophosphokinase
rpoC	16131818	5	0.524155118	DNA-directed RNA polymerase subunit β'
trkA	16131169	7	0.524186946	NAD-binding component of TrK potassium transporter
yleA	16128644	4	0.531161859	2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol oxygenase
bioB	16128743	2	0.537304585	biotin synthase
mopA	16131968	14	0.538111355	Cpn60 chaperonin GroEL, large subunit of GroESL

ftsA	16128087	4	0.561163653	ATP-binding cell division protein involved in recruitment of FtsK to Z ring
dnaK	16128008	47	0.567455791	chaperone Hsp70, co-chaperone with DnaJ
fnr	16129295	3	0.57922401	DNA-binding transcriptional dual regulator
yhdM	16131171	3	0.580850588	zntR; DNA-binding activator in response to Zn(II)
cyaA	16131658	13	0.584152043	adenylate cyclase
ybaQ	90111140	5	0.58587175	predicted DNA-binding transcriptional regulator
rplL	16131816	3	0.595151008	50S ribosomal subunit protein L7/L12
ycbW	90111193	2	0.597165235	predicted protein
def	16131166	2	0.60003943	peptide deformylase
infB	16131060	4	0.618479744	fused protein chain initiation factor 2
priA	16131773	5	0.630008516	Primosome factor n' (replication factor Y)
exbB	16130904	5	0.643756921	membrane spanning protein in TonB-ExbB-ExbD complex
talB	16128002	2	0.645098731	transaldolase B
pnp	49176320	6	0.647908465	polynucleotide phosphorylase/polyadenylase
rpoS	16130648	17	0.667068109	RNA polymerase, sigma S (sigma 38) factor
yebM	16129811	4	0.668445383	znuC; zinc transporter subunit: ATP-binding component of ABC superfamily
tpx	16129285	4	0.685171835	lipid hydroperoxide peroxidase
lexA	16131869	3	0.686666104	DNA-binding transcriptional repressor of SOS regulon
rne	16129047	14	0.688418677	fused ribonucleaseE: endoribonuclease/RNA-binding protein/RNA degradosome binding protein
iscU	16130454	2	0.694653655	iron-sulfur cluster assembly scaffold protein
ribB	16130937	4	0.704497687	3,4-dihydroxy-2-butanone-4-phosphate synthase
gnd	16129970	3	0.742868242	6-phosphogluconate dehydrogenase
nuoG	49176206	5	0.829906167	NADH:ubiquinone oxidoreductase, chain G
nadB	16130499	10	0.844571937	quinolinate synthase
cysA	16130348	9	0.87368292	sulfate/thiosulfate transporter subunit
rpsJ	16131200	4	0.895161755	30S ribosomal protein S10
sucA	16128701	6	0.895430889	2-oxoglutarate dehydrogenase E1 component
ptsI	16130342	8	0.912767986	phosphoenolpyruvate-protein phosphotransferase
pepB	90111453	2	0.959706497	aminopeptidase B
acnB	16128111	29	0.968978144	bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase

typA	49176434	2	0.970655092	GTP-binding protein
lepA	16130494	2	0.983569602	GTP-binding protein
hflB	16131068	4	1.025161703	ATP-dependent metalloprotease
ppx	16130427	4	1.065069406	exopolyphosphatase
nusA	16131061	4	1.144782961	transcription termination/anti-termination L factor
sixA	16130273	2	1.199472353	phosphohistidine phosphatase
nuoC	16130221	2	1.241774469	bifunctional NADH:ubiquinone oxidoreductase subunit C/D
deaD	90111550	5	1.405830318	ATP-dependent RNA helicase
pgk	16130827	3	1.500470386	phosphoglycerate kinase
clpB	16130513	9	1.500986032	protein disaggregation chaperone
ygjD	16130960	6	1.536278459	putative DNA-binding protein
gyrA	16130166	8	1.579162812	DNA gyrase (type II topoisomerase), subunit A
hscA	16130451	6	1.579644686	DnaK-like molecular chaperone
purA	16131999	2	1.592173157	adenylosuccinate synthetase
recN	49176247	13	1.651572683	recombination and repair protein
nuoB	16130222	2	1.652159808	NADH:ubiquinone oxidoreductase, chain B
hsdR	16132171	5	1.717417207	endonuclease R Type I restriction enzyme
gltB	16131102	45	1.776028408	glutamate synthase, large subunit
ygfZ	16130800	2	2.603434256	Hda suppressor
atpF	16131604	2	2.608348472	F0 sector of membrane-bound ATP synthase, subunit b
eno	16130686	2	3.218067447	enolase
yfiD	16130504	4	3.46501676	autonomous glycyl radical cofactor
yghJ	49176293	2	4.022076732	predicted inner membrane lipoprotein
oppD	49176090	2	4.114054838	oligopeptide transporter subunit
rseA	16130497	4	6.856055241	anti-sigma factor E

¥ The substrates with < 2 peptides have been removed from the list. Potential SspB-interacting substrates are shown in bold (with SILAC ratios > 2).
‡This column shows the number of peptides for which both heavy and light species were detected.
*This column shows the SILAC ratio, which is the average ratio of heavy-to-light peptides after normalizing to amount of $ClpP^{trap}$ in the two strains.

II. Validation of Results

Despite certain caveats (discussed below), the results of the preliminary run look promising. Many of the trapped proteins are true ClpX substrates, such as Dps and RseA (Flynn et al., 2004; Stephani et al., 2003). Indeed, Dps is degraded by ClpXP in a ClpX N domain-dependent

manner and therefore competes with SspB for binding the N domain (Meyer et al., *in prep*). Therefore, it was expected that Dps would be trapped more in the absence of SspB, which was indeed the case in our study with the Dps SILAC ratio being considerably low (0.13) indicating over-representation in the *sspB⁻* strain (Table A1.1). The adaptor RssB-mediated σ^S degradation is also N domain-dependent leading to higher levels of σ^S trapped in the *sspB⁻* strain. Similarly, the very high SILAC ratio of RseA (6.86) shows that the anti-sigma factor is trapped more in the *sspB⁺* strain, as would be expected from an SspB-interacting substrate.

