Stony Brook University

Trans-translation Dependent Degradation of mRNAs and Proteins

A Dissertation Presented

by

Zhiyun Ge

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Biochemistry and Structural Biology

(Biochemistry and Molecular Biology)

Stony Brook University

December 2011

Stony Brook University

The Graduate School

Zhiyun Ge

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

A. Wali Karzai - Dissertation Advisor
Associate Professor, Department of Biochemistry and Cell Biology

Rolf Sternglanz - Chair Person of Defense
Distinguished Professor, Department of Biochemistry and Cell Biology

Todd Miller
Professor, Department of Physiology and Biophysics

David Thanassi
Professor, Department of Molecular Genetics and Microbiology

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Trans-translation Dependent Degradation of mRNAs and Proteins

by

Zhiyun Ge

Doctor of Philosophy

in

Biochemistry and Structural Biology

(Biochemistry and Molecular Biology)

Stony Brook University

2011

The translation of mRNAs without in-frame stop codons results in the following problems: the ribosomes are stalled and sequestered on the defective mRNA, the incomplete polypeptides produced are potentially toxic to the cells, and the defective mRNA can result in more futile translational cycles. Bacteria have developed a unique system, called *trans*-translation to resolve all of the above-mentioned problems. The key players of the *trans*-translation system are tmRNA and its protein cofactor SmpB. One part of my work showed that the system mediates the degradation of non-stop mRNAs by recruiting RNase R to stalled ribosomes through the interaction between the C-terminal lysine-rich domain of RNase R and the *trans*-translation machinery. In addition, the SmpB-tmRNA system adds a peptide tag to the C-termini of the nascent products of mRNAs lacking an in-frame stop codon, marking them for proteolysis. The other part of my work showed that in *Escherichia coli*, in addition to the ClpXP system, Lon protease also degrades tmRNA-tagged proteins, but with much lower efficiency. I then studied a unique case in *Mycoplasma pneumoniae* (*Mp*) where Lon is the primary protease degrading the tmRNA proteins due to the absence of Clp family proteases. I identified two discrete signaling motifs in the *Mp*-tmRNA tag for Lon recognition. I also showed evidence that the *Mp*-Lon and *Mp*-tmRNA tag have co-evolved to allow enhanced Lon binding to the tag and more efficient degradation of the *Mp*-tmRNA tagged proteins. The discrepancy in the substrate specificity between *Ec*-Lon and *Mp*-Lon promoted us to further explore the general substrate specificity of Lon in difference bacterial species. I was also interested in the role of Lon protease in the pathogenesis of *Yersinia pestis*. I performed a genome wide profiling of Lon substrates and identified novel substrates, many of which are master regulators of gene expression in *Yersinia*. I also identified and compared the preferred cleavage sites of Lon protease in multiple substrates. We are now working on deciphering the biological significance of the degradation of these

substrates by Lon, with the hope of shedding more light on the role Lon is playing in *Yersinia* pathogenesis.

Table of contents

List of figures

List of tables

Acknowledgments

This work wouldn't have possibly been done without the support and encouragement of family, friends, and colleagues. I want to start by thanking my parents, grandparents, and my sister for their constant support and encouragement during this process.

I want to thank all lab members of the Karzai lab for their helpful discussion of my work. I especially want to thank Pretti Mehta and Jamie Richards for working with me side by side on the work presented in Chapter 2. I learned a great deal from them. It had been a great pleasure working with them. I also want to thank all members of the Center for Infectious Diseases for their generous help on reagents and equipment.

I want to thank Dr. Atonius Koller for his help and collaboration on the work presented in Chapter 4. He did all the mass spectrometry work and helped me analyze the data. The project could never have gone that far without his expertise and generous help.

I would like to thank my committee members Drs. Sternglanz, Miller, and Thanassi. Their useful insight into my work has helped guide me through my graduate research.

Lastly, I would like to thank my thesis advisor Dr. A. Wali Karzai. I couldn't have done this without his support and guidance. I learned from him not only the way to do research but also how to become a better person. His enthusiasm for science will always be an inspiration for me.

Chapter 1: Introduction

1.1 Introduction

While cells undergo and complete countless translation cycles to produce proteins that are essential for cell survival and health, they also encounter numerous conditions where defective mRNAs, such as those missing in frame stop codons (non-stop mRNAs) are produced. Non-stop mRNAs can be produced as a result of mutations, premature transcription termination, post-transcriptional cleavage of the mRNA, and incomplete mRNA processing. These non-stop mRNAs still possess the ability to be translated. However, the translation of these mRNAs create various problems, including the stalling and subsequent sequestration of ribosomes because factors required for translation termination cannot be successfully recruited without stop codons, the production of incomplete polypeptides that can be potentially toxic to the cells, as well as futile translation cycles if the defective mRNA is not removed promptly. This problem universally occurs in both prokaryotic and eukaryotic cells. Here we explore the system developed in bacteria called *trans*-translation, which comprehensively resolves all above-mentioned problems. The key players of the trans-translation machinery are tmRNA and its protein cofactor SmpB. This system is activated, upon the event of ribosome stalling on non-stop mRNAs, and recognizes the stalled ribosome. The tmRNA-SmpB machinery then tags the nascent incomplete polypeptide for directed proteolysis, recruits RNase R to the stalled ribosomes to degrade the non-stop mRNAs, and works to terminate translation and release the stalled ribosomes.

1.2 Prokaryotic Protein Synthesis

Translation refers to the synthesis of proteins from mRNA templates. To be able to efficiently translation thousands of proteins with high fidelity is essential for the survival and health of the cell. As a result, the cells adopt various mechanisms to ensure the integrity of this process. The ribosome is the key component of the translation machinery. Ribosomes in prokaryotes are comprised of two subunits of different sizes, a smaller subunit (the 30S subunit) and a larger subunit (the 50S subunit). The 30S subunit is composed of over 1500 nucleotides and 21 proteins [1-3] and plays key roles in translation initiation, decoding, and controlling the fidelity of codon-anticodon interactions. The 50S subunit contains two RNA chains with a total of approximately 3000 nucleotides and 34 proteins and is responsible for peptide bond formation and the channeling of the nascent proteins through the ribosomal exit tunnel. The assembled, functional 70S ribosome contains three binding sites for tRNA molecules, the aminoacyl site (A-site), the peptidyl site (P-site), and the exit site (E-site) [2, 3].

Normal translation consists of three steps: initiation, elongation, and termination (Figure 1.1). In bacteria, when translation is being initiated, the 30S subunit of the ribosome interacts with the Shine-Delgarno sequence on the mRNA, which is complementary to the anti-Shine Delgarno sequence at the 3' end of 16S rRNA[4]. Three initiation factors: IF-1, IF-2, and IF-3 are also required for this process. IF-3 helps in the selection of initiator tRNA (fMet-tRNA^{fMet}) by destabilizing the binding of other tRNAs in the P site of the ribosome [1, 2, 5-9]. IF-2 is a GTPase that binds preferentially to fMet-tRNA^{fmet}. IF-1 increases the affinity of IF-2 for the ribosome [10]. The current model is that IF-1 binds to the A-site of the 30S ribosome, IF-2 binds over the A-site, the initiator tRNA is bound to the AUG start codon of the mRNA in the P-site, and IF-3 is in the E site [1, 2, 4, 10-12]. The association of the 50S subunit to form the 70S complex is facilitated by IF2 and the subsequent release of IF-1 and IF-3. Once the 30S subunit and the 50S subunit associate, the 70S ribosome complex is formed. At this point, the P-site is occupied by the fMet-tRNA^{fMet}, which is bound to the AUG codon on the mRNA, and the A-site now contains the tRNA corresponding to the next codon, which serves to start the elongation phase of translation [1, 2].

Translation elongation refers to the process where the nascent polypeptide is elongated one amino acid at a time via the decoding of the mRNA. The fidelity of translation elongation is ensured by many mechanisms. First of all, aminoacyl-tRNA synthetases, are responsible for the addition of the correct amino acid to the corresponding tRNA. Once formed, the aminoacylated tRNAs are then bound by elongation factor-Tu (EF-Tu), which is a GTPase. A ternary complex containing EF-Tu-GTP-aminoacyl-tRNA then forms [2, 12, 13] and is brought to the ribosomal A-site to interact with the codon residing in the A-site, which is to be decoded. When the correct tRNA enters the A-site and interacts with the mRNA codon, elements within the decoding center of the ribosome (primarily rRNA nucleotides A1492, A1493 and G530) detect proper Watson-Crick base pairing between the codon and anticodon and undergo a conformational change, which indicates the correct tRNA is bound in the ribosomal A-site. This signal is subsequently communicated to the GTPase center in the large ribosomal subunit and activates the GTPase domain on EF-Tu, resulting in the hydrolysis of GTP by EF-Tu [1, 14, 15]. The GDP bound EF-Tu has decreased affinity for tRNAs and thus dissociates from the ribosome, which allows the aminoacyl tRNA to fully enter the A-site in a process known as accommodation. The aminoacyl stem of the tRNA is positioned in the peptidyl-transferase center of the ribosome to allow the rapid transferring of the peptide chain to the A-site tRNA, leaving a deacylated tRNA in the P-site and a peptidyl tRNA in the A-site [1, 2, 12, 15, 16]. Another GTPase, Elongation Factor-G (EF-G), is required for the translocation of the tRNAs and the mRNA [1, 3, 17]. After translocation, the deacylated tRNA is moved into the E-site, the peptidyl tRNA is in the P-Site awaiting the transfer of the polypeptide chain, and the A-site containing the next codon is available for the entry of the next cognate tRNA-EF-Tu-GTP complex.

The last event in the prokaryotic translation is termination. Translation terminates when the ribosome encounters one of the three stop codons found in bacteria (UAA, UAG and UGA). Release factors (RF) recognize stop codons in the ribosome A-site. There are two classes of release factors. RF-1 and RF-2 belong to the class 1 release factor, each recognizing two of the three stop codons. RF-1 recognizes UAA and UAG, and RF-2 recognizes UAA and UGA [18]. RF-1 and RF-2 mediate hydrolysis and release of the peptide from the tRNA in the P-site [19]. The class 2 release factor, RF-3 then binds the ribosome and induces release of RF-1 or RF-2 [14, 20]. This 70S ribosome complex is disassembled by ribosome recycling factor (RRF) along with EF-G [21-23]. Finally, IF-3 is required for the removal of the deacylated tRNA from the 30S ribosomal subunit [24].

1.3 *Trans*-Translation

There are many instances where the ribosome finds itself translating an mRNA lacking sufficient signals to complete a translation cycle. For example, mRNAs lacking an in frame stop codon cannot successfully undergo translation termination. Non-stop mRNAs can result from gene mutations, DNA damage, pre-mature transcription termination, as well as cleavage of mRNAs by ribonucleaseses. When translation is initiated on these non-stop mRNAs, the ribosome can read the mRNA until it reaches the 3' end of the mRNA. The lack of the stop codon then results in the failure of the translation machinery to enter the termination phase, and thus causes the ribosome to stall at the 3' end of the non-stop mRNA. This has several consequences, including the sequestration of a large number of ribosomes, the production of incomplete polypeptides, which upon release to the cytoplasm can potentially be toxic to the cells, and the defective mRNA can lead to more futile translation cycles if not removed promptly (Figure 1.2A).

Bacteria have developed a unique system, called *trans*-translation to resolve the problems associated with ribosome stalling on the non-stop mRNA. The key players of the *trans*-translation process are tmRNA and SmpB. tmRNA is a bifunctional RNA, which can function both as a tRNA and an mRNA. SmpB is the protein cofactor of tmRNA and is essential to all tmRNA functions. Upon the event of ribosome stalling, a tmRNA-SmpB-EF-Tu complex recognizes and binds to the stalled ribosomes，tmRNA then functions as a tRNA and the nascent polypeptide is transferred to the alanine charge on tmRNA. tmRNA then acts as a surrogate mRNA and translation then resumes on the open reading frame on the mRNA domain of tmRNA. The translation of this ORF adds a degradation tag to the nascent polypeptide (hereinafter designated as tmRNA tagged proteins) and translation termination occurs when the ribosome reaches the stop codon in the tmRNA ORF. The ribosome is then released by release factors and the tagged protein is directed for proteolysis by cellular energy-dependent proteases.

In addition, tmRNA also facilitates the decay of the non-top mRNAs through the coordinated action of RNase R (Figure 1.2B).

The work presented here is mainly focused on the *trans*-translation mediated non-stop mRNA decay and the energy-dependent proteases involved in degrading tmRNA tagged proteins. I will further discuss these aspects of *trans*-translation below.

1.3.1 tmRNA

tmRNA is a unique bi-functional RNA molecule that possess activities similar to both tRNA and mRNA. The tRNA-like domain of tmRNA shares both sequence and structural similarity to $tRNA^{Ala}$ [25-29]. The tRNA-like domain of tmRNA consists of three domains found in $tRNA^{Ala}$: the amino acid acceptor stem, the TΦC arm and the D-loop. However, there is no anticodon arm in tmRNA. Like $tRNA^{Ala}$, tmRNA can be charged with alanine by alanyl-tRNA synthetase (Ala-RS) [28, 30]. Additional secondary structural elements in tmRNA include four RNA pseudoknots (PK1-PK4), whose specific functions are unclear. However, it was indicated in several previous studies that PK1 is important for tmRNA structure and possibly its function [28, 31-35]. The mRNA domain of tmRNA is located between pseudoknots 1 and 2 in the tmRNA, and it contains an open reading frame (ORF) encoding a 10-amino acid degradation tag (ANDENYALAA in *E. coli*) followed by two tandem UAA stop codons (Figure 1.3).

1.3.2 SmpB

SmpB is a small basic RNA binding protein, which binds tmRNA with high affinity and specificity. It is also essential for all known tmRNA functions [36, 37]. For instance, SmpB is required for stable interaction of tmRNA with stalled ribosomes [37, 38]. The structure of SmpB from *Aquifex aeolicus* determined in 2002 by NMR showed that it is predominantly an antiparallel β-barrell, with an embedded oligonucleotide-binding (OB) fold [39]. In addition, the crystal structure of SmpB in complex with the tRNA domain of tmRNA from *Aquifex aeolicus* was solved in 2003 and provided significant insight into how the tmRNA-SmpB complex binds to the stalled ribosome [40]. On the surface of the protein, there are clusters of residues, which are responsible for binding to tmRNA [36, 38, 40, 41]. Mutations of these residues abolish tmRNA binding both *in vivo* and *in vitro* [36]. SmpB also possesses an unstructured C-terminal tail that is essential for SmpB function. Extensive evidence has been provided to show that the C-terminal tail of SmpB facilitates the tRNA-like activity of tmRNA, by mimicking an anti-codon arm [42, 43]. SmpB variants that bear mutations in the C-terminal tail, or lack the C-terminal tail fail to support *trans*-translation, although they are fully capable of binding tmRNA and delivering it to the stalled ribosomes [43].

1.3.3 Other factors in *trans*-translation

Other than tmRNA and SmpB, the *trans*-translation process also requires various other factors. First of all, Alanyl-tRNA synthetase is required to charge tmRNA with alanine. Then the GTP bound form of EF-Tu recognizes the alanylated-tmRNA and a quaternary complex of SmpB-tmRNA-EF-Tu-GTP, binds to the A site of the stalled ribosomes. Ala-RS recognition and charging is essential for the function of tmRNA since non-aminoacylated tmRNA is not recognized by EF-Tu and does not associate with 70S ribosomes [28]. EF-Tu binds to the tRNA-like domain of tmRNA in the amino acid acceptor arm and protects the ester linkage of the aminoacylated molecule from hydrolysis. EF-Tu is also important for delivery and initial binding of tmRNA to stalled ribosomes [44]. The first transpeptidation reaction allowing the addition of the tmRNA alanine charge can occur in the absence of EF-Tu *in vivo* but at considerably reduced rates as compared to reactions where EF-Tu is present [45]. The complete mechanism of *trans*-translation is illustrated in Figure 1.4**.**

1.3.4 Biological importance of *trans*-translation

The *smpB* and *ssrA* (tmRNA coding) genes are present in all bacteria species examined to date [46-49]. Although it is not essential for the survival of most bacteria species, the SmpB-tmRNA translation quality-control system has been shown to be essential in pathogenic *Neisseria gonorrheae*, *Mycobacterium genitalium*, and *Mycobacterium pneumonia* [48, 50]*.*

In general, when lacking the SmpB-tmRNA quality-control system, bacteria are more sensitive to environmental stress and sublethal concentrations of translation-specific antibiotics, and bacteria also show a slow recovery from carbon starvation [51-55]. *Salmonella typhimurium* and *Yersinia pseudotuberculosis* lacking this system exhibit attenuated virulence. *Y. pseudotuberculosis* lacking the SmpB-tmRNA translational quality-control mechanism is severely defective in the ability to express and secrete *Yersinia* outer proteins (Yops) and is unable to cause lethal disease in a mouse infection model [55]. A similar phenotype is also observed in *Yersinia pestis* [56].

The SmpB-tmRNA system also plays a role in modulating gene expression through maintenance of intracellular concentrations of regulatory factors, including transcriptional activators and repressors [57, 58]. Consistent with this role, a number of regulatory proteins have been shown to be substrates of *trans*-translation, including LacI repressor, the λ-cI repressor, YbeL, GalE, and RbsK [59-61].

1.4 Bacteria mRNA decay

In both prokaryotic and eukaryotic cells, the turnover rate of cellular mRNAs has been proven to be another one of the means by which cells regulate gene expression, in addition to the differential rates of mRNA transcription and translation. However, mRNAs in prokaryotes exhibit significantly lower half-lives compared to the mRNAs in eukaryotes. The stability of a certain transcript has major impact on the protein expression levels as a longer-lived transcript is evidently subjected to more translation cycles than a shorter-lived transcript. The rate of mRNA turnover is closely related to the nucleases involved in the degradation, the directionality of degradation, and the sequence and structural features of the specific transcript [62-64]. Several mechanisms are employed in controlling the half-life a specific transcript by making it more or less susceptible to degradation. Over the years, two degradation pathways for prokaryotic mRNAs, the 5'- and 3'- end dependent pathways, have been extensively studied. However, the full details of the mRNA decay pathways remain to be elucidated.

1.4.1 5'-end dependent mRNA cleavage by endoribonucleases

One well understood mRNA decay pathway in *E. coli* is initiated by the cleavage of the mRNA by an endoribonuclease at one or multiple specific internal sites in a 5'-end dependent manner. Many *E. coli* endoribonucleases have been shown to possess the ability to initiate mRNA decay. However, RNase E has always been thought as the main ribonuclease responsible for 5'-end dependent mRNA cleavage [65]. This is supported by the evidence that the half-lives of many transcripts increase in *E. coli* in the absence of RNase E [66-68]. Other endoribonucleases such as RNase G, RNase Z, and RNase III have also been implicated in initiating mRNA decay, but in a more limited and specialized manner [69-72].

In the case of RNase E cleavage of the transcript, the rate of cleavage is largely dependent on the intrinsic factors of the transcript, especially the 5'-end. These features include the phosphorylation state of the extreme 5'-end, the secondary structure within the 5'-end, and whether the transcript is actively being translated by ribosomes. There is strong evidence that when a stem-loop structure is present at the 5'-end of the transcript, the half live of the transcript is lengthened due to the fact that the structure blocks the access of 5'–end dependent endoribonuclease to the transcript. One good example is the *ompA* transcript [73]. In recent years, new evidence surfaced that in addition to the secondary structure within the 5' end of the transcript, the phosphorylation state of the 5'-end is also an important factor that affects the cleavage of the transcript by RNase E. RNase E prefers a 5'- monophosphorylated substrate and exhibits limited activity against 5'- end triphosphorylated substrates [74]. However, the natural substrates of RNase E are all 5'-end triphosphorylated. Unless converted to monophosphorylated 5'-end, the cellular mRNAs are very poorly cleaved. These facts promoted a revision of the existing model of 5'-end dependent mRNA decay, especially for the initiation step. In a paper published by Celesnik

et al., researchers described a conversion of the natural triphosphorylated 5'–end to a monophosphorylated form by the removal of the 5'-pyrophosphate prior the RNase E cleavage [75, 76]. This newly identified step is also the rate-limiting step in the entire mRNA decay process. After the removal of the 5'-pyrophosphate, which allows RNase E to cleave the mRNA, the cleavage products are rapidly degraded by other exoribonucleases.

The complete degradation of the cellular mRNAs is a cooperative event between the 5'-end dependent endoribonuclease and various other exonriboucleases. However, there are no known processive 5'- 3' exribonuclease in *E. coli* [77]. The final degradation of the mRNAs is carried out by processive 3'–5' exribonucleases.

1.4.2 3'-end mRNA degradation by 3'-to-5' exonucleases

E. coli possess many different exoribonucueases, among which RNase II and PNPase are thought to be mainly responsible for the 3'-5' exoribonucleolytic decay of cellular RNAs. Both enzymes processively degrade substrates from the 3'-end, generating individual mononucleotides until only 2-5 nucleotides are left at the 5'-end. Oligoribonuclease then takes over and converts the short oligonucleotides into individual mononucleotides [78], (Figure 1.5).
RNase II and PNPase share a lot of similarities in their ribonucleolytic activities. For instance, both enzymes require at least 6-10 nucleotides of single stranded region at the 3'-end of the substrate to latch on [79]. However, there is established difference between these two enzymes in their ability to deal with RNA secondary structure. RNase II can degrade single stranded RNA without secondary structure but is unable to work through any internal structures present in its substrates. PNPase favors single stranded RNA and pauses at an internal stem-loop structure. But it has also been reported that PNPase, with the assistance of poly (A) polymerase I, and possibly RhlB helicase, can degrade through RNA secondary structures [80]. Another 3'-5' exonuclease, RNase R, has recently been indicated to be involved in the decay of some highly structured RNAs in *E. coli*, such as the REP (Repeated Extragenic Palindromic) – stabilizer [81]. RNase R shares extensive sequence homology with RNase II, especially in the catalytic domain. However, RNase R has two additional domains, a N-terminal domain and a C-terminal lysine-rich domain. Unlike RNase II, RNase R is able to efficiently unwind and degrade through RNA secondary structure, provided that there is a single stranded 3'-end to allow RNase R to bind to the substrate. The role of RNase R in mRNA decay is yet to be elucidated. Existing evidence suggest that RNase R may more actively participate in mRNA decay under stress conditions and it may have activity that neither RNase II nor PNPase can fulfill [82, 83].

1.4.3 Non-stop mRNA decay

Gene mutation, premature transcription termination, DNA damage, mRNA damage and translational errors can all lead to the generation of mRNAs lacking an in-frame stop codon. Due to the various problems the non-stop mRNAs can cause, the timely removal of these non-stop mRNA is of significant importance. It was recently reported that tmRNA promotes the decay of mRNAs that lead to ribosome stalling [84]. Previous work in our lab then identified RNase R as the RNase responsible for the degradation of the non-stop mRNAs in a *trans*-translation dependent manner [85, 86]. In addition, we also found that specific sequence elements within tmRNA are required to facilitate the decay of non-stop mRNAs by RNase R. A combination of these results suggests that a functional *trans*-translation system and RNase R are both required for the timely removal of the non-stop mRNAs. A possible mechanism was proposed for tmRNA-mediated decay of non-stop mRNAs, suggesting that the interactions of tmRNA, the ribosome, and RNase R grants access of RNase R to the defective mRNAs and facilitates the binding and degradation of the non-stop mRNA by RNase R. We further explored this hypothesis in this study.

1.5 Directed proteolysis of proteins generated from non-stop mRNAs

One major role of the *trans*-translation system is to remove the incomplete polypeptides resulting from the translation of non-stop mRNAs. The tmRNA-SpmB system achieves this by co-translationally adding a tag to the C-terminus of the nascent polypeptide, which contains signals required for directing the tagged proteins to a number of cellular energy dependent proteases for proteolytic degradation [87]. In *E. coli*, tmRNA tagged proteins are mostly degraded by ClpXP *in vivo*, although several other energy dependent protease also possess the ability to recognize and degrade them [88].

1.5.1 AAA+ Proteases

In order to meet the needs of all cellular activities and cope with environmental change, cellular proteins are subjected to various modifications. The degradation of damaged or misfolded proteins not only helps to remove potentially toxic proteins, but also contributes to recycle amino acids when environmental nutrition is depleted. In addition, the degradation of unwanted regulatory proteins is essential in maintaining the homeostasis of the cell. The vast majority of the protein degradation process is carried out by energy-dependent proteases. These are powerful degradation machineries that can degrade hundreds of cellular proteins, yet their proteolytic activity is tightly regulated to avoid accidental degradation of useful proteins.

In bacteria, cellular energy dependent proteases such as ClpXP, ClpAP, Lon, FtsH and HslUV belong to the AAA+ family. The main components of the

AAA+ proteases include the ATPase domain/subunit and the protease domain/subunit. The ATPase domain/subunit goes through cycles of ATP binding and hydrolysis to generate energy which can be translated into mechanical force [89]. These ATPases form a hexmaric ring with a central pore that is too narrow for folded proteins and can only accommodate unfolded polypeptide chains. Associated with the ATPase domain/subunit is the protease domain/subunit, which contains the active site of the protease. The protease domain/subunit also forms oligomeric rings with the active site secluded in an interior chamber and is only accessible through the central pores of the ATPase hexameric ring. The unique architecture of the AAA+ proteases contributes significantly to the substrate selection and degradation. In the case of Clp family proteases, the ATPase subunits form a hexameric ring that is stacked on the heptameric ring formed by the peptidase subunits. In the case of Lon and FtsH, the functional proteolytic machinery consists of six identical subunits, each of which contains an ATPase domain, a peptidase domain, and an additional N-terminal domain.

All AAA+ proteases adopt common strategies in substrate recognition and degradation, as well as the regulation of the protease activity. First of all, most substrates contain sequence motifs that mark them for recognition by AAA+ proteases (Figure 1.6). For example, ClpXP has 5 classes of sequence motifs that it recognizes: C-motif1 (LAA-COOH), similar to tmRNA tagged proteins; C-motif2 (RRKKAI-COOH); N-motif1 (polar-T/Φ-Φ-basic-Φ); N-motif2 (NH_2-Met-basic-Φ-Φ-Φ-X_5-Φ); and N-motif3 (Φ-X-polar-X-polar-X-basic-polar) [90]. The *E. coli* tmRNA tag belongs to the C-motif1 class. These sequence motifs mediate the degradation of the substrates by AAA+ proteases either directly by binding to the substrate recognition domain of the protease, or indirectly by binding to a cofactor and being delivered to the protease. For example, the *E. coli* tmRNA tagged proteins can bind directly to ClpXP through the C-terminal –LAA-COO^- motif, or they can be recognized first by the ClpXP cofactor, SspB, and delivered ClpXP through the interaction between SspB and ClpXP [91-94]. SspB recognizes the N-terminal portion of the tag leaving the C-terminal sequence accessible for ClpX binding. Cells also regulate the proteolytic activity of the AAA+ proteases by limiting the accessibility of the recognition motifs in substrates in certain condition, and exposing them in other conditions [90, 92, 95-104].

Once the AAA+ proteases successfully bind to the recognition motif in a substrate, a mechanical force generated from cycles of ATP binding and hydrolysis is applied to the substrate, which first brings the engaged substrate closer to the entrance of the central pore. The protease then attempts to unfold the substrate and translocate it through the narrow central pore by applying more mechanical force onto the substrate [105, 106]. At this point, some substrates will unfold cooperatively and allow subsequent translocation and degradation, while other substrates, despite the remarkable force applied, may still be able to resist unfolding and dissociate from the protease [107-111]. The unfolding power of different AAA+ protease varies intrinsically and is also dependent on the

characteristics of the substrate [92, 99, 104, 112]. Nonetheless, in the event where the AAA+ protease can successfully unfold and translocate the substrate, extensive conformational changes in the ATPase compartment occur as it goes through cycles of ATP binding and hydrolysis. These structural changes can be transmitted to the substrates through several loops that protrude into the ATPase central pore [113-120]. Although it is currently widely accepted that these pore loops play essential roles in substrate recognition and translocation, how the ATP powered loop movement promote substrate unfolding and translocation remains unknown.

Unfolded polypeptide chains that travel through the central pore of the ATPase compartment and reach the protease chamber, are cleaved, generating short peptides that are released back to the cytosol. Different AAA+ proteases adopt different active sites. For instance, ClpP has a His-Asp-Ser catalytic triad, while Lon has a Ser-Lys catalytic dyad [119, 121].

1.5.2 Lon protease

Lon protease was first discovered in *E. coli* over 30 years ago [122]. It is a cellular energy dependent protease and belongs to the AAA+ superfamily. Bacterial Lon protease (LonA subfamily) has three distinct domains: A large N-terminal domain that is thought to be involved in substrate recognition and binding [123], an ATPase domain that is responsible for hydrolyzing ATP and providing energy for the unfolding and translocation, and a C-terminal protease domain that is covalently attached to the ATPase domain (Figure 1.7). The protease domain contains the Ser-Lys catalytic dyad. Another family of Lon proteases (LonB family), mostly found in archaeal, contains a trans-membrane domain and is membrane associated. Due to the lack of the crystal structure of the full-length Lon protease, the mechanistic details of Lon protease in substrate degradation, as well as its oligomeric architecture remain largely unknown. However, cryo-EM studies and the crystal structure of individual domains of Lon suggest that it functions as a hexamer [123-125]. One unique feature of Lon protease that is not observed in other AAA protease is that Lon has DNA binding activity and DNA binding can stimulate the ATPase activity of Lon in vitro [126, 127].

Lon protease is involved a wide range of cellular activities. Although Lon is not essential in many bacteria species, probably due to the redundancy in substrate specificity among several AAA+ proteases, *E. coli* Lon mutants exhibit distinct phenotypes such as UV sensitivity due to the accumulation of cell division inhibitor SulA [128], and mucoid-like morphology due to the stabilization of RcsA, which is the transcriptional activator of capsule synthesis genes [129].

Lon protease plays an essential role in the bacterial stress response. In the event of DNA damage and bacterial SOS response, several factors are

induced, whose levels are kept very low under normal conditions due to their potential harmful effects to the cell. One of these factors is SulA, which can bind FtsZ and inhibit septum formation to delay cell division and allow time for DNA repair to occur. Upon completion of DNA repair, SulA is rapidly removed from the cells by the activity of Lon protease [130] and cell division is resumed. Lon mutants suffer filamentation and eventually cell death due to the accumulation of SulA when SOS response is induced. Lon's involvement in the SOS response is crucial but not limited to SulA degradation. Lon degrades two proteins, UmuC and UmuD, very rapidly under normal conditions to keep their presence at a minimal level. However, in the case of severe DNA damage, UmuD undergoes an autoproteolytic cleavage of 24 amino acids at the N-terminus, which makes UmuD resistant to Lon degradation and subsequent stabilization of UmuC by UmuD'. This allows the assembly of a lower fidelity DNA polymerase V that is required to transverse the lesion and allow normal DNA replication to resume [131, 132]. Lon is also required for the fitness and homeostasis of the bacteria in various other environmental stresses, such as nutritional depletion and acidic environment. In addition to the bacterial stress response, Lon also plays an essential role in bacterial communication [133, 134], bacteria motility [135-138] and biofilm formation [139, 140]. In recent years, more and more linkages have been established between Lon protease activity and the pathogenecity of several bacteria species such as *Yersinia pestis* [141], *Salmonella enterica* [142], *Brucella abortus* [143], *Campylobacter jeujeni* [135], and *Pseudomonas syringae* [144]. Several of these bacterial species possess a type three secretion system (TTSS) that injects virulence factors into the host cells. Lon mediates TTSS expression by degrading several key regulators that control TTSS expression. In the case of *Yersinia pestis*, Lon degrades the transcriptional repressor of TTSS, YmoA [141] and another TTSS thermal regulator RovA [145] at 37°C. Moreover, there is also evidence that Lon regulates the secretion of TTSS effectors by degrading the excessive effectors that are not secreted [146, 147].

1.6 Figures

Figure 1.1

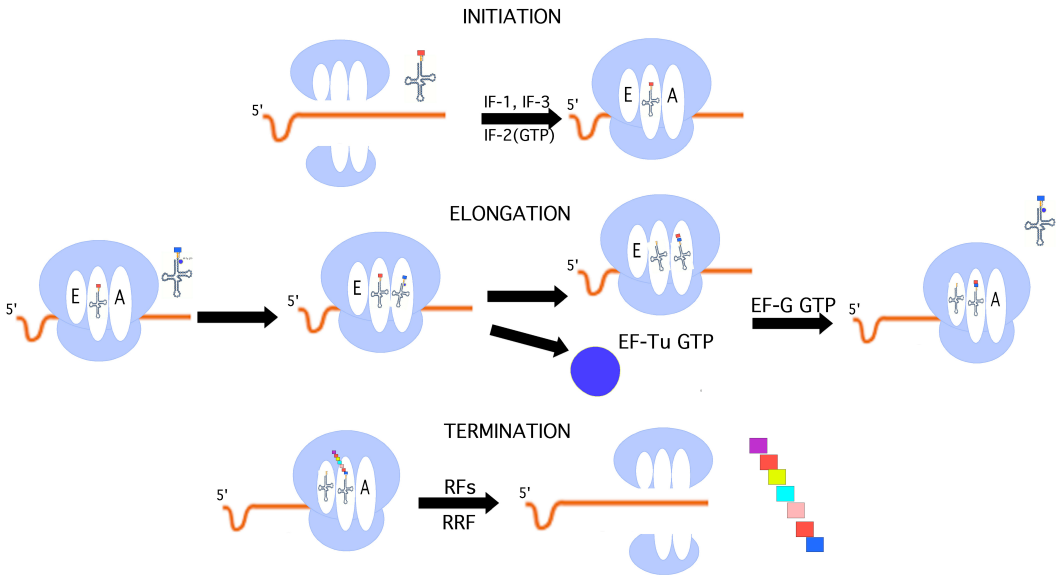


Figure 1.1: Three steps of normal translation: initiation, elongation and termination phases. See text for details.