The proteins that were preferentially trapped in the presence of SspB could be potential SspB-modulated substrates. Similar to NRseA and ssrA-tagged substrates, SspB may directly interact with these substrates and deliver them to ClpXP for degradation. Using a SILAC ratio > 2 as cut-off, candidates for this group of substrates include OppD, YghJ, YfiD, Eno, AtpF, and YgfZ (Table A1.1). However, because of general over-representation of membrane-bound proteins in trapping experiments, Eno, YfiD, and YgfZ are more likely to be ClpX substrates. To validate this result, the *in vivo* degradation of these substrates would have to be tested in wild-type, *clpX⁻*, and *sspB⁻* strains to probe possible roles of ClpX and SspB. The three candidates can also be purified and tested for *in vitro* degradation mediated by ClpX and SspB. In fact, affinity-tagged YgfZ was purified and its degradation tested *in vitro*. Because YgfZ is a putative folate-binding protein (Teplyakov et al., 2004), degradation was also tested in the presence of folate. However, there was no observable degradation under the conditions tested here (Fig. A1.3). It is possible that additional factors that were not present in the *in vitro* experiment may be required for YgfZ degradation. Another possibility is that YgfZ gets degraded only under specific conditions which were not mimicked in this assay. Investigating *in vivo* degradation of YgfZ may be able to address these concerns and provide more conclusive results.

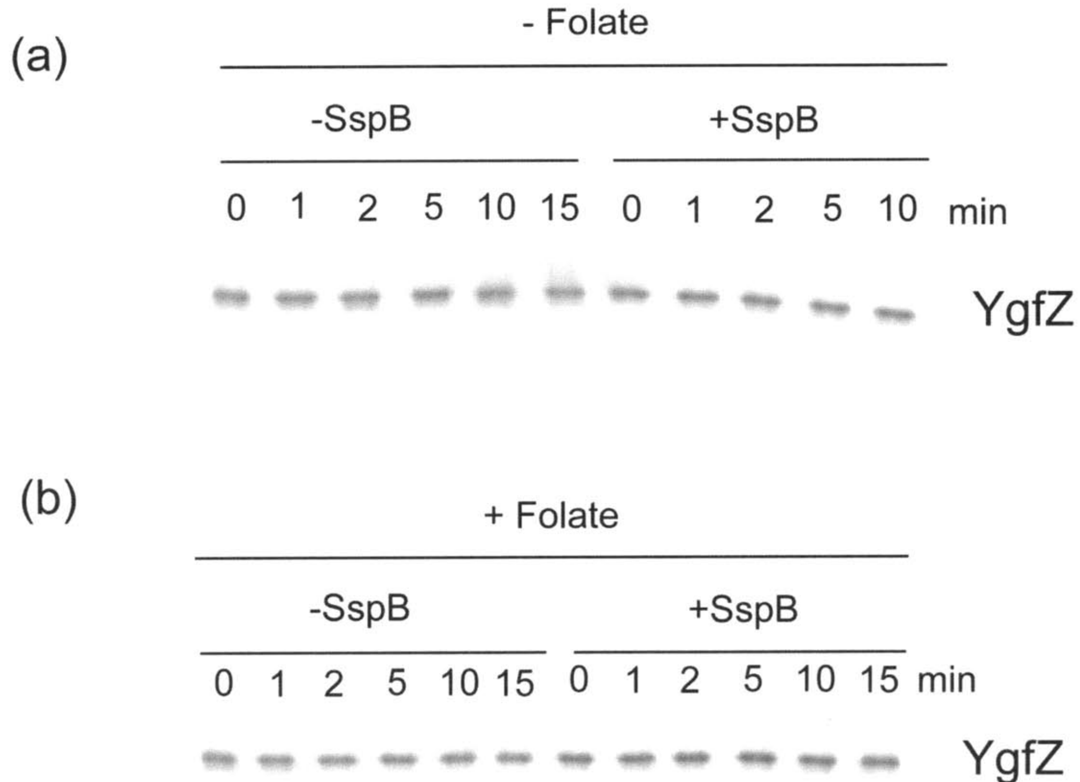


Figure A1.3. *In vitro* degradation of YgfZ by ClpXP. Western blots showing degradation of YgfZ with and without SspB in the (a) absence of folate and (b) presence of folate.

III. Caveats

There were certain caveats to the trapping experiment, which must be considered when analyzing the results of this work. The experiment was done once and should be repeated before starting the intensive process of testing degradation of individual substrates trapped in the assay. There were certain technical difficulties when doing this experiment. Interestingly, the *sspB*$^{-}$ strain had a slightly faster growth rate compared to the *sspB*$^{+}$ strain. This result was consistently observed in multiple growth experiments. When harvesting cells expressing the ClpP^{trap}, the two strains were grown to the same OD_{600} followed by a 2-hour-induction of protein expression before mixing the two strains. Because of the different growth rates, mixing of equal amounts of cells of the two strains may not have been very accurate. This problem was

reflected in the different amounts of ClpP^{trap} purified from the two strains (look below for explanation). However, to compensate for the different ClpP^{trap} levels in the two strains, the SILAC ratios reported here were normalized to the [H/L] ratio of peptides corresponding to the ClpP^{trap}.

In addition, the yield of the trapping experiment was modest and it may be ideal to scale up the experiment to ensure maximal trapping. The myc$_3$-TEV-His$_6$-ClpP^{trap} did not bind well to the anti-myc beads and therefore there was considerable loss of protein at that purification step (Materials and Methods section). The inadequate binding to anti-myc beads may be because of oversaturation of the beads. A similar problem was observed in the nickel-NTA-binding step used to remove ClpP^{trap}. During the purification of ClpP^{trap}-associated proteins, there was a step to remove the ClpP^{trap} but the substantial amount of ClpP^{trap} still present in the final sample indicate that all the His$_6$-ClpP^{trap} did not bind to the nickel-NTA resin. It may be better to use more beads when repeating the experiment to eliminate the ClpP^{trap}, although not removing the trap provides the added advantage of having an internal control to normalize the SILAC ratios, as was done in this work.

The proteomic screen also leaves open the possibility of false positives resulting from non-specific interactions of proteins with the ClpP^{trap}. Abundant proteins are especially likely to be trapped even if they are not ClpXP substrates. To remove these false positives from the list of potential ClpX substrates, the ClpP^{trap} can be expressed in a *clpX*$^-$ *clpA*$^-$ strain, which should trap any non-specific binders.

MATERIALS AND METHODS

Trapping Strains

For the proteomic screen, mutations (*clpA::kan* and *clpP::cat* for both *sspB*$^+$ and *sspB*$^-$ strains; *sspB::kan* for the *sspB*$^-$ strain) were introduced into the starting strain *E. coli* W3110 *ΔsmpB ΔleuB* (Neher et al., 2006) using P1 transductions to generate the two strains: W3110 *ΔsmpB ΔleuB clpA::kan clp::cat* (TC57) and W3110 *ΔsmpB ΔleuB ΔsspB clpA::kan clp::cat* (TC58). The plasmid pJF105 expressing myc$_3$-TEV-His$_6$-ClpP^{trap} (Flynn et al., 2003) was then transformed into the TC56 and TC57 to form the ClpP^{trap} over-expression strains TC58 (*sspB*$^+$) and TC59 (*sspB*$^-$) respectively.