Figure 1.2

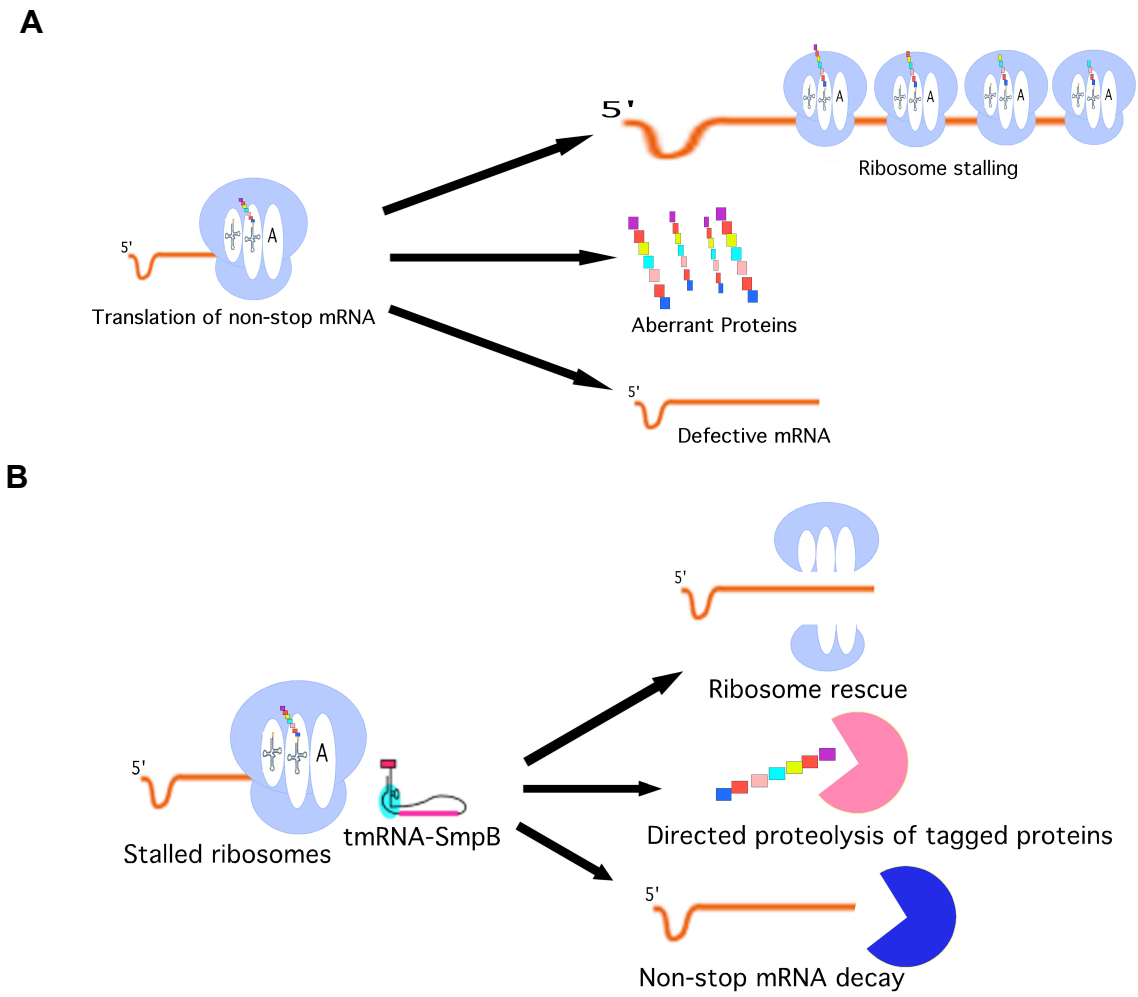


Figure 1.2: A) Problems associated with non-stop mRNAs: stalling and sequestration of ribosomes, release of aberrant and potentially toxic protein products, and more futile translation cycles if the defective mRNA is not removed promptly. B) The tmRNA-SmpB mediated translation quality-control system addresses the three major problems associated with stalled ribosomes through release of stalled ribosomes, directed proteolysis of the aberrant polypeptide and facilitating the decay of the damaged message.

Figure 1.3

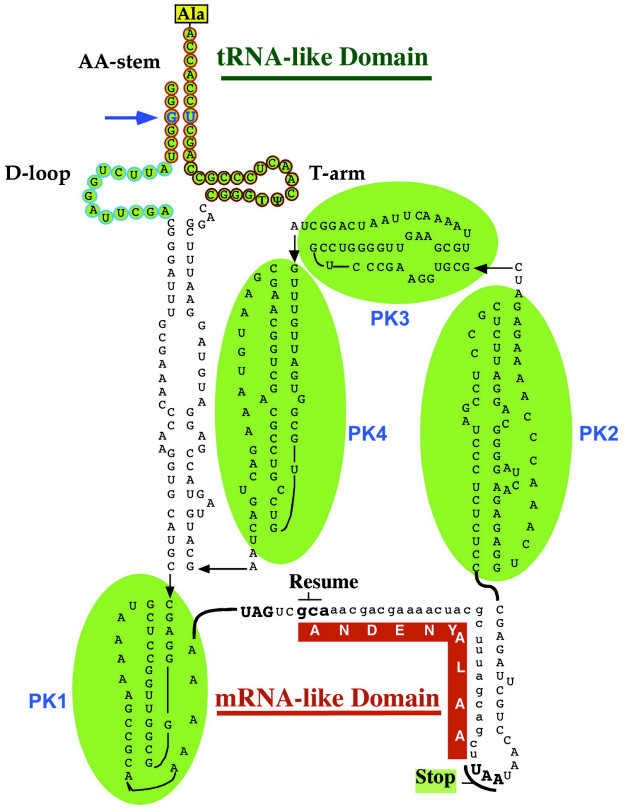


Figure 1.3: Proposed secondary structure of *E. coli* tmRNA. The tRNA- and mRNA-like properties are highlighted, including the amino acid acceptor stem, the D-loop, and the T-arm. The mRNA-like domain, with the tag peptide coding sequence, and the UAA termination signal are highlighted. The four pseudoknots are labeled PK1−4. Figure and legend adopted from Dulebohn *et al*. 2007.

Figure 1.4

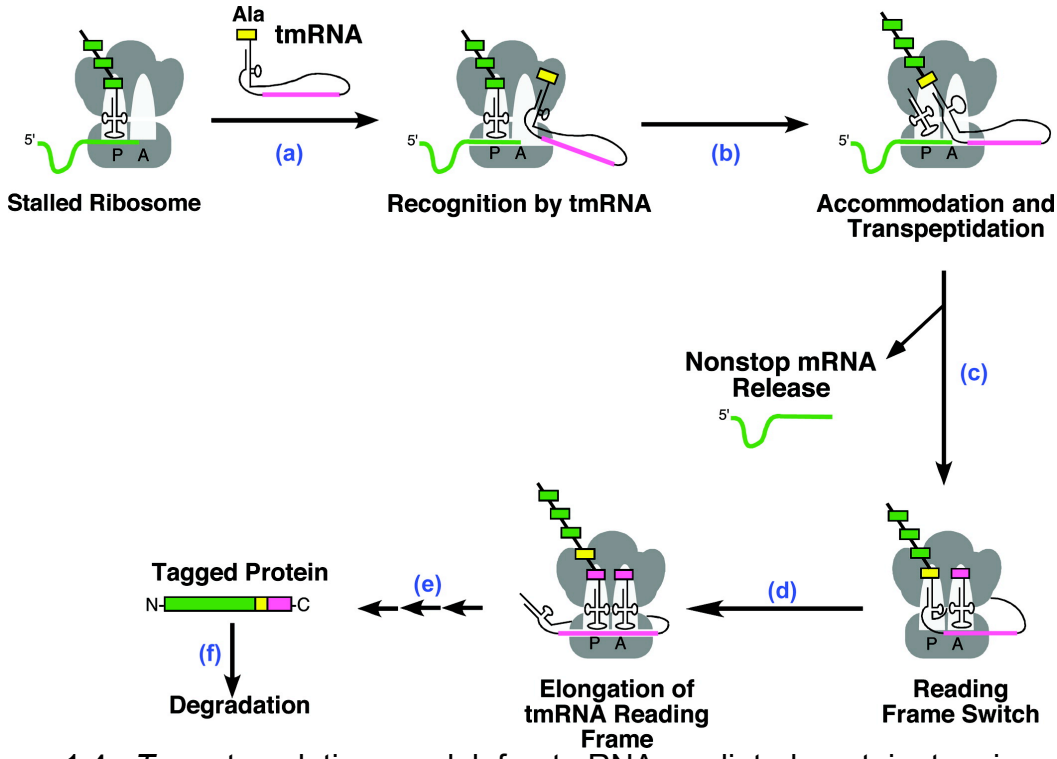


Figure 1.4: *Trans*-translation model for tmRNA-mediated protein tagging and ribosome rescue. A ribosome stalls on an incomplete or untranslatable message, leading to (a) the recruitment of aminoacylated tmRNA to the ribosomal A site and (b) transfer of the nascent chain to the alanine-charged tRNA-like domain of tmRNA. A message-switching event (c) then replaces the faulty mRNA with an open reading frame within tmRNA (d), which is translated until a stop codon is reached (e) and the tagged protein is released and degraded by C-terminal specific proteases (f). Figure and legend adopted from Dulebohn *et al*. 2007.

Figure 1.5

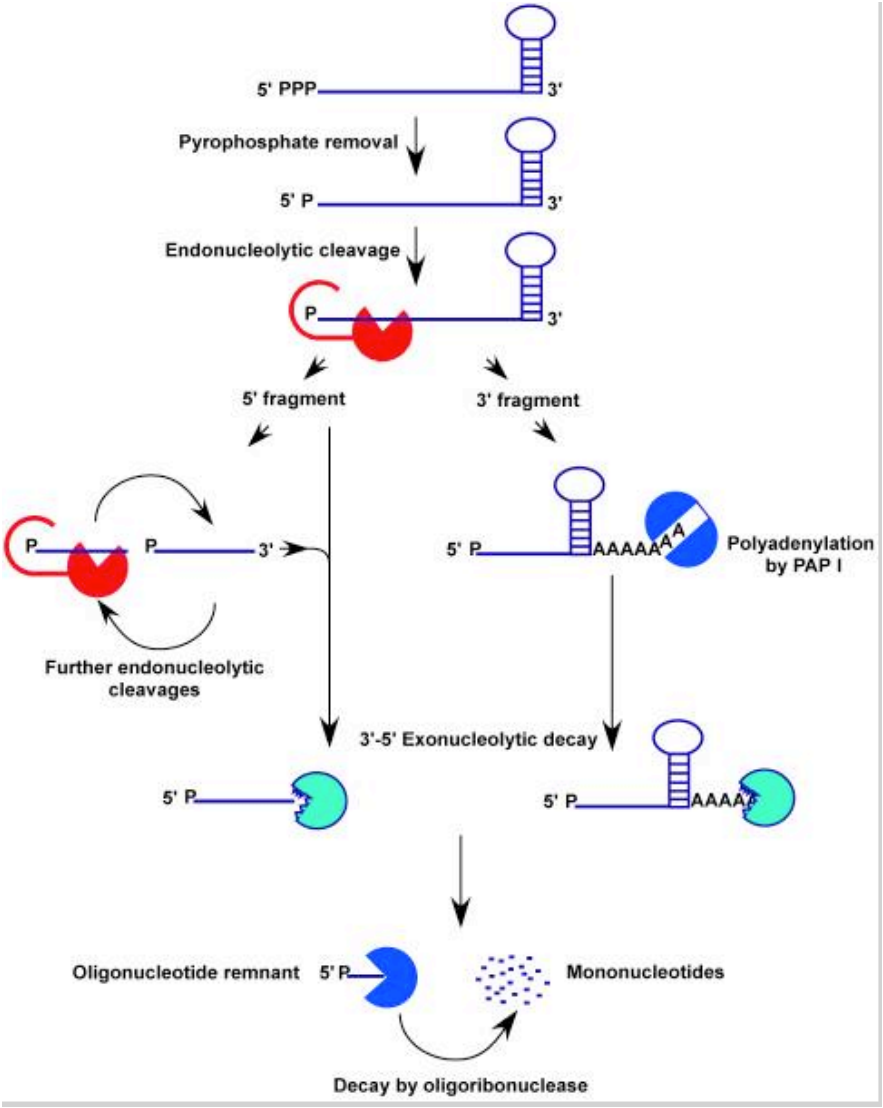


Figure 1.5: Schematic representation of a major pathway for mRNA decay in *E. coli*. A typical primary transcript possesses a single stranded, triphosphorylated 5′-terminus and a 3′-end with a stem–loop structure. Prior to internal cleavage by RNase E there is an RppH-dependent pyrophosphate removal step at the 5′-terminus. Internal, endonucleolytic cleavages are performed by RNase E, which requires the monophosphorylated 5′-end for catalytic activity. The monophosphorylated 5′-fragment is then subject to further endonucleolytic cleavages or 3′-5′ exonucleolytic decay by exoribonucleases such as RNase II, RNase R, and PNPase. Fragments that contain a 3′ stem–loop structure are polyadenylated by poly (A) polymerase (PAP I), allowing 3′-5′ exonucleolytic decay to be initiated by PNPase or RNase R. The final oligoribonucleotide product of this process is degraded to individual nucleotides by oligoribonuclease. Figure and legend adopted from Richards *et al*. 2008.

Figure 1.6

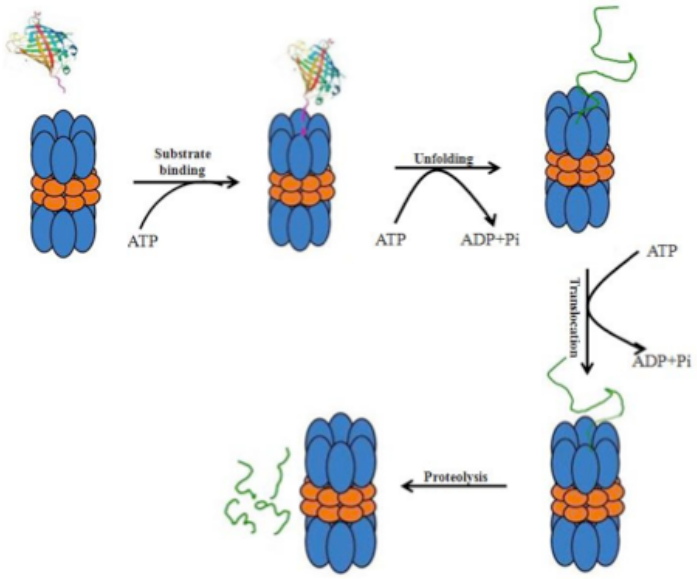


Figure 1.6: Degradation mechanism of AAA+ proteases. The protease complex, usually comprised of multiple ATPase subunits as well as multiple protease subunits, and shaped like a barrel with a central pore, recognizes certain motifs in the substrate and binds to it. The protease then uses the energy generated from ATP hydrolysis to unfold the protein and translocate the unfolded peptide chain through the central pore into the protease chamber, where peptide bond cleavage takes place. The final degradation products, the short peptides, are then released from the protease chamber.

Figure 1.7

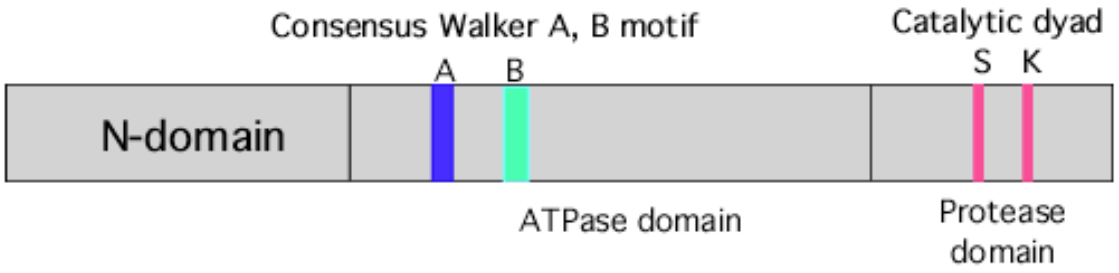


Figure 1.7: Primary structure of Lon. Lon has three distinct domains: the N-terminal domain, the ATPase domain, and the protease domain. Lon is a serine protease and the active site consists of a serine-lysine dyad in the protease domain.

Chapter 2 Non-stop mRNA decay initiates at the ribosome

*** The work contained within this chapter was published with the work of Preeti Mehta and Jamie Richards in Molecular Microbiology (2010) 78(5).**

2.1 Summary

The translation machinery deciphers genetic information encoded within mRNAs to synthesize proteins needed for various cellular functions. Defective mRNAs that lack in-frame stop codons trigger nonproductive stalling of ribosomes. We investigated how cells deal with such defective mRNAs, and present evidence to demonstrate that RNase R, a processive 3'-to-5' exoribonuclease, is recruited to stalled ribosomes for the specific task of degrading defective mRNAs. The recruitment process is selective for nonstop mRNAs and is dependent on the activities of SmpB protein and tmRNA. Most intriguingly, our analysis reveals that a unique structural feature of RNase R, the C-terminal lysine-rich (K-rich) domain, is required for both ribosome recruitment and targeted nonstop mRNA decay activities of the enzyme. These findings provide new insights into how a general RNase is recruited to the translation machinery and highlight a novel role for the ribosome as a platform for initiating nonstop mRNA decay.

2.2 Introduction

An enduring question in molecular biology concerns how evidently general enzymes, with a broad spectrum of potential targets, can recognize and process a select set of substrates that lack distinct sequence or structural features. We have been interested in how one such general enzyme (RNase R), which recognizes and degrades a wide assortment of cellular RNAs, is targeted selectively and in a spatio-temporally controlled manner to a defined set of substrates. We have recently shown that RNase R is directed to some of its mRNA substrates via a process that is dependent on the activities of SmpB protein and tmRNA [148]. The SmpB-tmRNA system orchestrates a versatile translational quality control process – termed *trans*-translation – that resolves challenges associated with decoding damaged or defective mRNAs. For instance, bacterial mRNAs that lack in-frame stop codons (nonstop mRNAs) still serve as decoding templates for the translation machinery. Ribosomes are fully capable of engaging these nonstop transcripts and executing the initiation and elongation phases of translation. However, the lead ribosome, having decoded all of the available coding information without encountering any in-frame termination signals, will stall at the 3'-end of the nonstop transcript. Stalling of the lead ribosome impedes the progress of all trailing ribosomes decoding the same message, resulting in potential accumulation of a large number of idled translation complexes. Buildup of such stalled ribosomes can give rise to two potentially perilous consequences for the cell. Firstly, sequestration of a large

number of ribosomes away from the active translating pool will result in significant loss of translational capacity. Secondly, the incomplete polypeptide products of these nonstop transcripts might be harmful to the cell. The SmpB-tmRNA quality control system has evolved to solve these problems by recognizing and rescuing stalled ribosomes, and directing the addition of a proteolytic tag to the C-termini of the incomplete polypeptide products [37, 40, 57, 58, 149-155].

We have recently shown that in addition to the ribosome rescue and directed proteolysis functions of the *trans*-translation system there is a requirement for SmpB protein and tmRNA in promoting the selective decay of defective mRNAs [85, 148]. For instance, an mRNAs that lacks a stop codon is much more efficiently degraded than a related transcript that possesses a stop codon. This selective mRNA decay, in addition to requiring SmpB protein and tmRNA, also requires the presence of RNase R [148]. This finding has led to the suggestion that RNase R might also be recruited to stalled ribosomes, either alongside SmpB and tmRNA or at an early stage of the *trans*-translation process [156].

The three-dimensional (3D) architectural organization of RNase R is not known. However, 3D-structural information for RNase II, a close paralogue of RNase R, is available [157, 158]. Based on available biochemical and structural data, RNase II can be separated into several functionally distinct domains (Figure 2.1A): an N-terminal segment containing two cold shock domains (CSD), a central region of ~400 residues spanning the nuclease (RNB) domain, and a C-terminal S1 domain. Based on sequence homology and functional similarity, RNase R is predicted to have similar overall domain architecture and adopt an analogous three-dimensional fold. As illustrated in Figure 2.1A, in addition to the CS, catalytic RNB, and S1 domains, RNase R possesses two distinct structural domains, the N-terminal putative helix turn helix (HTH) and the C-terminal lysine rich (K-rich) domains, that are not present in other RNases. Although RNase R is distinguished from other 3'-to-5' exonuclease by its ability to participate in the tmRNA-facilitated degradation of defective mRNAs [148], not much is known about how RNase R is directed to defective mRNAs, or which of its distinctive domains (Figure 2.1A) are responsible for its unique functions.

In this study, we investigated the potential recruitment of RNase R, a key component of the nonstop mRNA decay machinery, to stalled ribosomes and evaluated the functional contributions of the two unique RNase R domains to the *trans*-translation process. Our analysis demonstrates that RNase R is indeed recruited to ribosomes stalled on a nonstop mRNA and that the K-rich domain is essential both for recruitment to stalled ribosomes and selective decay of the defective transcript. Our data also demonstrate that enrichment of RNase R is specific to ribosomes translating nonstop mRNAs and is reliant on the presence of functional SmpB protein and tmRNA.

2.3 Results

2.3.1 Construction of RNase R truncation variants to evaluate *in vivo* mRNA decay

The mechanism by which RNase R selectively degrades nonstop mRNAs is not known. Sequence analysis shows that despite extensive sequence and structural similarity to RNase II, RNase R contains two distinct domains – the putative N-terminal HTH domain and the C-terminal K-rich domain – that are not present in other RNases (Figure2.1A). We postulated that one or both of these unique domains could serve as a possible determinant for the characteristic specificities exhibited by RNase R. To investigate whether these domains play a role in *trans*-translation, we generated N- and C- terminal truncated variants of RNase R (Figure 2.1B) and examined their *in vivo and in vitro* effects on RNase R related functions. We constructed a range of C-terminal K-rich domain variants (RNR^{789}, RNR^{770}, RNR^{751}, RNR^{759}, RNR^{749} and RNR^{723}), each with an incremental deletion of the lysine rich domain. RNase R^{723}, for instance, is missing the last 75 amino acids at the C-terminus, which encompasses the entire K-rich domain. RNase R^{734} and R^{744} lack the last 65 and 55 amino acids at the C-terminus, respectively. We also constructed RNase $R^{\Delta N}$, which is missing 64 amino acids at the N-terminus, encompassing the entire putative helix-turn-helix domain.

2.3.2 Degradation of *λ−cI-N* nonstop reporter mRNA by truncated RNase R variants

To assess the role the unique domains of RNase R in nonstop mRNA decay, we evaluated the N- and C-terminal truncation variants for their *in vivo* mRNA decay activity using plasmid encoded reporter transcripts. These transcripts are comprised of the coding sequence for the N-terminal domain of the bacteriophage λ *cI* gene followed by the strong *trp* operon transcriptional terminator (*trpA*t*)*. To ensure evaluation of a nonstop mRNA specific process, we utilized two related reporter variants (Figure 2.1C), a nonstop transcript lacking in-frame stop codons (*λ−cI-N* nonstop), and a "normal" transcript possessing a stop codon (*λ−cI-N* stop). The *λ−cI-N* nonstop transcript has been well characterized and is known to promote ribosomes to stall at its 3' end, whereas the variant containing a stop codon is readily translated and promotes efficient recycling of the translation machinery. Competence of RNase R variants in reporter mRNA decay was analyzed in a strain lacking the chromosomal copy of *rnr* and bearing plasmids for controlled expression of both the reporter transcript and one of the RNase R constructs. Expression of RNase R, and its truncation variants, was from a pACYC-based plasmid bearing the *rnr* coding sequence under the tight control of the arabinose inducible promoter.

We analyzed the decay rate of the *λ–cI-N* nonstop reporter mRNA in the presence of either a plasmid encoded full-length RNase R (pRNR^{WT}) or a control empty vector (pBAD). As shown in Figure 2.2, we observed rapid degradation of the nonstop mRNA in the presence of RNase R, with a half-life ($T_{1/2}$) of 0.8 ± 0.1 min. The stability of this transcript was substantially increased in the absence of RNase R, with $T_{1/2}$ of 4.5 ± 0.6 min (Figure 2.2 and Table 2.1). Analysis of the *λ–cI-N* nonstop reporter in the presence of RNase R variants with truncations proximal to the C-terminus (RNase R^{789}, RNase R^{770} and RNase R^{751}) showed mRNA half lives virtually identical to that of wild-type RNase R, with $T_{1/2}$ ranging between 0.8 ± 0.1 and 1.1 ± 0.1 min (Figure 2.2, Table 2.1), suggesting that the last 48 amino acids of the C-terminal domain did not play a major role in the selective degradation of nonstop mRNAs. The next three variants (RNase R^{744}, RNase R^{734} and RNase R^{723}) with larger deletions of the C-terminal domain, showed a clear decrease in selective mRNA decay activity compared to wild-type RNase R (Figure 2.2 and Table 2.1). Decay of the *λ–cI-N* nonstop reporter transcript was considerably slowed in the presence of these truncated RNase R variants, as compared to RNase R^{770} and full-length RNase R. More specifically, the *in vivo* half-lives of the *λ-cI-N* nonstop reporter in the presence of RNase R^{744}, RNase R^{734} or RNase R^{723} were 1.7 ± 0.2 min, 2.5 ± 0.3 min and 2.7 ± 0.3 min, respectively. We did not observe any significant effect on nonstop mRNA decay with RNase R$^{\Delta N}$, an RNase R variant lacking the N-terminal HTH motif (data not shown). These results suggested that the C-terminal domain of RNase R, in particular the region between residues 723 and 744 of the K-rich domain, played a crucial role in the *trans*-translation mediated degradation of nonstop mRNAs.

We reasoned that if the observed mRNA degradation defect was specific to nonstop mRNAs then the presence or absence of RNase R, or loss of the K-rich domain, should not have any effect on the stability of a related stop codon containing transcript. To test this hypothesis, we examined the stability of *λ–cI-N* stop reporter in an *rnr* $^{-}$ strain in the presence of plasmid encoded full-length RNase R, a K-rich domain deletion variant (pRNR^{723}), or a control empty vector. This analysis showed that the stability of the *λ–cI-N* stop reporter was not affected by the presence or absence of RNase R, as this reporter exhibited similar half-lives in cells lacking RNase R activity, expressing wild-type RNase R, or expressing the RNase R K-rich domain deletion variant (Figure 2.3, and Table 2.1). These results demonstrated that RNase R dependent mRNA decay was specific to nonstop mRNAs, and was mediated by the C-terminal K-rich domain of the protein. Overall, the stop codon containing mRNA was much more stable, decayed at a slower rate and was not affected by the presence or absence of the K-rich domain of RNase R.

2.3.3 Expression and steady state levels of RNase R and its variants

It was possible that the differences in activity we observed with the K-rich domain truncation variants of RNase R were due to differences in the expression level or stability of these proteins. To directly address this question, we examined the expression level of each construct in an *rnr*$^-$ background. Steady state levels of RNase R truncation variants were determined in cells grown to mid-log phase in media containing 0.01% arabinose, using western blot analysis with an RNase R specific antibody. For all of the plasmid borne constructs examined, RNase R polypeptides were expressed at notably higher levels than that of the full-length protein present in wild-type cells (Figure 2.4). However, a direct correlation between differences in expression of these RNase R variants and their activity in mRNA decay was not evident. More precisely, among the constructs examined, the three shortest RNase R variants that exhibited the greatest defect in targeted mRNA decay (RNase R^{744}, RNase R^{734} or RNase R^{723}) had higher steady state levels of RNase R protein in the cell, as compared to the longer and fully functional variants (RNase R^{789} and RNase R^{770}). These results indicated that the loss of targeted mRNA decay activity of K-rich truncation variants was not due to lower expression level or stability of these proteins.

2.3.4 Exoribonuclease Activity of RNase R variants *in vitro*

It was conceivable that the loss of selective mRNA decay and the increased stability of the nonstop reporter transcripts in cells expressing the most defective K-rich domain variants (RNase R^{723} and RNase R^{734}) was due to impairment of their catalytic activity. To assess this possibility, we examined the exonuclease activity of each variant in an *in vitro* RNA decay assay. To this end, we expressed and purified, to near homogeneity, full-length RNase R and several of its truncation variants (RNase R^{723}, RNase R^{734}, RNase R^{744}, and RNase R^{DN}). As a substrate for the *in vitro* RNA decay assay, we used a 32 nucleotide long single-stranded RNA (ss32), predicted to lack stable secondary structures. RNase II, the other RNase R family member, is fully capable of degrading unstructured RNA segments but is unable to degrade through structured regions in RNAs. Therefore, we included RNase II as a positive control for degradation of this single stranded unstructured RNA. The ribonuclease activity of each purified enzyme was assessed by its ability to degrade the ss32 substrate. The degradation reaction mixtures contained the RNA substrate and the appropriate buffer components, and were initiated by addition of the respective RNase. Degradation reactions were terminated at the indicated time points and the products were resolved on denaturing polyacrylamide gels (Figure 2.5). The activity of each RNase examined was indicated by the disappearance of the full-length substrates and appearance of lower molecular weight bands. Purified RNase R, RNase R^{DN}, and the K-rich domain truncation variants (RNase R^{744}, RNase R^{734}, and RNase R^{723}) were all equally capable of degrading the ss32 single-stranded substrate (Figure 2.5A). Similarly, purified RNase II was fully capable of degrading this substrate, confirming that the ss32 substrate did not possess stable secondary structure elements (Figure 2.5A). These results

demonstrated that the nuclease activity of RNase R truncation variants was not affected by the loss of the N-terminal HTH or the C-terminal K-rich domain.

It could be hypothesized that perhaps the requirement for RNase R function in degradation of nonstop mRNAs was due to the ability of RNase R to unwind and degrade local RNA secondary structures. Such a structure, the *trpA*t transcriptional terminator stem-loop, was present at the 3'-end of our *in vivo* reporters (Figure 2.1C). This possibility would also imply that the C-terminal K-rich domain of RNase R played a role in unwinding local RNA secondary structure elements, and that the loss of key parts of this domain in RNase R^{723}, RNase R^{734}, and RNase R^{744} resulted in the inability of these variants to support nonstop mRNA decay. To examine this possibility, we utilized two RNA substrates: *trpA*t-U_{10}, a 32 nucleotide RNA containing the *trpA*-terminator stem loop structure followed by a stretch of uridines at its 3'-end; and *trpA*t-A_{10}, a 32 nucleotide long RNA containing the *trpA*-terminator stem loop followed by a stretch of adenines at its 3'-end. We designed these substrates to mimic the endogenous structures of reporter mRNA substrates, and to assess whether the presence of a stretch of poly-U or poly-A tail had any effect on degradation of RNA substrates by RNase R. RNase II is known to lack the ability to degrade structured RNAs. We therefore included RNase II in our analysis to serve as a control and to ensure that the integrity of the structured RNA substrates was maintained. Assessment of the *trpA*t-U_{10} RNA showed that RNase R and its truncation variants were fully capable of degrading this substrate despite the presence of the internal stem-loop structure (Figure 2.5B). In sharp contrast, RNase II was unable to degrade through the secondary structure of this RNA, and managed to remove only a small part of the unstructured 3'-end (Figure 2.5B, first panel on the left). Similarly, our analysis of the poly-A-containing *trpA*t-A_{10} substrate showed that RNase R and all of its truncation variants were equally capable of degrading this RNA substrate, whereas RNase II was unable to complete the degradation process (Figure 2.5C). A comparison of these two substrates also indicated that RNase R preferred RNA substrates with a poly-A tail to substrates with a poly-U tail (Figure 2.5B and 2.5C). Taken together, the *in vitro* RNA degradation assays clearly demonstrated that loss of the C-terminal K-rich domain of RNase R affected neither its ribonuclease activity nor its ability to unwind internal secondary structure elements.

2.3.5 The C-terminal K-rich domain of RNase R is required for recruitment to stalled ribosomes

The alternative and decidedly more intriguing possibility for the observed increased stability of the nonstop mRNA in cells expressing the K-rich domain truncation variants was that these proteins suffered a defect in the *trans*-translation mediated recognition of the defective mRNA on stalled ribosomes. This possibility was compelling as it suggested that whereas full-length RNase R could be recruited to stalled ribosomes, its C-terminal truncation variants lacked

the requisite functional domain needed for such targeted recruitment. In order to ascertain whether RNase R was recruited to stalled ribosomes and to determine whether the K-rich domain played a role in this process, we performed stalled ribosome enrichment assays under high stringency conditions [38, 159]. These assays are reliant on the expression of the *λ-cI-N* nonstop transcript, which encodes for a protein that contains an N-terminal hexa-histidine (His_6) epitope and lacks in-frame stop codons, thus promoting ribosomes stalling. We surmised that ribosomes translating this reporter should have the His_6-epitope displayed on their surface and stall at the 3'-end of the nonstop message. The delay in the release of the attached His_6-polypeptide will thus permit the capture and enrichment, via affinity chromatography on a Ni-NTA column, of stalled ribosome from a pool of total ribosomes. We postulated that if RNase R was recruited to stalled ribosomes then it should be enriched in the captured ribosome fractions of cells expressing the *λ-cI-N* nonstop reporter. Furthermore, we reasoned that if the recruitment process was specific to nonstop transcripts then RNase R enrichment should not be as prominent in cells expressing the otherwise identical stop-codon containing *λ-cI-N* stop transcript. Moreover, to verify enrichment of stalled ribosomes, we routinely examined the captured ribosome fractions for the presence and enrichment of SmpB protein [38]. In all cases examined, SmpB protein was present and enriched only on ribosomes translating a nonstop mRNA. Analysis of the *λ-cI-N* nonstop reporter in *rnr* $^-$ cells expressing the plasmid borne full-length enzyme showed that RNase R was indeed recruited to stalled ribosomes expressing the nonstop reporter (Figure 2.6). We observed a consistent and reproducible increase in the amount of ribosome associated RNase R in these cells. In contrast, analysis of the related *λ-cI-N* stop reporter showed only background level association of RNase R with captured ribosomes (Figure 2.6). We did not observe any enrichment above and beyond this background. The background level of association might be due to either the routine sampling of ribosome by RNase R or the unintended production of defective transcripts during expression of the two plasmid borne genes.

Quantification of ribosome enrichment data, from several independent experiments, clearly showed RNase R recruitment to be substantially enhanced, by greater than 2.5-fold, in cells expressing the nonstop reporter as compared to the stop-codon containing transcript. Most significantly, analysis of these reporter transcripts in *rnr* $^-$ cells expressing the K-rich domain truncation variant (RNase R^{723}) showed only background level ribosome association of this defective variant in cells expressing either the nonstop or the stop codon containing reporter transcript (Figure 2.6), suggesting that enhanced recruitment of the truncated RNase R protein was blocked. Based on these findings, we concluded that the nonstop mRNA decay defect associated with the K-rich domain truncation variants of RNase R was due to the failure of these variants to be productively recruited to stalled ribosomes.

2.3.6 The recruitment of RNase R to stalled ribosomes is SmpB-tmRNA dependent

We have previously shown that the selective degradation of nonstop mRNAs by RNase R was dependent on the presence of SmpB and tmRNA [148]. In light of our current findings, we inferred that RNase R recruitment to stalled ribosomes should also be SmpB and tmRNA dependent. We therefore sought to determine whether recruitment of RNase R to ribosomes stalled on a nonstop transcript was dependent on *trans*-translation. To this end, we expressed *λ-cI-N* nonstop, or the related *λ-cI-N* stop reporter, in either an *smpB*$^{-}$*ssrA*$^{-}$ strain or the otherwise isogenic parental W3110 strain and isolated enriched stalled ribosomes. This analysis showed that in the W3110 parental strain RNase R was enriched in the captured ribosome pool of cells expressing the nonstop reporter transcript (Figure 2.7). In sharp contrast, RNase R was not enriched in the stalled ribosome fractions of *smpB*$^{-}$*ssrA*$^{-}$ cells, beyond the level of the control *λ-cI-N* stop transcript (Figure 2.7). These results confirmed the conclusion that SmpB and tmRNA functions were required for the selective and nonstop mRNA dependent recruitment of RNase R to stalled ribosomes.