Protein Purification

The strains TC58 and TC59 were grown in defined media as described in Neher et al. (2006). TC58 was grown in media supplemented with the heavy isotope of leucine (Cambridge Isotopes) whereas TC59 was grown in media with the regular leucine. The $ClpP^{trap}$ was over-expressed for 2 hours at 30°C and equal amounts of cells were mixed. The cells were resuspended in S1 buffer (50 mM sodium phosphate buffer (pH 8), 300 mM NaCl, 10% glycerol, 5 mM imidazole) followed by lysis using a French Press (25,000 psi). Protease inhibitor cocktail III (Calbiochem) was added and the lysate spun down to remove cell debris. The supernatant was incubated with nickel-NTA agarose beads (Qiagen) for 2 hours at 4°C. The beads were washed with S1 buffer followed by W20 buffer (S1 buffer with 20 mM imidazole). Fractions were eluted with W500 (S1 buffer with 500 mM imidazole) and the concentrated fractions buffer-exchanged into PBS (150 mM NaCl, 20 mM sodium phosphate (pH 7.3)) using PD-10 desalting columns (GE Healthcare). The eluate was incubated for 2 hours at 4°C with anti-myc agarose beads (Santa Cruz Biotechnology), pre-equilibrated in PBS. The beads were washed sequentially with PBS, PBS-T (PBS + 0.1% Tween 20), and TEV reaction buffer (1 M Tris Cl (pH 8), 10 mM EDTA). Recombinant TEV protease (Invitrogen) was then added to the beads resuspended in TEV reaction buffer. After incubation for 30 minutes at room temperature, the beads were spun down and the supernatant was collected (elution 1). The step was repeated to collect 3 more eluted fractions. The most concentrated fractions were pooled and dialyzed in two steps into buffered 2 M urea followed by 8 M urea. The dialyzed sample (in buffered 8 M urea) was incubated with nickel-NTA beads to remove $ClpP^{trap}$ and the flow-through collected and concentrated using Amicon tubes (Millipore).

The *ygfZ* gene was amplified by PCR from *E. coli* W3110 genomic DNA and cloned into the over-expression vector pET28b. The resulting plasmid was transformed into *E. coli* BL21 (DE3)/pLysS strain. The N-terminally His-tagged YgfZ was over-expressed at 37°C and the cell pellets resuspended in 50 mM sodium phosphate buffer (pH 8), 300 mM NaCl, 10 mM imidazole, 1mM DTT (lysis buffer). The cells were lysed using lysozyme and treated with benzonase (Novagen) to remove nucleic acids. The lysate was cleared by centrifugation and the supernatant was incubated for 1 hour at 4°C with nickel-NTA agarose beads, pre-equilibrated in lysis buffer. The beads were collected by centrifugation, resuspended, and washed sequentially with lysis buffer and wash buffer (lysis buffer containing 20 mM imidazole). Bound protein was eluted in five fractions using buffer containing 500 mM imidazole. Fractions

containing YgfZ were identified by SDS-PAGE, buffer-exchanged into 20 mM Tris Cl (pH 7.5), 100 mM NaCl, 1 mM DTT using PD-10 columns, pooled, and the concentration determined by UV absorption at 280 nm.

In vitro Degradation Assay

YgfZ (100 nM), with or without folic acid (300 nM), was incubated with *E. coli* ClpXP (0.3 μM $ClpX_6$; 0.8 μM $ClpP_{14}$) in the presence or absence of *E. coli* SspB (300 nM monomer) at 30°C in PD buffer plus an ATP-regeneration system (Flynn et al., 2001). PD buffer contained 25 mM HEPES-KOH (pH 7.6), 5 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.032% NP-40, and 10% glycerol. The ATP-regeneration system contained 5 mM ATP, 50 μg/mL creatine kinase, and 5 mM creatine phosphate. Because of difficulty in resolving the bands corresponding to YgfZ and creatine kinase, a western blot was used to monitor YgfZ degradation. The samples were run on a 12.5% Tris-Glycine gel and transferred to a PVDF membrane (Millipore). The membrane was incubated with primary anti-His rabbit antibody (1:5000) followed by goat anti-rabbit antibody (1:5000). The blot was then developed using ECF substrate (GE Healthcare) and scanned on a Typhoon 9400 (GE Healthcare).

APPENDIX II

In Vitro Degradation of the ribosomal protein S7 by ClpXP

INTRODUCTION

Energy-dependent protein degradation by bacterial proteases plays essential roles in protein quality control and various regulatory processes in a cell. The proteases are compartmentalized with separate ATPase and proteolytic components (see Chapter 1). The ATPases belong to the AAA+ (ATPases associated with various cellular activities) super-family, members of which share a characteristic core ATPase domain. The AAA+ proteins can associate with proteolytic components to form active proteases such as ClpXP. ClpXP consists of the hexameric AAA+ protein ClpX and the tetradecameric serine peptidase ClpP. Both components are ring-like and stack to form a barrel-like structure. The active site serine residues in ClpP are sequestered in the internal chamber, to which access is controlled by ClpX. ClpX recognizes, unfolds, and translocates substrates into the active site chamber of the protease where the polypeptide is hydrolyzed by ClpP (Sauer et al., 2004).

It is very important for intracellular proteases to attain high levels of specificity given their access to proteins involved in key biological processes. Both domains of ClpX, the N-terminal domain and the ATPase domain, are involved in substrate selection. ClpX is typically able to identify substrates through recognition signals located near their N- or C-termini. One well-characterized recognition signal is an 11-residue degradation tag (degron) known as the ssrA tag. A specialized RNA called ssrA directs addition of the degradation signal to the C-terminus of incomplete polypeptides when ribosomes stall during translation (Karzai et al., 2000). ClpXP is able to recognize molecular elements in the ssrA tag and remove these aberrant polypeptides from the cell.

ClpX and other proteases also use additional proteins known as adaptors to modulate substrate selection. These adaptors can influence substrate choice and degradation rates by proteases. Multiple ClpX-specific adaptor proteins have been identified in *E.coli*. They interact simultaneously with the N-terminal domain of ClpX and specific substrates to facilitate ClpXP-mediated degradation. In general, adaptors have the potential to affect substrate specificity as well as expand the repertoire of substrates that can be degraded by proteases. Therefore, studying adaptor-mediated substrate selection should provide a more complete picture of how proteolysis is regulated to suit the cell's needs.

The ClpX-specific adaptor SspB functions as a delivery protein by tethering substrates to ClpX and increasing degradation rates by lowering the Michaelis-Menten constant (K_M) at low

substrate concentrations (Dougan et al., 2003; Levchenko et al., 2000; Wah et al., 2003). SspB consists of a folded N-terminal domain followed by a flexible linker and a short C-terminal tail region. The folded domain forms the substrate-binding domain (SBD) whereas the C-terminal tail region is responsible for interacting with the N domain of ClpX and is thus known as the XB region (ClpX-binding) (Dougan et al., 2003; Wah et al., 2003). *E. coli* SspB directly interacts with two known ClpX substrates: ssrA-tagged proteins and the N-terminal cleavage fragment of the extracytoplasmic stress protein RseA (NRseA) (Flynn et al., 2004; Flynn et al., 2001; Levchenko et al., 2000). There is surprisingly limited sequence similarity between the SspB-binding regions on the ssrA tag and NRseA (Flynn et al., 2004; Levchenko et al., 2005). Although the substrates bind overlapping sites on SspB, they bind in opposite orientations and share only one common interaction with the adaptor (Levchenko et al., 2005; Levchenko et al., 2003; Song and Eck, 2003). Thus, although there are two well-characterized SspB binding partners in *E.coli*, it is currently impossible to predict interaction between ClpX substrate(s) and SspB based on purely amino acid sequence motifs (Fig A2.1).

There is experimental evidence, however, suggesting that the degradation of additional ClpXP substrates may be affected by SspB (Flynn et al., 2004). NRseA was identified as an SspB-modulated ClpX substrate when Flynn et al. (2004) compared the ClpX substrate profiles of cells in the presence and absence of SspB by two-dimensional gel electrophoresis and mass spectrometry. Their results indicated that NRseA was preferentially degraded by ClpXP in the presence of SspB and further studies showed that SspB directly interacts with NRseA to increase NRseA's degradation rate. There were additional ClpX substrates which, based on the proteomic analysis of Flynn et al. (2004), appeared to be positively affected by SspB although follow-up experiments were not carried out to verify whether or not the adaptor was participating directly in degradation of these substrates.

Work done by Butland et al. (2005) identified S7, a protein in the 30S ribosomal subunit, amongst a number of potential SspB-interacting proteins. In addition, previous experiments have indicated that S7 interacts with ClpXP *in vivo* (Neher et al., 2006). Interestingly, there is compelling evidence that proteolysis may play a role in regulating the levels of ribosomal proteins (Flynn, 2004; Kuroda et al., 2001; Nishii et al., 2005; Petersen, 1990). This regulation is particularly important during amino acid starvation conditions when ribosomal proteins are degraded to generate amino acids for synthesizing stress-response enzymes. Therefore, S7 seemed a reasonable candidate ClpXP substrate, which may additionally be interacting with the adaptor SspB.

The S7 level in the cell is tightly regulated, which is perhaps not surprising given its role in initiating assembly of the 30S ribosomal subunit (Nowotny and Nierhaus, 1988). S7 is a translational repressor of its own operon (*str* operon), which codes for the ribosomal proteins S7 and S12 as well as the elongation factors EF-Tu and EF-G (Dean et al., 1981; Saito et al., 1994). Over-expression of S7 confers a growth defect, possibly due to translational repression of the *str* mRNA (Fredrick et al., 2000; Robert and Brakier-Gingras, 2001). Given the key role of S7 in ribosomal assembly, it is plausible that proteolysis may also be involved in regulating S7 levels in the cell. One possibility is that degradation is used to remove free S7, thereby alleviating its inhibitory effect on the *str* operon.

Here, we use an *in vitro* approach to investigate the possible roles of ClpXP and SspB on S7 proteolysis.

RESULTS AND DISCUSSION

The first step in testing S7 as a potential ClpXP and SspB substrate was to check *in vitro* degradation of purified S7 by ClpXP (Fig A2.1a). ClpXP was able to degrade S7 and the degradation rate appeared to be enhanced by SspB. However, there were technical difficulties in this assay. There was nucleic acid contamination of the purified protein despite treatment with nucleases during the purification procedure. Because of high UV absorbance of the contaminant at 260 nm, it was not possible to obtain an accurate measurement of S7 concentration. In addition, S7 had low solubility during the purification process resulting in loss of protein at different stages of the procedure. The purification was repeated using higher salt concentration in all buffers and, although S7 solubility improved in the presence of high salt, the purification was still not optimal.

To bypass the problem of nucleic acid contamination, S7 was purified under denaturing conditions and then refolded. *In vitro* degradation assays were repeated using this version of S7 and, as expected, ClpXP degraded S7. However, the previously observed enhancement of S7 degradation rate by SspB was difficult to reproduce (Fig A2.1b). Although there appeared to be a slight increase in degradation rate in the presence of the adaptor, the increase was not as substantial as that observed in earlier experiments. It is possible that the S7 concentration used in this assay was high enough to saturate the protease, thereby making it difficult to observe any rate enhancement by the adaptor. The nucleic acid contaminant in the natively purified S7

may have had an effect, making the result difficult to interpret. Additional experiments are required to confirm a role for SspB in S7 degradation.

Although it is unclear if SspB plays a role in S7 degradation, *in vitro* degradation of S7 by ClpXP was consistently observed. To elucidate features of the S7-ClpX interaction, degradation was tested using a variant ClpX without the N domain. This ClpXΔN variant forms an active protease with ClpP but cannot interact with adaptor proteins and some substrates such as the phage MuA protein and the DNA-binding protein Dps (Abdelhakim et al., 2008; Meyer et al., *in prep;* Thibault et al., 2006). Interestingly, ClpXΔNP was unable to degrade S7 with or without the nucleic acid contaminant, indicating a possible role of ClpX N domain in selection of S7 (Fig A2.1c, d).