2.4 Discussion

RNases are essential for processing and maturation of stable RNAs and for the disposal of defective or unwanted transcripts. In *E. coli*, 3'-to-5' exoribonucleolytic degradation is performed by three enzymes, RNase II, PNPase and RNase R. RNase II and RNase R are members of the RNR family and have a significant level of similarity in terms of primary sequence and domain architecture (Figure 2.1A), yet there are key functional differences between these enzymes. For example, RNase R has the ability to degrade RNAs containing secondary structure elements, while RNase II is incapable of degrading such structured RNAs. Similarly, RNase R has the unique capability to degrade defective mRNAs in a *trans*-translation dependent manner, whereas RNase II does not appear to be involved. What might account for the distinctive capabilities of RNase R? There are two possibilities. Firstly, it is possible that one or more of the analogous structural domains present in both enzymes (CS, S1, and the nuclease domain) have evolved to acquire novel functionalities in RNase R. Alternatively, one of the unique structural features of RNase R (the N-terminal HTH or the C-terminal K-rich domain) may confer these abilities to the enzyme.

The first of these explanations has been shown to be correct for the ability of RNase R to degrade structured RNAs. It was shown that the core nuclease domain of RNase R was capable of binding and degrading structured RNAs, albeit with much lower efficiency. The CS and S1 domains of RNase R were shown to participate in both substrate positioning and RNA binding [160-162]. We tested the second of these hypotheses in studies described herein, and have presented compelling evidence to demonstrate that RNase R is recruited to

ribosomes stalled on nonstop mRNAs, and that the unique K-rich domain of the enzyme is essential for the targeted degradation of defective transcripts.

We have previously shown that RNase R can degrade defective mRNAs in a SmpB-tmRNA dependent manner [148]. Here, we demonstrate that this function is mediated by the C-terminal K-rich domain of RNase R, whereby deletion of this domain resulted in a clear loss in its ability to accelerate the decay rate of a nonstop transcript. The decay rate of an otherwise identical transcript bearing an in-frame stop codon was not affected by the deletion of this domain. The K-rich domain deletion variants of RNase R were catalytically active, as evidenced by *in vitro* data, where absence of this domain had no effect on the decay of RNAs, both with and without 3' secondary structure elements. Furthermore, our evidence suggests that the role of the K-rich domain in the degradation of nonstop mRNA involves directing RNase R to stalled ribosomes. These data show a clear increase in the amount of RNase R associated with ribosomes stalled on a nonstop transcript, with the increase reliant on the presence of SmpB protein and tmRNA. These findings are entirely consistent with the recognition of stalled ribosome by the *trans*-translation machinery and the engagement of RNase R to degrade to causative defective mRNA. In addition, the increase in RNase R concentration on stalled ribosomes is dependent on the K-rich domain of the enzyme. This implies that RNase R is recruited to its role in nonstop mRNA decay through interactions mediated by its C-terminal K-rich domain.

Interestingly, we observed that RNase R was also associated with ribosomes translating a functional mRNA bearing a stop codon. This background association was independent of the type of reporter (nonstop or stop) or the RNase R variant used. These results suggest a mechanism whereby RNase R routinely samples translating ribosomes. Upon ribosome stalling the K-rich domain makes additional contacts with the stalled complex, perhaps with newly unmasked elements of the stalled ribosome, components of the *trans*-translation machinery, or the defective mRNA. The consequence of these interactions would be an increase in the residence time (i.e. enrichment) of RNase R on the stalled ribosome and the targeted delivery of RNase R to the causative defective mRNA. This model provides an intriguing parallel to other enzymes involved in the processing of nascent polypeptides, whereby bacterial ribosomes have been shown to serve as a platform for initiating of a number of co- and post-translational processing events. For instance, enzymes required for deformylation and removal of the initiator methionine, chaperones required for assisted folding and maturation of nascent polypeptides, and factors required for targeting and export of proteins have all been shown to interact with translating ribosomes (for a review see [163]). The spatio-temporally regulated recruitment of these factors to the ribosome is thought to be important for proper processing, targeting and folding of nascent polypeptides. The data presented here suggest an analogous mechanism, in which ribosomes serve as a platform for initiating targeted proteolysis and the selective degradation of causative defective mRNAs.

These findings are significant as they represent the first known case of recruitment of an exoribonuclease to bacterial ribosomes and provide new insights into the mechanisms by which bacteria deal with problems arising from defective mRNAs. The view that emerges from these studies suggests that bacteria have evolved an integrated solution to problems associated with nonproductive ribosome stalling via the SmpB-tmRNA mediated *trans*-translation process, which in addition to rescuing stalled ribosomes and tagging proteins for directed proteolysis, also recruits RNase R to stalled ribosomes for the specific task of degrading the associated mRNAs, thus coupling decoding of defective transcripts with their selective and spatio-temporally controlled removal from the cell.

2.5 Materials and Methods

2.5.1 Media and growth conditions

Cells were grown aerobically in LB at 37°C unless specified. Antibiotics were used at the following concentrations when needed: chloramphenicol (30 μg/ml), kanamycin (50 μg/mL), and ampicillin (100 μg/mL). When required IPTG was used at 1mM and arabinose at 0.01% or 0.02% (w/v). Bacterial growth was monitored by optical density measurements at 600 nm (OD_{600}).

2.5.2 Strains and plasmids

Escherichia coli strain W3110 *Δrnr* (DE3) was used for the expression of RNase II, RNase R, and its truncation variants. Ribosome enrichment assays [38, 159] were performed in strains MG1655 *rnr::KAN* or W3110 *smpB⁻ssrA⁻ rnr::KAN*. Reporter mRNA degradation was assessed in *E. coli* strain MG1655 *rnr::KAN* harboring pACYC/BAD-RNase R, or derivatives thereof, and either pλ-cI-N nonstop or pλ-cI-N stop reporter plasmids. The pACYC/BAD-RNase R plasmid was constructed by cloning the RNase R coding sequence between the *Nde* I and *Xho* I sites in MCS2 of pACYC Duet-1 vector (Novagen). To place the inserted coding sequence under the control of the pBAD promoter a *BspH* I / *Hind* III fragment of pBAD18, containing the *araC* gene and the pBAD promoter region, was cloned upstream of RNase R between the *Nco* I and *Hind* III sites of pACYC Duet-1. C-terminal truncation variants of RNase R were construction by introducing a stop codon after codons for amino acid 723, 734, 744, 751, 770 and 789. These variants are referred as RNR^{723}, RNR^{734}, RNR^{744}, RNR^{751}, RNR^{770} and RNR^{789} respectively. The N-terminal truncation variant was constructed by deleting the coding sequence for amino acids 1-64 and is referred as $RNR^{\Delta N}$.

To construct the pλ-cI-N nonstop reporter, six histidine codons (His_6-epitope) were introduced before the coding sequence of N-terminal domain of the

λ-*cI* gene [149]. Two tandem stop codons were inserted before the *trp*A terminator in pλ-cI-N to generate the pλ-cI-N stop reporter construct.

The *RNase II* gene was cloned into the *NcoI* and *BamHI* sites of pET28b for over-expression and purification purposes. The plasmid pET-RNR for expression of WT RNase R was a kind gift from Dr Murray Deutscher. The RNase R N-terminal truncation variant was constructed by removing the coding region for the first 87 amino acids at the N-terminus. The various RNase R C-terminal truncation variants of RNase R were constructed by introducing stop codons at the desired sites, using standard site directed mutagenesis protocols.

2.5.3 Measurement of mRNA decay

E. coli cells harboring a reporter plasmid and an RNase R expression plasmid (pBAD-RNR), or its truncation variants, were grown in media containing 0.01% arabinose to an OD_{600} of 0.6, at which point 1 mM IPTG was added to induce expression of the reporter mRNA. Expression was allowed to progress for 1hr before the addition of rifampicin (0.45 mg/ml) to inhibit transcription. Aliquots were taken at indicated time points, quick chilled and placed on ice. Cells were pelleted and total cellular RNA was extracted using TRI-Reagent (Molecular Research Center). Total RNA was subjected to gel electrophoresis on 1.5% (w/v) formaldehyde-agarose gels. Resolved RNA was transferred to Hybond N+ (GE Healthcare) overnight using capillary transfer. The reporter mRNA was probed using the biotinylated DNA oligonucleotide '500 probe' (5'-TTCATAAATTGCTTTAAGGCGACGTGCGTCCTCAAGCTGCTCTTGTGTTA-3'). Hybridizations were carried out for 16-20 hours at 42°C. Equal loading of RNA was confirmed by ethidium bromide staining of the gel and comparison of 16S rRNA bands, or by using a 16S rRNA probe (5'-TCAGATGCAGTTCCCAGGTTGAGCCCGGGGATTTCACATCTGACTTAACA-3'). Blots were developed using the BrightStar® BioDetect™ Nonisotopic Detection Kit (Applied Biosystems) and were exposed to film, bands were quantified by scanning and subsequent analysis with Image-J (http://rsbweb.nih.gov/ij/). mRNA half-lives were calculated by linear regression analysis of data, obtained from at least three independent experiments.

Levels of WT and truncated RNase R produced in these experiments were estimated by running aliquots from cells grown under similar conditions as for the decay experiments prior to rifampicin addition on an SDS-PAGE. The gel was electro-blotted onto a PVDF membrane, probed with rabbit polyclonal anti-RNase R antibodies and developed by the ECL method (GE Healthcare). Bands were quantified by scanning and subsequent analysis with Image-J. The RNase R expression level of each construct was calculated as a percentage of wild-type expression.

2.5.4 Expression and Purification of RNase II, RNase R, and RNase R truncation variants

Strain W3110 *Δrnr* (DE3) harboring pET28b-RNB, pET-RNR^{WT}, or one of the RNase R truncation variants, was grown at 37°C to OD_{600} 0.6 and the expression of all constructs was induced by addition of IPTG to a final concentration of 1.0 mM. Growth was continued for an additional 3 hours and cells were harvested by centrifugation at 3,700 x g for 60 min at 4°C. The resulting cell pellets were stored at –80°C. Frozen cell pellets were resuspended in an appropriate buffer, lysed on ice by four 30s pulses of sonication with 30s interval between each pulse. 100 µl of 100 mM PMSF was added after each pulse to prevent protein degradation during cell lysis. Cell debris was removed by centrifugation of the lysed cells at 10,000 x g for 1 hour at 4°C. The resulting supernatant was further cleared by passage through a 0.45 µm syringe filter. All purification steps were performed at 4˚C using an ÄKTA-FPLC system (GE healthcare), unless indicated otherwise.

To purify RNase II, the frozen cell pellets were thawed in Buffer A (50 mM Tris-HCl pH7.5, 50 mM KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, and 2 mM β-ME). Cell were lysed and indicated above and the filtered supernatant was loaded onto a 5 mL Hi-Trap-Blue HP column (GE healthcare) equilibrated in Buffer A. RNase II was eluted upon application of a linear gradient of 50 mM to 1.0 M KCl in Buffer A. Fractions containing RNase II were pooled and concentrated. At this point, the RNase II protein is associated with RNA, presumably its cellular substrates. The protein-RNA complex was incubated at 37°C for 30 minutes to allow the bound RNA to be digested by RNase II. The solution was then cooled to 4°C and the KCl concentration was reduced to 50 mM before loading the protein sample onto a Mono-Q HR 10/10 column (GE healthcare), pre-equilibrated in Buffer A. RNase II was eluted upon application of a linear gradient of 50 mM to 1.0 M KCl in Buffer A. Fractions containing RNase II were pooled, concentrated and loaded onto a Superdex 75 10/300 GL size-exclusion column equilibrated in Buffer B (50 mM Tris-HCl pH7.5, 300 mM KCl, 1 mM $MgCl_2$, 1 mM DTT, and 10% (v/v) glycerol). Fractions containing RNase II were pooled, concentrated, portioned into small aliquots and stored at -80°C.

To purify RNase R^{WT} and the N-terminal truncation variant RNase R$^{\Delta N}$, the frozen cell pellets of each variant were thawed in Buffer C (20 mM HEPES pH7.5, 300 mM KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, and 2 mM β-ME). Cells were lysed and the filtered supernatant was loaded onto a 5 mL Hi-Trap Blue HP column (GE healthcare) equilibrated in Buffer C. RNase R was eluted upon application of a linear gradient of 300 mM to 1.0 M KCl in Buffer C. RNase R containing fractions were pooled concentrated and incubated at 37°C for 30 minutes to allow the bound RNA to be digested by the RNase R. The solution was then cooled to 4°C and the KCl concentration was reduced to 300 mM before loading onto a Mono S HR 10/10 anion-exchange column (GE healthcare), equilibrated in Buffer C. RNase R was eluted upon application of a

linear gradient of 300 mM to 1.0 M KCl in Buffer C. Fractions containing RNase R were pooled, concentrated and loaded on to a Superdex 75 10/300 GL size-exclusion column equilibrated in Buffer B. Fractions containing RNase R^{WT} and RNase $R^{\Delta N}$ were pooled and concentrated, respectively. Protein aliquots were stored at -80°C.

To purify RNase R C-terminal truncation variants RNase R^{723}, RNase R^{734}, and RNase R^{744}, the frozen cell pellets were thawed in Buffer D (50 mM Tris-HCl pH7.5, 300 mM KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, and 2 mM β-ME). Cells were lysed and the filtered supernatant was loaded onto a 5 mL Hi-Trap Blue HP column (GE healthcare) equilibrated in Buffer D. RNase R was eluted from the column upon application of a linear gradient of 300 mM to 1.0 M KCl in Buffer D. Fractions containing RNase R were pooled, concentrated and incubated at 37°C for 30 minutes to allow the bound RNA to be digested by RNase R. The solution was then cooled to 4°C and KCl concentration was reduced to 300 mM before loading the sample onto a Mono-Q HR 10/10 column (GE healthcare) equilibrated in Buffer D. RNase R was eluted upon application of a linear gradient of 300 mM to 1.0 M KCl in Buffer D. Fractions containing RNase R were pooled, concentrated and loaded onto a Superdex 75 10/300 GL size-exclusion column equilibrated in Buffer B. Fractions containing RNase R^{723}, RNase R^{734}, and RNase R^{744} were pooled and concentrated. Protein aliquots were stored at -80°C.

2.5.5 Preparation of oligoribonucleotide substrates

Three different oligoribonucleotide substrates were prepared: a single stranded RNA with no secondary structure (ss32: 5'-CCCCACCACCAUCACUUAAAAAAAAAAAAAAA-3'); a single stranded RNA with an internal step-loop structure and a poly-A overhang at its 3'-end (*trpA*t-A_{10}: 5'-gcagcccgccuaaugagcgggcAAAAAAAAAA-3'); and a single stranded RNA with a poly-U overhang at its 3'-end (*trpA*t-U_{10}: 5'-gcagcccgccuaaugagcgggcUUUUUUUUUU-3'). These substrates were all labeled at their 5' ends using T4 polynucleotide kinase and [γ-^{32}P] ATP. Unincorporated ATP was removed by passing the kinase reaction mix through a G-25 Sephadex quick spin column (Roche).

2.5.6 *In vitro* RNase activity assay

A typical *in vitro* RNase activity assay was carried out in a 30 μL reaction mixture containing 20 mM Tris-HCl pH8.0, 300 mM KCl, 0.25 mM $MgCl_2$, 1 mM DTT, 5 μM RNA substrate, and 0.2 μM of the designated RNase. The RNase was added last to the reaction mixture to initiate the degradation reaction. The reaction mixtures were incubated at 37°C, aliquots were taken from the mixture at indicated time points and the reaction was terminated by addition of 1 volume

of gel loading buffer containing 50% formamide, 15% formaldehyde, 10% glycerol, 0.025% bromophenol blue, and 0.025% xylene cyanol. Degradation products were resolved on 20% denaturing polyacrylamide gel containing 7.5 M urea. Conversion of the substrate to products was visualized by autoradiography.

2.5.7 Stalled ribosome enrichment

Reporter mRNAs from either pλ-cI-N stop or pλ-cI-N nonstop were over-expressed in either MG1655 *rnr*::*KAN* or W3110 *smpB*$^{-}$*ssrA*$^{-}$ *rnr*::*KAN* complemented with designated plasmid borne RNase R, under a pBAD promoter. Cells were grown in media containing 0.02% arabinose to OD_{600} of 1.0 and the expression of reporter constructs were induced by the addition of IPTG to a final concentration of 1 mM. Cells were allowed to grow for another 45min before harvesting by centrifugation at 3,700 x g for 1hr. The resulting cell pellets were resuspended in Buffer E (50 mM Tris-HCl pH 7.5, 300 mM NH_4Cl, 20 mM $MgCl_2$, 2 mM β-ME, and 10 mM Imidazole) and lysed by French Press. Crude ribosomes were pelleted by ultracentrifugation at 30,000 x g for 16-18hrs through a 32% sucrose cushion in Buffer E. Ribosomes translating the reporter mRNAs were purified from crude ribosomes by Ni^{2+}-NTA affinity chromatography. The ribosome-associated RNase R was detected by western blot analysis using polyclonal anti-RNase R antibody. The amount of RNase R associated with ribosomes translating the stop or non-stop reported mRNA was then quantified using Image J. The amount of RNase R (WT of RNR^{723}) associated with ribosomes translating the stop mRNA were then set as 1.0 (background association) and the fold increase of RNase R translating the non-stop mRNA was calculated according to their respective background.

2.6 Figures

Figure 2.1

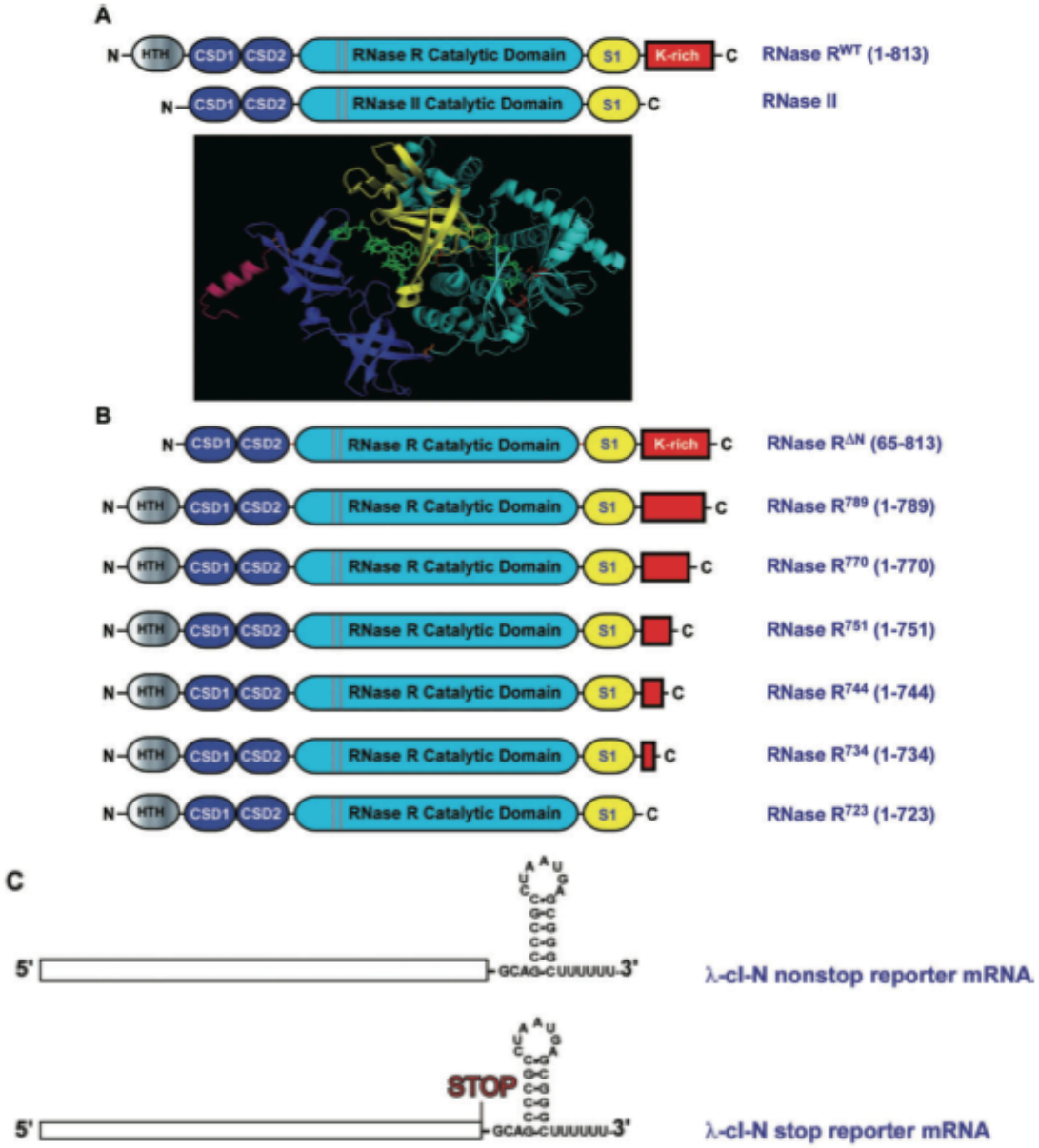


Figure 2.1: A) Schematic representation of the domain architecture of RNase R and RNase II. RNase R and RNase II share extensive similarity in the N-terminal cold-shock, central nuclease and C-terminal S1 domains. RNase R has two additional domains, a N-terminal putative helix-turn-helix (HTH) domain and a C-terminal lysine-rich (K-rich) domain. B) RNase R truncation variants used in this study, missing either the N-terminal HTH domain or various lengths of the C-terminal K-rich domain. C) Schematic representation of the *λ-cI-N* nonstop and *λ-cI-N* stop reporters. The *λ-cI-N* coding region is represented as a rectangle and the nucleotide sequence of the *trp*A terminator, located at the 3'-end of the transcript, is shown.

Figure 2.2

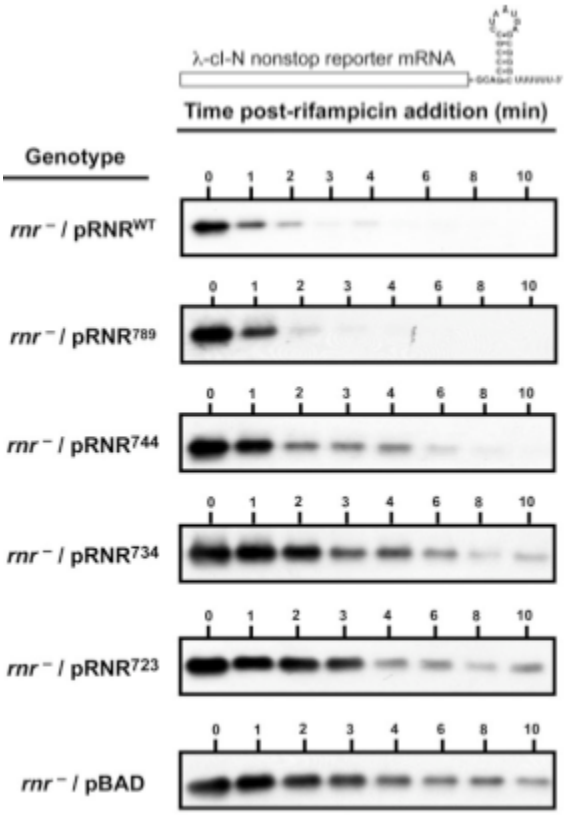


Figure 2.2: Decay rates for *λ-cI-N* nonstop reporter mRNA as estimated from rifampicin chase experiments**.** Aliquots of cells expressing the *λ-cI-N* nonstop and RNase R, or one of its C-terminal truncation variants, were taken at indicated time points post rifampicin addition and used for RNA purification. Total RNA isolated from equal number of cells was used for northern blots, which were hybridized with the *λ-cI-N* specific probe. Bands on the blot were quantified by scanning and subsequent analysis with Image-J (http://rsbweb.nih.gov/ij/). mRNA half-lives were calculated by linear regression analysis of data obtained from at least three independent experiments. pACYCDuet-1 was used as the vector control.

Figure 2.3

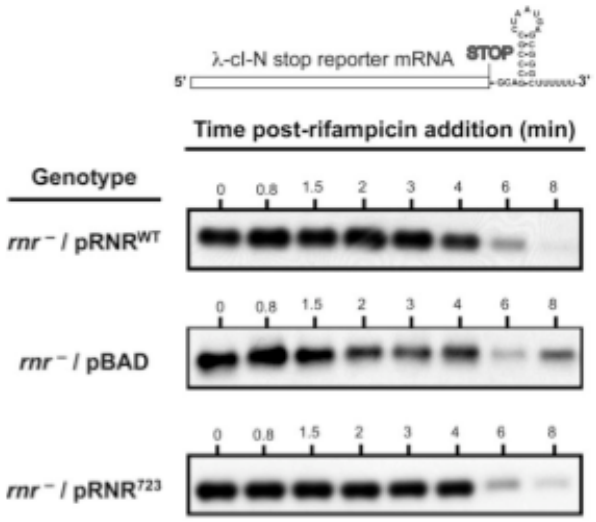


Figure 2.3: Decay rates for *λ-cI-N* stop reporter mRNA as estimated from rifampicin chase experiments. Aliquots of cells expressing RNase R or RNR^{723} and *λ-cI-N* stop were taken at indicated times post rifampicin addition and used for RNA purification. Total RNA isolated from equal number of cells was used for northern blots, which were hybridized with a ***λ-cI-N*** specific probe. Bands on the blot were quantified by scanning and subsequent analysis with Image-J (http://rsbweb.nih.gov/ij/). mRNA half-lives were calculated by linear regression analysis of data obtained from at least three independent experiments. pACYCDuet-1 was used as the vector control.

Figure 2.4

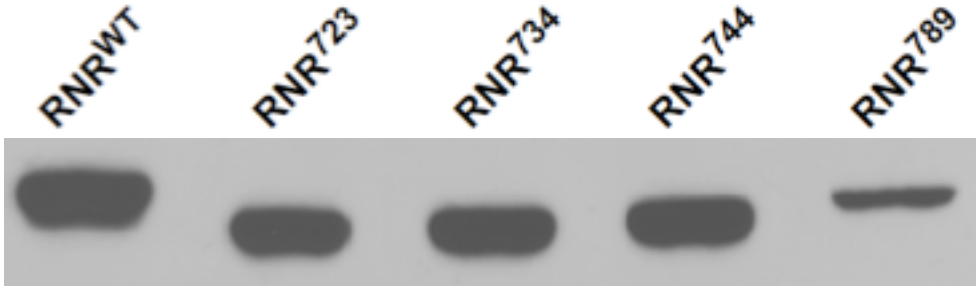


Figure 2.4: Evaluation of the steady state levels of wild type RNase R and its truncation variants. The intracellular levels of RNase R and its variants was assessed in strains and under conditions used for degradation and enrichment assays. Briefly, cells were grown in the presence of 0.01% arabinose to an OD_{600} of 0.6. Expression of the reporter was induced with 1mM IPTG for 1hr. Equal number of cells were lysed in SDS sample buffer, boiled for 10 min and subjected to electrophoresis on a denaturing polyacrylamide gel. The resolved proteins were transferred to a PVDF membrane and examined by western blot analysis using RNase R specific antibodies. The only major difference observed was in the intracellular levels of 789 and WT, which nevertheless show similar rates of degradation. All the other variants show comparable steady state levels of protein in the cell.

Figure 2.5

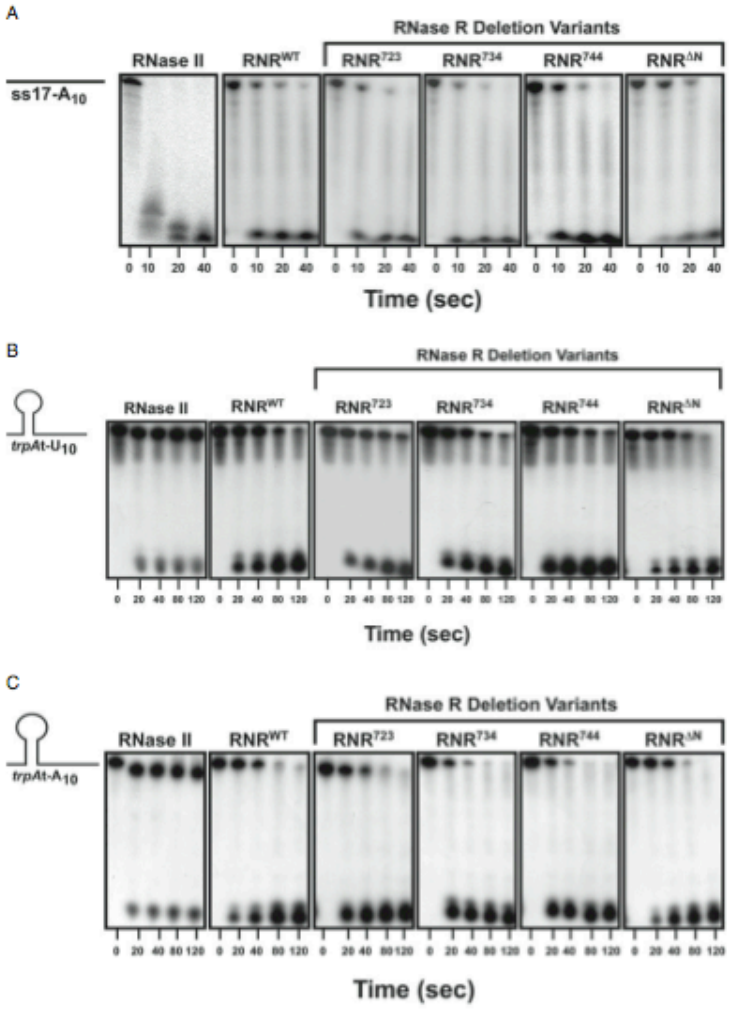


Figure 2.5: A) Analysis of the ability of RNase II, RNase R, and RNase R truncation variants to degrade a single stranded RNA substrate (ss32) *in vitro*. The ^{32}P-labeled ss32 substrate was used to examine the degradation capacity of these RNases *in vitro*. Samples from each degradation reaction were taken at designated time points and resolved by electrophoresis on a denaturing polyacrylamide gel and visualized by autoradiography. B) The ability of RNase II, RNase R, and RNase R truncation variants to degrade a RNA substrate with an internal stem-loop structure followed by a 3' poly-U overhang. The experiment was performed as described in Figure 4A. [C]. The ability of RNase II and RNase R, and RNase R truncation variants to degrade a RNA substrate with an internal stem-loop structure followed by a 3' poly-A overhang.

Figure 2.6

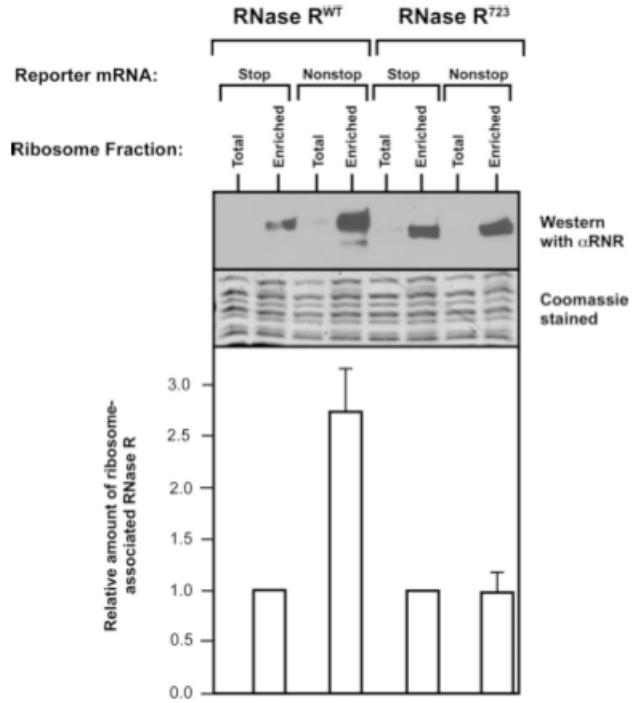


Figure 2.6: RNase R is selectively enriched in ribosome fractions translating a *λ-cI-N* nonstop mRNA and the enrichment is dependent on the presence of the C-terminal K-rich domain of RNase R. Ribosomes translating reporter mRNAs, either the normal *λ-cI-N* stop or a *λ-cI-N* nonstop mRNA, were isolated from the total cellular ribosome pool, utilizing the N-terminal His_6 tag encoded on the reporter protein. Protein components of the isolated ribosomes were resolved by electrophoresis on a 10% SDS-polyacrylamide gel and the presence of RNase R, or the RNase R^{723} truncation variant, was detected by western blot analysis using RNase R specific antiserum.

Figure 2.7

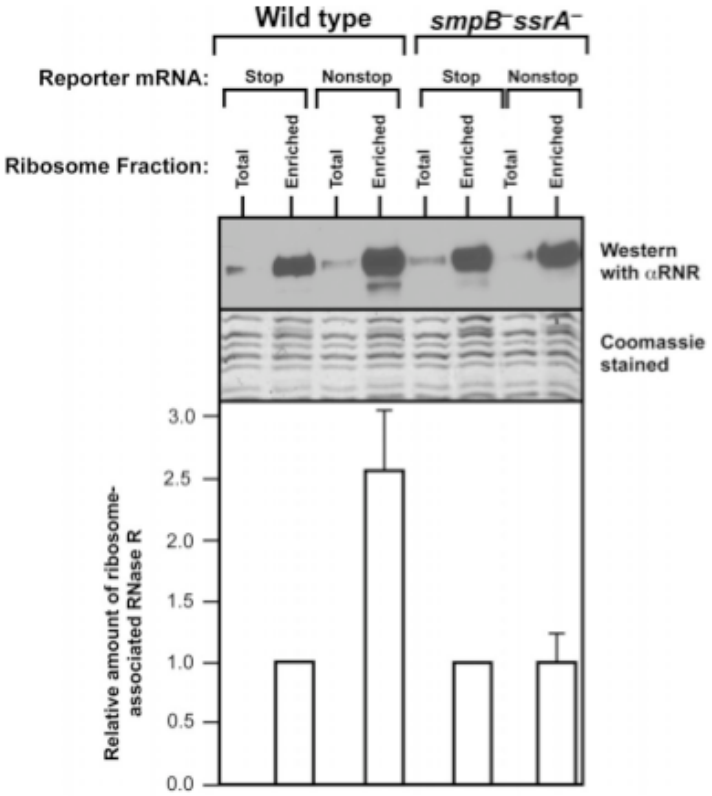


Figure 2.7: The recruitment of RNase R to stalled ribosomes is dependent on the presence of tmRNA and SmpB protein. Reporter mRNAs, *λ-cI-N* stop or a *λ-cI-N* nonstop, were expressed in either *ssrA*⁻*smpB*⁻ strain or its isogenic parental strain. The isolation of ribosomes translating the reporter mRNAs and the detection of RNase R in these ribosomes were performed as described in the Figure 2.5.