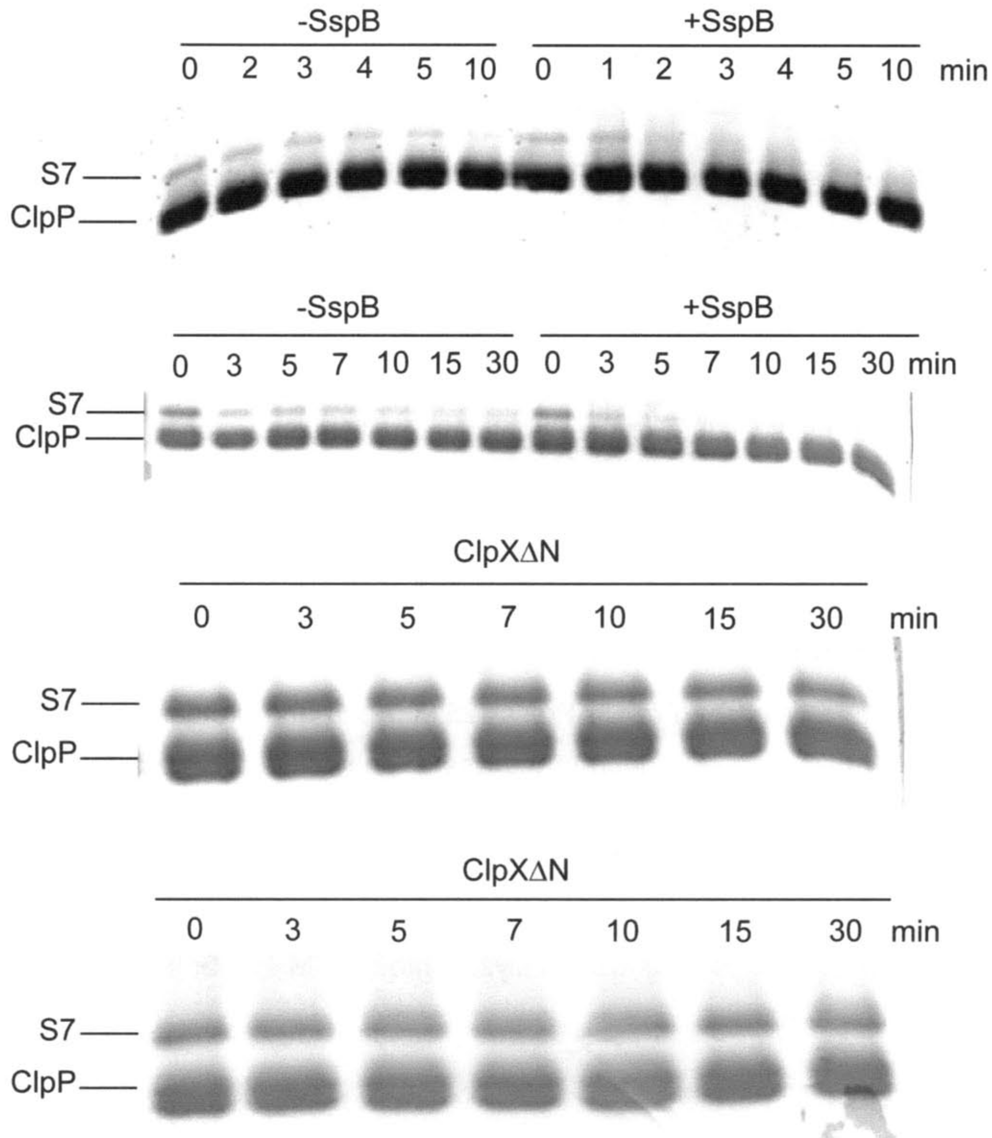


Figure A2.1. *In vitro* degradation of the N-terminally His-tagged ribosomal protein S7.

(a) *In vitro* degradation of S7 (purified under native conditions) by ClpXP, with and without SspB.
(b) *In vitro* degradation of S7 (purified under denaturing conditions) by ClpXP, with and without SspB.
(c) *In vitro* degradation of S7 (purified under native conditions) by ClpXΔNP, with and without SspB.
(d) *In vitro* degradation of S7 (purified under denaturing conditions) by ClpXΔNP, with and without SspB.

MATERIALS AND METHODS

Cloning and over-expression of S7

The *rpsG* (S7) gene was amplified by PCR from *E.coli* W3110 genomic DNA and cloned into the over-expression vector pET28b to purify an N-terminally His-tagged fusion protein. The plasmid was transformed into *E.coli* BL21(DE3)/pLysS cells and S7 expression induced for 2 hours at 37 °C. For the native purification protocol, cells were resuspended in 50 mM sodium phosphate buffer (pH 8), 500 mM NaCl, 10 mM imidazole (3 mL buffer/g of cells) and kept frozen at -80°C until ready for protein purification. For the denaturing purification protocol, the cell pellets were frozen (without resuspension).

S7 Purification under native conditions

The thawed cells were lysed by French Press (25,000 psi) and protease inhibitor cocktail III (Calbiochem) (0.67 µL/mL lysate) added to the lysate. The cell debris was spun down and the supernatant added to Ni-NTA beads (Qiagen). After incubation for 1 hour at 4°C, the beads were washed with 50 mM sodium phosphate buffer (pH 8), 500 mM NaCl, 20 mM imidazole. Fractions were eluted with 50 mM sodium phosphate buffer (pH 8), 500 mM NaCl, 250 mM imidazole and the concentrated ones were dialyzed into 50 mM Tris Cl (pH 8.5), 1 M KCl, 10% glycerol, 0.01% Triton X-100.

S7 Purification under denaturing conditions

The cell pellets were resuspended in lysis buffer (0.1 M sodium phosphate, 10 mM Tris, 6 M GuHCl (pH 8)) and spun down. The supernatant was incubated with Ni-NTA beads for 1 hour at room temperature and then washed with lysis buffer before elution with 0.2 M acetic acid, 6 M GuHCl. The fractions containing protein were refolded by slowly diluting into excess buffer (50 mM Tris Cl (pH 8.5), 1 M NaCl, 10% glycerol, 0.01% Triton X-100) on a stir-plate. The protein was then concentrated using spin columns (Amicon) and the concentration determined by UV absorption at 280 nm.

In vitro Degradation Assays

The assays were carried out at 30°C in PD buffer, which contained 25 mM HEPES-KOH (pH 7.6), 5 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.032% NP-40, and 10% glycerol. The ATP-regeneration system, containing 5 mM ATP, 50 µg/mL creatine kinase, and 5 mM creatine phosphate, was added to the reaction mix. The concentrations of *E.coli* $ClpX_6$ and $ClpP_{14}$ were 0.3 µM and 0.8 µM respectively. The ClpXΔN concentration was 0.3 µM and *E.coli* SspB concentration was either 0.15 µM or 0.5 µM (dimer equivalents). Concentration of S7 (purified under denaturing conditions) in the reaction was 5 µM. It was not possible to accurately calculate concentration of natively purified S7 because of nucleic acid contamination resulting in high absorbance at 260 nm. The amount of S7 added to the reaction was estimated so as it would be in the easily detectable range on 12.5% Tris-Glycine gels stained with Coomassie Blue stain.

REFERENCES

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

Almiron, M., Link, A.J., Furlong, D., and Kolter, R. (1992). A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. Genes & development *6*, 2646-2654.

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

Barkow, S.R. (2009). Mechanistic Studies of the AAA+ Molecular Motor ClpXP. In Department of Chemistry (Cambridge, MIT), p. 115.

Bolon, D.N., Grant, R.A., Baker, T.A., and Sauer, R.T. (2004a). Nucleotide-dependent substrate handoff from the SspB adaptor to the AAA+ ClpXP protease. Molecular cell *16*, 343-350.