Table 2.1

RNase R expression conctruct	Substrate	*In vivo* $t_{1/2}$ (min)
pRNR^{WT}	**λ-cI-N non-stop**	**0.8±0.1**
pRNR^{789}	**λ-cI-N non-stop**	**1.0±0.1**
pRNR^{744}	**λ-cI-N non-stop**	**1.7±0.2**
pRNR^{734}	**λ-cI-N non-stop**	**2.5±0.3**
pRNR^{723}	**λ-cI-N non-stop**	**2.7±0.3**
pBAD	**λ-cI-N non-stop**	**4.5±0.6**
pRNR^{WT}	**λ-cI-N stop**	**4.6±0.1**
pRNR^{723}	**λ-cI-N stop**	**4.9±0.1**
pBAD	**λ-cI-N stop**	**4.8±0.6**

Table 2.1: *In vivo* mRNA decay activity of K-rich domain deletion variants of RNase R.

Chapter 3 Co-evolution of multipartite interactions between an extended tmRNA tag and a robust Lon protease in *Mycoplasma*

*** The work contained within this chapter was published in Molecular Microbiology (2009) 74(5).**

3.1 Summary

Messenger RNAs that lack in-frame stop codons promote ribosome stalling and accumulation of aberrant and potentially harmful polypeptides. The SmpBtmRNA quality control system has evolved to solve problems associated with nonstop mRNAs, by recognizing and rescuing stalled ribosomes and directing the addition of a degradation tag to the C-termini of the associated polypeptides, marking them for directed proteolysis. In *E. coli*, tmRNA tagged proteins are degraded primarily by the ClpXP protease system. We have shown that the AAA+ Lon protease also participates in degradation of tmRNA tagged proteins, but with much lower affinity. Here, we present a unique case of enhanced recognition of an extended *Mycoplasma pneumoniae (Mp)* tmRNA tag by the *Mp*-Lon protease. We demonstrate that *Mp*-Lon can efficiently and selectively degrade *Mp*-tmRNA tagged proteins. Most significantly, our studies reveal that the larger (27 amino acid long) *Mp*-tmRNA tag contains multiple discrete signaling motifs for efficient recognition and rapid degradation by Lon. We propose that higher affinity multipartite interactions between *Mp*-Lon and the extended *Mp*-tmRNA tag have co-evolved from pre-existing weaker interactions, as exhibited by Lon in *E. coli*, to better fulfill the function of *Mp*-Lon as the sole soluble cytoplasmic protease responsible for the degradation of tmRNA tagged proteins.

3.2 Introduction

Protein degradation has emerged as a key cellular mechanism for regulation of a diverse array of physiological processes. In prokaryotes, energy-dependent proteases play a major role in re-sculpting the bacterial proteome, in maintaining ideal concentrations of critical regulatory proteins, and in the disposal of unwanted, damaged or misfolded proteins. Bacterial energy-dependent proteases are grouped into four families, named after their representative members: ClpXP/ClpAP, HslUV (ClpYQ), FtsH (HflB), and Lon [164]. The Clp and HslUV proteases are two-component proteases consisting of a chaperone subunit, ClpA, ClpX, or HslU, and a peptidase subunit, ClpP or HslV. The ClpA, ClpX, and HslU chaperones are ATPases that are critical for substrate recognition, unfolding and translocation into the ClpP or HslV peptidase. The activities of the ATP-driven chaperone components help determine the substrate range and specificity for ClpAP, ClpXP, and HslUV proteases. Unlike the two-component proteases, the ATP-dependent Lon protease forms a hexameric ring that is derived from a single polypeptide, which carries both the chaperone and

peptidase functions. Each Lon monomer has three distinct domains: the amino-terminal domain that is implicated in substrate binding, the ATPase domain that contains the Walker A and B motifs for ATP binding and hydrolysis, and the peptidase domain located at the carboxyl terminus of the protein.

For the vast majority of protein, the sequence determinants recognized by energy dependent proteases appear to be present in the primary sequence of each individual substrate, which are either constitutively accessible for protease recognition or become available under defined cellular transitions or environmental conditions [106, 128, 164]. For one group of substrates, the tmRNA-tagged proteins, a defined protease recognition module is added to their C-termini. The tmRNA-mediated tagging and ribosome rescue system is the only known biological process that co-translationally appends a degradation tag to the C-termini of proteins to target them for directed proteolysis [149, 155, 156, 165, 166]. The idea of a degradative role for tmRNA function originated from the realization that the C-terminal residues of the peptide sequence encoded by the mRNA-like domain of tmRNA are similar to recognition determinants of intracellular proteases [167-170]. Subsequent studies showed that energy-dependent proteases are important contributors to this process [171, 172]. Early work on the proteolytic function of the tmRNA-mediated *trans*-translation process had shown that ClpXP, ClpAP, and FtsH were involved in disposal of tmRNA-tagged proteins in a tag-specific manner [171, 172]. In *E. coli*, ClpXP has persisted as the principal protease responsible for degradation of tmRNA-tagged proteins *in vivo*. The inner membrane-bound protease FtsH has narrower specificity against tmRNA-tagged proteins than ClpXP and ClpAP, and is active mainly on unstable [173] and locally available substrates [174, 175]. Recent work from our lab has demonstrated a role for the Lon ATP-dependent protease in degradation of tmRNA-tagged proteins [88]. Although Lon contributes more to the *in vivo* degradation of tmRNA tagged proteins than ClpAP and FtsH, its contributions are still far less than those made by the ClpXP system [88]. These findings suggest that the *E. coli* ClpXP, in coordination with its SspB cofactor, has much higher affinity for tmRNA-tagged proteins than the *E. coli* Lon protease. However, this selective delivery and high affinity for targeted degradation of tmRNA-tagged proteins need not be true in all bacterial species. Indeed, recent surveys of protease homologs and orthologs in eubacteria have revealed that the ATP-dependent Lon protease is more strongly conserved than other bacterial energy-dependent proteases, including the Clp family of proteases (Tripathi *et al*., 2008). For instance, the small genome bacterial species *Mycoplasma* lack both the ClpAP and ClpXP protease systems, yet they possess a Lon protease homolog. In contrast to the variable conservation of bacterial energy-dependent proteases, the SmpB-tmRNA system is strictly conserved and is presumably universally utilized in eubacteria to co-translationally tag proteins for directed proteolysis. This presents an interesting quandary. If the *E. coli* tmRNA tag recognition model for ClpXP and Lon proteases is universally true, how are *Mycoplasma* tmRNA-tagged proteins targeted for efficient degradation?

Despite the significance of Lon in re-sculpting the bacterial proteome, there is a paucity of information about how Lon is targeted to specific substrates, how these substrates are recognized and degraded, or what sequence or structural features are important for recognition by Lon. We have been interested in understanding how Lon protease recognizes and degrades specific substrates. In particular, we have been interested by the potential link between the Lon protease and the *trans*-translation systems of *Mycoplasma* species (*Mycoplasma pneumoniae* and *Mycoplasma genitalium*). We have been especially intrigued by the expanded nature and sequence composition of the *Mycoplasma* tmRNA tag, as it is considerably longer (27 amino acid residues, compared to the 11 amino acid residue *E. coli* tmRNA tag) and contains an unusual preponderance of charged residues (Figure 1A). Thus, we thought it conceivable that the *Mycoplasma* tmRNA tag and the *Mycoplasma* Lon protease might have co-evolved to facilitate rapid and efficient recognition and degradation of tmRNA tagged proteins. In this report, we present evidence to demonstrate that the *Mycoplasma* Lon protease is a highly robust protease that has co-evolved with tmRNA to detect an expanded set of recognition elements within the *Mp*-tmRNA peptide tag sequence for efficient and rapid clearance of tagged proteins. Our findings also have inplications for how Lon recognizes both its normal and tagged substrates.

3.3 Results

3.3.1 Construction of a reporter protein harboring the *Mycoplasma pneumonia* tmRNA tag

To investigate the relationship between *Mycoplasma pneumoniae* (*Mp*) tmRNA-tagged proteins and Lon protease, we fused the *Mp*-tmRNA tag sequence to the 3´ end of the N-terminal domain of the λ-cI gene to generate a reporter substrate (λ-cI-N-ssrA$_{MP}$) for *Mp*-Lon. We cloned, expressed, and purified the resulting λ-cI-N-ssrA$_{MP}$ protein product. However, the purified λ-cI-N-ssrA$_{MP}$ protein resolved into two distinct bands in 15% tris-tricine gel, suggesting a mixture of two related proteins with a molecular weight difference of 1-2 kDa. In order to determine the identity of the two proteins, we performed MALDI-TOF mass spectrometry analysis. This analysis showed the higher molecular weight band (15102 Da) to be the full-length λ-cI-N-ssrA$_{MP}$ reporter and the lower molecular weight band (13893 Da) to be a truncated reporter protein carrying the N-terminal part of the *Mp*-tmRNA tag but missing the C-terminal 11 residues (Figure 3.1). These results suggested that part of the *Mp*-tmRNA tag sequence served as a recognition site for one or more *E. coli* peptidases, leading to the cleavage of the C-terminal 11 amino acids of the tag. Analysis of this reporter in ClpXP and Lon deficient strains suggested that these proteases were not responsible for generating the shorter product. The cleavage product is henceforth referred to as λ-cI-N-ssrA$_{MP-11}$.

3.3.2 Purified *Mp*-Lon degrades *Mp*-tmRNA tagged proteins selectively and with high efficiency

The C-terminal end of the *E. coli* tmRNA tag (*Ec*-tag) is the most information rich part of the tag, with binding elements for ClpXP, ClpAP and Lon proteases [88, 91, 100, 156, 165, 171, 176-179]. Since our purified λ-cI-N-ssrA$_{MP}$ reporter protein contained a mixture of both full-length λ-cI-N-ssrA$_{MP}$ and truncated λ-cI-N-ssrA$_{MP-11}$, we reasoned that perhaps the shorter reporter might serve as an internal control for our *in vitro* proteolysis assay, as it might be missing some key C-terminal Lon recognition determinants. We, therefore, tested the ability of purified *Mp*-Lon to degrade the *Mp*-tmRNA tagged reporter proteins. As shown in Figure 3.2C, *Mp*-Lon was fully capable of rapidly and selectively degrading full-length λ-cI-N-ssrA$_{MP}$. The *Mp*-Lon mediated degradation process was dependent on the presence of two key elements. First, the reaction was absolutely dependent on the presence of ATP, suggesting that ATP bonding and hydrolysis was required for substrate binding and unfolding. Second, reporter protein degradation was dependent on the presence of full-length *Mp*-tmRNA tag. This assertion was confirmed by the fact that *Mp*-Lon was unable to degrade either the truncated λ-cI-N-ssrA$_{MP-11}$ (internal control) or the untagged λ-cI-N reporter protein (Figure 3.2). Thus, the addition of *Mp*-tmRNA tag converted non-substrate proteins to efficient substrates for degradation by *Mp*-Lon protease. In addition, the fact that *Mp*-Lon was unable to efficiently degrade the truncated λ-cI-N-ssrA$_{MP-11}$ reporter suggested that the C-terminal 11 amino acids of the *Mp*-tmRNA tag contain key elements for recognition and degradation by *Mp*-Lon. Furthermore, these experiments provided a convenient internal control (λ-cI-N-ssrA$_{MP-11}$) for future proteolysis assays.

3.3.3 *Mp*-Lon is a more effiencient protease

Despite extensive length and amino acid sequence differences between the *Ec*- and *Mp*-tmRNA tags, they still share some common attributes. For instance, the C-termini of both tags end with a mixture of aromatic and small hydrophobic amino acids (Figure 3.2A). Therefore, we endeavored to determine whether the *Mp*-Lon protease could recognize *Ec*-tmRNA tag, and whether the *Ec*-Lon protease could recognize *Mp*-tmRNA tag. To this end, we expressed and purified *Mp*-Lon, *Ec*-Lon, and two tagged reporter proteins: one carrying the *Ec*-tmRNA tag (λ-cI-N-ssrA$_{EC}$) and the other carrying the *Mp*-tmRNA tag (λ-cI-N-ssrA$_{MP}$). In *in vitro* proteolysis assays, using 100nM *Mp*-Lon hexamer, we observed rapid and efficient degradation of λ-cI-N-ssrA$_{EC}$ (Figure 3.2D, top panel), suggesting that *Ec*-tmRNA tag, although lacking many sequence features of *Mp*-tmRNA tag, was still capable of targeting substrates to proteolysis by *Mp*-Lon. In a similar analysis of the *Ec*-Lon protease, however, we noticed a compelling difference between the two proteases. When we used an equivalent

concentration (100nM hexamer) of *Ec*-Lon protease and a reporter protein with *Mp*-tmRNA tag (λ-cI-N-ssrA$_{MP}$), we observed only modest degradation of the tagged protein. However, when we increased the concentration of *Ec*-Lon to 400nM hexamer, efficient degradation of λ-cI-N-ssrA$_{MP}$ was observed (Figure 3.2D, lower panel). These data indicated that *Ec*-Lon has lower affinity for the *Mp*-tmRNA tag. The observed lower affinity of *Ec*-Lon was not due to loss of enzymatic activity during purification or storage. We observed similar proteolytic activities with several independent preparations of both *Mp*-Lon and *Ec*-Lon proteases, and both enzymes were equalily active against denatured Casein. These results suggested that in the absence of the Clp system the *Mycoplasma* Lon protease has evolved to become a much more efficient protease capable of specific recognition and selective degradation of *Mp*-tmRNA tagged proteins.

3.3.4 *Mp*-Lon recognizes multiple sequence motifs within the expanded *Mp*-tmRNA tag

Having established that *Mp*-Lon was active and fully capable of selectively recognizing and degrading *Mp*-tmRNA tagged proteins with high efficiency, we sought to gain insights into the extended nature of the *Mp*-tmRNA tag. We reasoned that perhaps in the absence of *clp* genes in *Mycoplasma*, the extended *Mp*-tmRNA tag sequence had evolved to provide enhanced recognition of tagged proteins by the *Mp*-Lon protease. Such an enhancing effect on Lon recognition could be achieved through expansion of tag sequence elements for direct Lon recognition, acquisition of mediatory sequences for cofactor binding, or a combination of both. To date, a protein cofactor for Lon protease has not been identified. Additionally, it is thought that evolutionary adaptation has resulted in massive gene loss in *Mycoplasma*. Therefore, it is less likely that a new cofactor that facilitates substrate delivery to Lon would have emerged. We, thus, hypothesized that in addition to the C-terminal sequence elements, the expanded *Mp*-tmRNA tag may have evolved added recognition elements for efficient recognition by *Mp*-Lon. To directly test the presence of any additional recognition determinants, we divided the 27-amino-acid long tmRNA tag into 9 regions and substituted, where possible, the native tag residues with aspartic acids (Figure 3.3A). The reporter proteins, carrying *Mp*-tmRNA tag variants with aspartic acids subtitutions at each of the 9 regions, were cloned, expressed, and purified. An interesting and immediate outcome of this analysis was that upon substitution of aspartic acid residues in regions 5-9 we observed a major change in the pattern of the purified reporter proteins (Figure 3.3B). More specifically, we observed a total absence of the shorter λ-cI-N-ssrA$_{MP-11}$ tagged protein, suggesting that the *E. coli* peptidase responsible for this cleavage was sensitive to introduction of aspartic acid residue in this region of *Mp*-tmRNA tag.

To ensure that all proteolysis assays had the same ratio of full-length λ-cI-N-ssrA$_{MP}$ reporter and the degradation resistant λ-cI-N-ssrA$_{MP-11}$ internal control, we constructed a reporter gene variant, by introducing a stop codon after codon

15 of the tag, which produced a reporter control that was identical to λ-cI-N-ssrA$_{MP-11}$ (Figure 3.2B). We expressed and purified the λ-cI-N-ssrA$_{MP-11}$ control protein and added it to each purified λ-cI-N-ssrA$_{MP}$ tag variant at a concentration that yielded equimolar ratios of the two proteins (at 1:1 molar ratio of each full-length λ-cI-N-ssrA$_{MP}$ tag variant and the λ-cI-N-ssrA$_{MP-11}$ control). We assessed the effect of aspartic acid substitutions on *Mp*-Lon recognition of *Mp*-tmRNA tagged reporters by subjecting each λ-cI-N-ssrA$_{MP}$ variant to *in vitro* proteolysis (Figure 3.3B). In order to better illustrate the effect of aspartic acid substitutions in each of the nine regions of the *Mp*-tmRNA tag, we quantified the remaining amount of full-length λ-cI-N-ssrA$_{MP}$ at each time point and have presented the 15-minute time point in bar-graph format as an indicator of their degradation efficiency. This analysis demonstrated that aspartic acid substitutions in the most C-terminal segment of the *Mp*-tmRNA tag (Regions 8 and 9) were very effective in preventing directed degradation of the λ-cI-N-ssrA$_{MP}$ reporter protein by *Mp*-Lon (Figure 3.3B). These results are consistent with our earlier observation that *Mp*-Lon is unable to efficiently degrade the λ-cI-N-ssrA$_{MP-11}$ control protein, and provide strong support for the conclusion that some key *Mp*-Lon recognition elements are located with the last 11 amino acid residues of *Mp*-tmRNA tag. Aspartic acid substitutions in Region 6 and region 7, although located within the last 11 amino acids of the *Mp*-tag, produced only a modest effect on degradation of the λ-cI-N-ssrA$_{MP}$ reporter by *Mp*-Lon.