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

Bouche, S., Klauck, E., Fischer, D., Lucassen, M., Jung, K., and Hengge-Aronis, R. (1998). Regulation of RssB-dependent proteolysis in Escherichia coli: a role for acetyl phosphate in a response regulator-controlled process. Molecular microbiology *27*, 787-795.

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

Bougdour, A., Wickner, S., and Gottesman, S. (2006). Modulating RssB activity: IraP, a novel regulator of sigma(S) stability in Escherichia coli. Genes & development *20*, 884-897.

Chien, P., Grant, R.A., Sauer, R.T., and Baker, T.A. (2007a). Structure and substrate specificity of an SspB ortholog: design implications for AAA+ adaptors. Structure *15*, 1296-1305.

Chien, P., Perchuk, B.S., Laub, M.T., Sauer, R.T., and Baker, T.A. (2007b). Direct and adaptor-mediated substrate recognition by an essential AAA+ protease. Proceedings of the National Academy of Sciences of the United States of America *104*, 6590-6595.

Clamp, M., Cuff, J., Searle, S.M., and Barton, G.J. (2004). The Jalview Java alignment editor. Bioinformatics (Oxford, England) *20*, 426-427.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome research *14*, 1188-1190.

De Las Penas, A., Connolly, L., and Gross, C.A. (1997). The sigmaE-mediated response to extracytoplasmic stress in Escherichia coli is transduced by RseA and RseB, two negative regulators of sigmaE. Molecular microbiology *24*, 373-385.

Dean, D., Yates, J.L., and Nomura, M. (1981). Identification of ribosomal protein S7 as a repressor of translation within the str operon of E. coli. Cell *24*, 413-419.

Donaldson, L.W., Wojtyra, U., and Houry, W.A. (2003). Solution structure of the dimeric zinc binding domain of the chaperone ClpX. The Journal of biological chemistry *278*, 48991-48996.

Dougan, D.A., Reid, B.G., Horwich, A.L., and Bukau, B. (2002). ClpS, a substrate modulator of the ClpAP machine. Molecular cell *9*, 673-683.

Dougan, D.A., Weber-Ban, E., and Bukau, B. (2003). Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX. Molecular cell *12*, 373-380.

Elsasser, S., and Finley, D. (2005). Delivery of ubiquitinated substrates to protein-unfolding machines. Nature cell biology *7*, 742-749.
Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D.A., and Bukau, B. (2006). ClpS is an essential component of the N-end rule pathway in Escherichia coli. Nature *439*, 753-756.
Erbse, A.H., Wagner, J.N., Truscott, K.N., Spall, S.K., Kirstein, J., Zeth, K., Turgay, K., Mogk, A., Bukau, B., and Dougan, D.A. (2008). Conserved residues in the N-domain of the AAA+ chaperone ClpA regulate substrate recognition and unfolding. The FEBS journal *275*, 1400-1410.
Farrell, C.M., Baker, T.A., and Sauer, R.T. (2007). Altered specificity of a AAA+ protease. Molecular cell *25*, 161-166.
Farrell, C.M., Grossman, A.D., and Sauer, R.T. (2005). Cytoplasmic degradation of ssrA-tagged proteins. Molecular microbiology *57*, 1750-1761.
Flynn, J.M. (2004). Substrate Selection by the ClpXP Protease: A Tail of Destruction. In Department of Biology (Cambridge, MA, Massachusetts Institute of Technology), p. 211.
Flynn, J.M., Levchenko, I., Sauer, R.T., and Baker, T.A. (2004). Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. Genes & development *18*, 2292-2301.
Flynn, J.M., Levchenko, I., Seidel, M., Wickner, S.H., Sauer, R.T., and Baker, T.A. (2001). Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. Proceedings of the National Academy of Sciences of the United States of America *98*, 10584-10589.
Flynn, J.M., Neher, S.B., Kim, Y.I., Sauer, R.T., and Baker, T.A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. Molecular cell *11*, 671-683.
Forman, M.S., Trojanowski, J.Q., and Lee, V.M. (2004). Neurodegenerative diseases: a decade of discoveries paves the way for therapeutic breakthroughs. Nature medicine *10*, 1055-1063.
Frank, E.G., Gonzalez, M., Ennis, D.G., Levine, A.S., and Woodgate, R. (1996). In vivo stability of the Umu mutagenesis proteins: a major role for RecA. Journal of bacteriology *178*, 3550-3556.
Fredrick, K., Dunny, G.M., and Noller, H.F. (2000). Tagging ribosomal protein S7 allows rapid identification of mutants defective in assembly and function of 30 S subunits. Journal of molecular biology *298*, 379-394.
Glynn, S.E., Martin, A.M., Nager, A.R., Baker, T.A., and Sauer, R.T. (2009). Crystal Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. Cell (in press).
Gonciarz-Swiatek, M., Wawrzynow, A., Um, S.J., Learn, B.A., McMacken, R., Kelley, W.L., Georgopoulos, C., Sliekers, O., and Zylicz, M. (1999). Recognition, targeting, and hydrolysis of the lambda O replication protein by the ClpP/ClpX protease. The Journal of biological chemistry *274*, 13999-14005.
Gonzalez, M., Frank, E.G., Levine, A.S., and Woodgate, R. (1998). Lon-mediated proteolysis of the Escherichia coli UmuD mutagenesis protein: in vitro degradation and identification of residues required for proteolysis. Genes & development *12*, 3889-3899.
Gonzalez, M., Rasulova, F., Maurizi, M.R., and Woodgate, R. (2000). Subunit-specific degradation of the UmuD/D' heterodimer by the ClpXP protease: the role of trans recognition in UmuD' stability. The EMBO journal *19*, 5251-5258.
Gottesman, S. (2003). Proteolysis in bacterial regulatory circuits. Annual review of cell and developmental biology *19*, 565-587.

Gottesman, S., Clark, W.P., de Crecy-Lagard, V., and Maurizi, M.R. (1993). ClpX, an alternative subunit for the ATP-dependent Clp protease of Escherichia coli. Sequence and in vivo activities. The Journal of biological chemistry *268*, 22618-22626.

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

Griffith, K.L., Shah, I.M., and Wolf, R.E., Jr. (2004). Proteolytic degradation of Escherichia coli transcription activators SoxS and MarA as the mechanism for reversing the induction of the superoxide (SoxRS) and multiple antibiotic resistance (Mar) regulons. Molecular microbiology *51*, 1801-1816.

Grimaud, R., Kessel, M., Beuron, F., Steven, A.C., and Maurizi, M.R. (1998). Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP. The Journal of biological chemistry *273*, 12476-12481.

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

Hansen, A.M., Qiu, Y., Yeh, N., Blattner, F.R., Durfee, T., and Jin, D.J. (2005). SspA is required for acid resistance in stationary phase by downregulation of H-NS in Escherichia coli. Molecular microbiology *56*, 719-734.

Hanson, P.I., and Whiteheart, S.W. (2005). AAA+ proteins: have engine, will work. Nature reviews *6*, 519-529.

Hersch, G.L., Burton, R.E., Bolon, D.N., Baker, T.A., and Sauer, R.T. (2005). Asymmetric interactions of ATP with the AAA+ ClpX6 unfoldase: allosteric control of a protein machine. Cell *121*, 1017-1027.

Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. Annual review of biochemistry *67*, 425-479.

Hoskins, J.R., Kim, S.Y., and Wickner, S. (2000). Substrate recognition by the ClpA chaperone component of ClpAP protease. The Journal of biological chemistry *275*, 35361-35367.

Hoskins, J.R., and Wickner, S. (2006). Two peptide sequences can function cooperatively to facilitate binding and unfolding by ClpA and degradation by ClpAP. Proceedings of the National Academy of Sciences of the United States of America *103*, 909-914.

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

Ishii, Y., and Amano, F. (2001). Regulation of SulA cleavage by Lon protease by the C-terminal amino acid of SulA, histidine. The Biochemical journal *358*, 473-480.

Jariel-Encontre, I., Bossis, G., and Piechaczyk, M. (2008). Ubiquitin-independent degradation of proteins by the proteasome. Biochimica et biophysica acta *1786*, 153-177.

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

Kang, S.G., Ortega, J., Singh, S.K., Wang, N., Huang, N.N., Steven, A.C., and Maurizi, M.R. (2002). Functional proteolytic complexes of the human mitochondrial ATP-dependent protease, hClpXP. The Journal of biological chemistry *277*, 21095-21102.

Karzai, A.W., Roche, E.D., and Sauer, R.T. (2000). The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. Nature structural biology *7*, 449-455.

Karzai, A.W., Susskind, M.M., and Sauer, R.T. (1999). SmpB, a unique RNA-binding protein essential for the peptide-tagging activity of SsrA (tmRNA). The EMBO journal *18*, 3793-3799.

Katayama, Y., Gottesman, S., Pumphrey, J., Rudikoff, S., Clark, W.P., and Maurizi, M.R. (1988). The two-component, ATP-dependent Clp protease of Escherichia coli. Purification, cloning, and mutational analysis of the ATP-binding component. The Journal of biological chemistry *263*, 15226-15236.

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

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

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

Kirstein, J., Moliere, N., Dougan, D.A., and Turgay, K. (2009). Adapting the machine: adaptor proteins for Hsp100/Clp and AAA+ proteases. Nat Rev Microbiol *7*, 589-599.

Klauck, E., Lingnau, M., and Hengge-Aronis, R. (2001). Role of the response regulator RssB in sigma recognition and initiation of sigma proteolysis in Escherichia coli. Molecular microbiology *40*, 1381-1390.

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

Lessner, F.H., Venters, B.J., and Keiler, K.C. (2007). Proteolytic adaptor for transfer-messenger RNA-tagged proteins from alpha-proteobacteria. Journal of bacteriology *189*, 272-275.

Levchenko, I., Grant, R.A., Flynn, J.M., Sauer, R.T., and Baker, T.A. (2005). Versatile modes of peptide recognition by the AAA+ adaptor protein SspB. Nature structural & molecular biology *12*, 520-525.

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

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

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

Levchenko, I., Smith, C.K., Walsh, N.P., Sauer, R.T., and Baker, T.A. (1997). PDZ-like domains mediate binding specificity in the Clp/Hsp100 family of chaperones and protease regulatory subunits. Cell *91*, 939-947.

Lo, J.H., Baker, T.A., and Sauer, R.T. (2001). Characterization of the N-terminal repeat domain of Escherichia coli ClpA-A class I Clp/HSP100 ATPase. Protein Sci *10*, 551-559.

Mackay, J.P., Kowalski, K., Fox, A.H., Czolij, R., King, G.F., and Crossley, M. (1998). Involvement of the N-finger in the self-association of GATA-1. The Journal of biological chemistry *273*, 30560-30567.

Marine, J.C., and Lozano, G. (2010). Mdm2-mediated ubiquitylation: p53 and beyond. Cell death and differentiation *17*, 93-102.

Martin, A., Baker, T.A., and Sauer, R.T. (2007). Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease. Molecular cell *27*, 41-52.

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

Martin, A., Baker, T.A., and Sauer, R.T. (2008b). Protein unfolding by a AAA+ protease is dependent on ATP-hydrolysis rates and substrate energy landscapes. Nature structural & molecular biology *15*, 139-145.

McGinness, K.E., Baker, T.A., and Sauer, R.T. (2006). Engineering controllable protein degradation. Molecular cell *22*, 701-707.

McGinness, K.E., Bolon, D.N., Kaganovich, M., Baker, T.A., and Sauer, R.T. (2007). Altered tethering of the SspB adaptor to the ClpXP protease causes changes in substrate delivery. The Journal of biological chemistry *282*, 11465-11473.

Meyer, A.S., Flynn, J.M., and Baker, T.A. (*in prep*). ClpXP degradation of the DNA-protection enzyme Dps requires auto-tethering to the enzyme.

Mika, F., and Hengge, R. (2005). A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of sigmaS (RpoS) in E. coli. Genes & development *19*, 2770-2781.

Missiakas, D., Mayer, M.P., Lemaire, M., Georgopoulos, C., and Raina, S. (1997). Modulation of the Escherichia coli sigmaE (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins. Molecular microbiology *24*, 355-371.

Mogk, A., Schmidt, R., and Bukau, B. (2007). The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. Trends in cell biology *17*, 165-172.

Muffler, A., Fischer, D., Altuvia, S., Storz, G., and Hengge-Aronis, R. (1996). The response regulator RssB controls stability of the sigma(S) subunit of RNA polymerase in Escherichia coli. The EMBO journal *15*, 1333-1339.

Nair, S., and Finkel, S.E. (2004). Dps protects cells against multiple stresses during stationary phase. Journal of bacteriology *186*, 4192-4198.

Navon, A., and Ciechanover, A. (2009). The 26 S proteasome: from basic mechanisms to drug targeting. The Journal of biological chemistry *284*, 33713-33718.

Neher, S.B., Sauer, R.T., and Baker, T.A. (2003). Distinct peptide signals in the UmuD and UmuD' subunits of UmuD/D' mediate tethering and substrate processing by the ClpXP protease. Proceedings of the National Academy of Sciences of the United States of America *100*, 13219-13224.

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

Nishii, W., Suzuki, T., Nakada, M., Kim, Y.T., Muramatsu, T., and Takahashi, K. (2005). Cleavage mechanism of ATP-dependent Lon protease toward ribosomal S2 protein. FEBS letters *579*, 6846-6850.

Nowotny, V., and Nierhaus, K.H. (1988). Assembly of the 30S subunit from Escherichia coli ribosomes occurs via two assembly domains which are initiated by S4 and S7. Biochemistry *27*, 7051-7055.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.

Park, E.Y., Lee, B.G., Hong, S.B., Kim, H.W., Jeon, H., and Song, H.K. (2007). Structural basis of SspB-tail recognition by the zinc binding domain of ClpX. Journal of molecular biology *367*, 514-526.

Petersen, C. (1990). Escherichia coli ribosomal protein L10 is rapidly degraded when synthesized in excess of ribosomal protein L7/L12. Journal of bacteriology *172*, 431-436.

Pines, J. (2006). Mitosis: a matter of getting rid of the right protein at the right time. Trends in cell biology *16*, 55-63.

Pratt, L.A., and Silhavy, T.J. (1996). The response regulator SprE controls the stability of RpoS. Proceedings of the National Academy of Sciences of the United States of America *93*, 2488-2492.

Pruteanu, M., and Baker, T.A. (2009). Proteolysis in the SOS response and metal homeostasis in Escherichia coli. Research in microbiology *160*, 677-683.

Robert, F., and Brakier-Gingras, L. (2001). Ribosomal protein S7 from Escherichia coli uses the same determinants to bind 16S ribosomal RNA and its messenger RNA. Nucleic acids research *29*, 677-682.

Saito, K., Mattheakis, L.C., and Nomura, M. (1994). Post-transcriptional regulation of the str operon in Escherichia coli. Ribosomal protein S7 inhibits coupled translation of S7 but not its independent translation. Journal of molecular biology *235*, 111-124.

Sauer, R.T., Bolon, D.N., Burton, B.M., Burton, R.E., Flynn, J.M., Grant, R.A., Hersch, G.L., Joshi, S.A., Kenniston, J.A., Levchenko, I., *et al.* (2004). Sculpting the proteome with AAA(+) proteases and disassembly machines. Cell *119*, 9-18.

Schmidt, R., Zahn, R., Bukau, B., and Mogk, A. (2009). ClpS is the recognition component for Escherichia coli substrates of the N-end rule degradation pathway. Molecular microbiology *72*, 506-517.

Schrader, E.K., Harstad, K.G., and Matouschek, A. (2009). Targeting proteins for degradation. Nature chemical biology *5*, 815-822.

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

Singh, S.K., Rozycki, J., Ortega, J., Ishikawa, T., Lo, J., Steven, A.C., and Maurizi, M.R. (2001). Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis. The Journal of biological chemistry *276*, 29420-29429.

Skerker, J.M., Prasol, M.S., Perchuk, B.S., Biondi, E.G., and Laub, M.T. (2005). Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. PLoS biology *3*, e334.

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

Stephani, K., Weichart, D., and Hengge, R. (2003). Dynamic control of Dps protein levels by ClpXP and ClpAP proteases in Escherichia coli. Molecular microbiology *49*, 1605-1614.

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

Tanaka, K., and Chiba, T. (1998). The proteasome: a protein-destroying machine. Genes Cells *3*, 499-510.

Teplyakov, A., Obmolova, G., Sarikaya, E., Pullalarevu, S., Krajewski, W., Galkin, A., Howard, A.J., Herzberg, O., and Gilliland, G.L. (2004). Crystal structure of the YgfZ protein from Escherichia coli suggests a folate-dependent regulatory role in one-carbon metabolism. Journal of bacteriology *186*, 7134-7140.

Thibault, G., Yudin, J., Wong, P., Tsitrin, V., Sprangers, R., Zhao, R., and Houry, W.A. (2006). Specificity in substrate and cofactor recognition by the N-terminal domain of the chaperone ClpX. Proceedings of the National Academy of Sciences of the United States of America *103*, 17724-17729.

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

Varshavsky, A. (1996). The N-end rule: functions, mysteries, uses. Proceedings of the National Academy of Sciences of the United States of America *93*, 12142-12149.

Wah, D.A., Levchenko, I., Baker, T.A., and Sauer, R.T. (2002). Characterization of a specificity factor for an AAA+ ATPase: assembly of SspB dimers with ssrA-tagged proteins and the ClpX hexamer. Chemistry & biology *9*, 1237-1245.

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

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

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

Wawrzynow, A., Wojtkowiak, D., Marszalek, J., Banecki, B., Jonsen, M., Graves, B., Georgopoulos, C., and Zylicz, M. (1995). The ClpX heat-shock protein of Escherichia coli, the ATP-dependent substrate specificity component of the ClpP-ClpX protease, is a novel molecular chaperone. The EMBO journal *14*, 1867-1877.

Weichart, D., Querfurth, N., Dreger, M., and Hengge-Aronis, R. (2003). Global role for ClpP-containing proteases in stationary-phase adaptation of Escherichia coli. Journal of bacteriology *185*, 115-125.

Williams, M.D., Ouyang, T.X., and Flickinger, M.C. (1994). Starvation-induced expression of SspA and SspB: the effects of a null mutation in sspA on Escherichia coli protein synthesis and survival during growth and prolonged starvation. Molecular microbiology *11*, 1029-1043.

Wojtkowiak, D., Georgopoulos, C., and Zylicz, M. (1993). Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of Escherichia coli. The Journal of biological chemistry *268*, 22609-22617.

Wojtyra, U.A., Thibault, G., Tuite, A., and Houry, W.A. (2003). The N-terminal zinc binding domain of ClpX is a dimerization domain that modulates the chaperone function. The Journal of biological chemistry *278*, 48981-48990.

Wolf, D.H., and Hilt, W. (2004). The proteasome: a proteolytic nanomachine of cell regulation and waste disposal. Biochimica et biophysica acta *1695*, 19-31.

Yakhnin, A.V., Vinokurov, L.M., Surin, A.K., and Alakhov, Y.B. (1998). Green fluorescent protein purification by organic extraction. Protein expression and purification *14*, 382-386.

Yeung, H.O., Kloppsteck, P., Niwa, H., Isaacson, R.L., Matthews, S., Zhang, X., and Freemont, P.S. (2008). Insights into adaptor binding to the AAA protein p97. Biochemical Society transactions *36*, 62-67.

Zhou, Y., and Gottesman, S. (1998). Regulation of proteolysis of the stationary-phase sigma factor RpoS. Journal of bacteriology *180*, 1154-1158.

Zhou, Y., Gottesman, S., Hoskins, J.R., Maurizi, M.R., and Wickner, S. (2001). The RssB response regulator directly targets sigma(S) for degradation by ClpXP. Genes & development *15*, 627-637.