Most interestingly, aspartic acid substitutions in region 5 of the reporter rendered it highly resistant to proteolysis by *Mp*-Lon, despite the fact that this reporter variant had the full complement of the C-terminal recognition elements (Figure 3.3B). These results suggested that amino acid residues in region 5 of the *Mp*-tmRNA tag constitute an additional critical recognition element for binding and degradation by *Mp*-Lon. Furthermore, aspartic acid substitutions in region 4, located just upstream of region 5, resulted in close to 50% reduction in the degradation rate of the reporter substrate, suggesting an intermediate level of contribution by region in recognition by *Mp*-Lon. In contrast, aspartic acid substitutions in regions 1, 2, or 3 did not have any substantial effect on the ability of *Mp*-Lon to degrade the λ-cI-N-ssrA$_{MP}$ reporter protein (Figure 3.3B). Taken together, these results lend strong support to the conclusion that *Mp*-Lon requires multiple signaling elements in the *Mp*-tmRNA tag for efficient recognition and degradation of tagged proteins. Key signaling elements for this multipartite interaction appear to be located in discrete sites, including the extreme C-terminus (regions 8 and 9), and in the middle of the tag (regions 4 and 5).

3.3.5 Cumulative contributions of key signal residues to efficient recognition of *Mp*-tmRNA tag

Having established that discrete *Mp*-tmRNA tag sequence elements were required for Lon recognition, we set out to assess the relative contributions of individual amino acids in each of the key signaling sites to recognition by *Mp*-

Lon. To this end, we made single aspartic acid substitutions in regions 5, 8 and 9 to confirm their importance to recognition by *Mp*-Lon (Figure 3.4A). We also made single aspartic acid substitutions in a control region (region 7) to reaffirm that it does not contain significant Lon recognition elements. We chose to target hydrophobic amino acid residues in each region, as it has been reported that *E. coli* Lon prefers these residues in some of its proteolytic substrates [99]. We targeted residues L14, I15, A16 in region 5, I22 in region 7, Y24 in region 8, and F26 in region 9. Site directed mutagenesis was used to introduce the desired substitutions. Each λ-cI-N-ssrA$_{MP}$ reporter protein variant was expressed and purified to near homogeneity. We assessed the effect of each individual tag variant on recognition and degradation by *Mp*-Lon in an *in vitro* proteolysis assay (Figure 3.4A). Quantification of the reporter substrate variants showed a range of modest effects on degradation by *Mp*-Lon but none of the single Asp substitution resulted in complete inhibition of *Mp*-Lon activity. The two aromatic residues in regions 8 and 9 (Y24 and F26) made larger contributions than hydrophobic residues in region 5. Among the residues in region 5, the relatively bigger hydrophobic residues (L14 and I15) contributed more than the smaller A16. Amino acid I22 in region 7, despite being hydrophobic, had no significant effect on substrate degradation by *Mp*-Lon (Figure 3.4A). This finding was consistent with the observation that region 7 was not a primary recognition determinant. These results support the conclusion that *Mp*-Lon recognizes multiple sequence elements within the extended tmRNA tag, and that individual residues within these region make small but important contribution to recognition by *Mp*-Lon. Therefore, the additive input of the individual signal elements contributes to the overall recognition of *Mp*-tmRNA tag by *Mp*-Lon.

The tmRNA tag was initially identified as a potential proteolysis signal based on similarity of its last 5 amino acids (YALAA) with recognition signal of C-terminal specific proteases [149]. Mutational analysis of the tmRNA tag in a number of bacterial species, including *E. coli* and *Yersinia pseudotuberculosis* [43, 55, 60, 91, 149, 171, 180], has confirmed the importance of these residues for recognition by cellular proteases. For instance, the last three amino acid residues (LAA-COO$^-$) of the *Ec*-tmRNA tag are of particular importance for recognition by the ClpXP system. Conversion of the ultimate and penultimate alanines of the tag to aspartic acid residues (from LAA to LDD) renders tagged proteins highly resistant to proteolysis by ClpXP. The last four residues of the *Mp*-tmRNA tag (YAFA) have a similar mix of large hydrophobic and small amino acids (Figure 3.2A). We have already demonstrated that the two aromatic amino acids were required for recognition by *Mp*-Lon (Figure 3.3C). We have also shown that A16 in region 5 of the *Mp*-tmRNA tag made no substantial contributions to recognition by *Mp*-Lon (Figure 3.4A). We were curious to know why alanines were conserved at the C-terminus of the tag and whether they also served as signals for recognition by *Mp*-Lon. To assess the importance of these residues, we constructed reporter proteins where the last four amino acids of the tag were independently changed from YAFA-COO$^-$ to either –DADA-COO$^-$ or –YDFD-COO$^-$ and evaluated the degradation of these substrates by *Mp*-Lon

(Figure 3.4B). Interestingly, both tag variants were highly resistant to proteolysis, suggesting that the two alanines at the C-terminus of the tag were as critical to recognition *Mp*-Lon as the two aromatic residues.

3.3.6 *Ec*-Lon is also sensitive to aspartic acid substitutions in multiple sequence elements within the expanded *Mp*-tmRNA tag

We have shown that *Ec*-Lon is capable of recognizing and degrading *Mp*-tmRNA tagged proteins (Figure 3.2D). However, efficient degradation requires higher concentrations of *Ec*-Lon, suggesting weaker affinity for proteins carrying *Mp*-tmRNA tagged. Gur and Sauer [99] have recently shown that *Ec*-Lon prefers large hydrophobic residues in its binding site and that *Ec*-Lon can unfold and degrade stably folded proteins provided they have accessible recognition signals. Since the last four amino acid residues of *Mp*-tmRNA tag (-YAFA) constitute an ideal binding site for *Ec*-Lon, we wondered whether the presence of these residues was necessary and sufficient for degradation of a substrate carrying *Mp*-tmRNA tag or whether *Ec*-Lon was also sensitive to changes in other parts of the *Mp*-tmRNA tag.

To address these questions, we evaluated the ability of *Ec*-Lon to degrade single aspartic acid substitution variants of the λ-cI-N-ssrA$_{MP}$ reporter protein (Figure 3.4C). We found that *Ec*-Lon was not affected by aspartic acid substitutions in regions 1 or 2, and was only moderately sensitive to substitutions is region 3 of the tmRNA tag. However, *Ec*-Lon was highly sensitive to single aspartic acid substitutions in regions 5 (L14D and I15D) and in region 8-9 (Y24D, and F26D) of the *Mp*-tag. Taken together, these data suggested that the low affinity interactions between the two were distributed over multiple regions of the tag peptide and any disrupting substitutions affected the propensity of *Ec*-Lon to bind and degrade the target substrate. These results also suggested that the presence of large hydrophobic residues at the C-terminus of a substrate was necessary but not sufficient for degradation and that *Ec*-Lon made similar, alberit lower affinity, contacts with additional internal elements of the *Mp*-tmRNA tag.

3.3.7 Steady-state kinetic analysis of tagged protein degradation by *Mp*-Lon and *Ec*-Lon

The finding that *Mp*-Lon and *Ec*-Lon proteases degraded tmRNA tagged proteins to different extents suggested that these proteases had different affinities for tagged proteins. Similarly, the fact that aspartic acid substitutions in multiple key regions of the *Mp*-tag affected degradation of the reporter protein by both Lon proteases suggested that these alterations had lowered the affinity of these enzymes for tagged substrates. To gain further insights into these enzymatic processes we carried out steady-state kinetic analysis of the degradation of tagged proteins by the *Mp*-Lon and *Ec*-Lon proteases. However,

to perform the steady-state analysis we needed to purify [^{35}S]-labeled λ-cI-NssrA$_{MP}$ reporter protein from a strain that did not produce the truncated λ-cI-NssrA$_{MP-11}$ product. Although this truncated reporter had served as an extremely useful internal control in our initial proteolysis assays, its presence as a [^{35}S]-labeled protein would adversely affect the accuracy of our kinetic measurements. Therefore, we needed to identify cellular proteases responsible for the cleavage activity. Since we expressed the reporter protein in a *clpX*$^-$*clpP*$^-$*lon*$^-$ strain, we knew that the cleavage activity was not due to Clp or Lon proteases. In subsequent studies we determined that DegP and Tsp were responsible for the cleavage activity and that the truncated reporter product was not generated in *degP*$^-$*prc*$^-$ strains (Figure 3.5). The identification of the proteases responsible for production of the truncated reporter protein enabled us to use the *degP*$^-$*prc*$^-$ strain for the expression and purification of homogeneous full-length λ-cI-NssrA$_{MP}$ reporter proteins needed for our steady-state kinetics analysis.

To determine directly whether *Mp*-Lon and *Ec*-Lon had different affinities for tagged reporter proteins, we purified full-length [^{35}S]-labeled λ-cI-N-ssrA$_{MP}$ reporter protein, and its aspartic acid substitution variants, and measured the rates at which proteolysis by Lon generated acid-soluble radioactive peptides. We determined steady-state rates of substrate degradation by *Mp*-Lon and *Ec*-Lon at a range of reporter protein concentrations and fitted the data to the Michaelis-Menten equation (Figure 3.6). This analysis showed *Mp*-Lon to have high affinity, as reflected by the Michaelis constant (K_M), for the λ-cI-N-ssrA$_{MP}$ reporter, with a K_M value of 0.50 ± 0.04 μM and V_{max} of 5.1± 0.1 min^{-1} Lon$_6^{-1}$ (Figure 3.6, and Table I). *Ec*-Lon exhibited greater than 30-fold lower affinity for the same λ-cI-NssrA$_{MP}$ substrate, with K_M value of 17.4 ± 2.0 μM and V_{max} of 6.4± 0.2 min^{-1} Lon$_6^{-1}$ (Figure 3.6B). A similar analysis of the degradation of a substrate carrying the *Ec*- tmRNA tag, λ-cI-N-ssrA$_{EC}$, showed *Mp*-Lon to have close to 10-fold lower affinity for this substrate as compared to the λ-cI-N-ssrA$_{MP}$ reporter, with a K_M value of 4.9 ± 0.4 μM and V_{max} of 5.1± 0.1 min^{-1} Lon$_6^{-1}$ (Figure 3.6C, and Table I). These data were fully consistent with our finding that *Mp*-Lon had higher affinity for *Mp*-tmRNA tagged proteins and degraded these substrates with high selectivity and efficiency.

Next, we examined the effect of aspartic acid substitutions in key Lon recognition determinants of the *Mp*-tmRNA tag on degradation by *Mp*-Lon. We chose three λ-cI-N-ssrA$_{MP}$ aspartic acid substitution variants for this analysis: two with alterations in regions of the tmRNA tag sequence (region 5 and region 8) that constituted major determinants for Lon recognition, and one with alterations in a region of the tmRNA tag (region 6) that had only a small effect on degradation by Lon (Figure 3.7). Steady-state kinetic analysis showed that aspartic acid substitutions in region 8 of the *Mp*-tmRNA tag had the largest effect on the Michaelis constant, increasing K_M from 0.50 ± 0.04 μM to 29.8 ± 2.7 μM (Figure 3.7A, and Table I). The next largest effect on the Michaelis constant was seen with aspartic acid substitutions in the internal region 5 of the tmRNA tag, increasing K_M by 18-fold to 9.0 ± 0.6 μM (Figure 3.7B, and Table I). Aspartic acid

substitution in region 6 of the tmRNA tag had only a modest effect on the Michaelis constant, increasing K_M to 1.3 ± 0.1 μM (Figure 3.7C, and Table I). We also attempted to measure steady-state kinetic parameters for the degradation of region 5 variant by *Ec*-Lon. However, due to the dramatic affect of aspartic acid substitutions in this region on the rate of degradation by *Ec*-Lon, we were unable to obtain initial velocity data, even at very high (100 μM) substrate concentrations. These results suggested that aspartic acid substitutions in the internal segment of the *Mp*-tmRNA tag had lowered the already weak affinity of *Ec*-Lon for the tagged substrate to such an extent that made the tagged protein virtually unrecognizable by the *E. coli* protease. Taken together, these data are consistent with the conclusion that *Mp*-Lon recognizes multiple sequence elements in the extended *MP*-tmRNA tag and that key recognition signals, which play a critical role in substrate binding and degradation by Lon, are not limited to the extreme C-terminus of the tag peptide. Furthermore, *Ec*-Lon also recognizes the same key elements, albeit with much lower affinity, and is dramatically impacted by substitutions that alter these determinants.

3.3.8 *Mp*-Lon can partially complement *E. coli lon*⁻ phenotypes

Next, we wished to ascertain whether *Mp*-Lon was also capable of recognizing other cellular substrates of *Ec*-Lon. The basic rationale for this analysis was to determine whether *Mp*-Lon had retained the capacity to recognize well-known substrates of *Ec*-Lon, or has its substrate specificity evolved away from gene products not present in *Mycoplasma*. We chose two well-known substrates of *Ec*-Lon, SulA and RcsA, as they are not substrates of the tmRNA system and facile assays for their distinct *in vivo* activities are available. SulA is a cell division inhibitor induced by DNA damage that is rapidly degraded by Lon upon completion of DNA repair. Accumulation of SulA in *E. coli lon*⁻ renders cells sensitive to UV light and methylmethane sulfonate (MMS), a DNA damaging agent [181]. To evaluate SulA stability, we assessed the capability of *Mp*-Lon to complement the sensitivity of *E. coli lon*⁻ cells to MMS. First, we assayed the sensitivity of *E. coli lon*⁻ mutant HDB98 and its otherwise isogenic HDB97 parental strain to MMS. As shown in Figure 3.8A, the parental HDB97 strain was able to grow on plates irrespective of the presence or absence of MMS. In contrast, the *lon*⁻ HDB98 strain was unable to grow in the presence of MMS (Figure 3.8B). To complement the *lon*⁻ phenotype of the HDB98 *lon*⁻ strain, we provided either *Ec*-Lon or *Mp*-Lon function on an arabinose-inducible complementation plasmid. Introduction of a plasmid harboring the *Ec*-Lon gene, when induced with 0.01% arabinose, fully complemented the MMS sensitive phenotype of the *lon*⁻ strain (Figure 3.8B). A plasmid carrying the *Mp*-Lon gene provided partial complementation of the MMS phenotype, but only when induced to a higher level, i.e. with 0.2% arabinose (Figure 3.8C). These results suggested that *Mp*-Lon recognizes SulA weakly and degrades it at much lower efficiency, thus providing only partial complementation of the MMS sensitive phenotype.

Another *Ec*-Lon substrate that we evaluated was RcsA, a positive regulator of capsular polysaccharide biosynthesis genes [182]. *E. coli* strain HDB97 and its isogenic *lon*$^-$ mutant HDB98 contain a chromosomal *cpsB::lacZ* fusion insert [183], which affords a convenient assay for monitoring RcsA function via β-galactosidase expression and activity. As shown in Figure 3.9, we observed high levels of β-galactosidase activity in the Lon deficient HDB98 strain, suggesting RcsA protein accumulated in these cells due to the absence of Lon function. Introduction of a plasmid carrying *Ec*-Lon, when induced with 0.01% arabinose, resulted in full complementation of this phenotype and reduced the β-galactosidase activity to background levels (Figure 3.9). Expression of plasmid-borne *Mp*-Lon reduced β-galactosidase activity to approximately 25% of *lon*$^-$ cells. Higher-level expression, with 0.1% arabinose, of *Mp*-Lon did not result in improved complementation of the RscA phenotype (Figure 3.9). These results indicated that *Mp*-Lon recognized RcsA with much lower efficiency as compared to *Ec*-Lon. Furthermore, these data suggest that in the absence of SulA and RcsA in *Mycoplasma*, *Mp*-Lon binding affinity has evolved and is thus unable to efficiently bind and degrade these substrates.

3.4 Discussion

Timely and efficient removal of unwanted, damaged or unfolded proteins via proteolysis is vital for the cell. This is highlighted by the fact that interference with proteolysis severely impairs the cell's ability to survive under adverse environmental conditions. Signals that target proteins for proteolysis are either intrinsic to the primary structure of the substrate proteins or are extrinsically appended to target them for proteolysis [106, 128]. The SmpB-tmRNA quality control system, which rescues ribosomes stalled on defective mRNAs, is the only known biological process that co-translationally appends a peptide tag to proteins associated with stalled ribosomes [58, 155, 156, 166]. The small peptide tag, encoded by the mRNA-like domain of tmRNA, contains signals that directs tagged proteins to proteolysis by various cellular proteases [88, 149, 171, 172]. While the specific peptide sequence encoded by tmRNA varies among bacterial species, studies carried with *E. coli* tmRNA have confirmed the ability of the tmRNA peptide tag to promote targeted proteolysis by the Tsp, ClpXP, ClpAP, FtsH, and Lon proteases [88, 149, 171, 172]. The tmRNA mediated translation quality control system is thus intimately linked, by providing substrates for directed proteolysis, to protease-mediated protein quality control pathways.

The genes encoding SmpB and tmRNA are represented in all species of bacteria with available genomic sequence data. The preservation of SmpB and tmRNA likely reflects the extent of conservation and the evolutionary significance of this unique quality control system. Genome wide mutagenesis results suggest that genes encoding tmRNA and SmpB are essential in *Mycoplasma genitalium* and *Mycoplasma pneumoniae* [184]. Interestingly, of all the soluble cytoplasmic

AAA+ proteases that recognize tmRNA tagged proteins, only Lon protease is present in *Mycoplasma*, and is presumably essential for its survival [184, 185].

3.4.1 A case for co-evolution of the tmRNA-tag and Lon protease in *Mycoplasma*

In this study, we have demonstrated that the *Mycoplasma pneumoniae* Lon protease recognizes and degrades *Mp*-tmRNA-tagged proteins, efficiently and selectively. Most significantly, our data demonstrate that recognition of the tmRNA tag by *Mp*-Lon is not limited to the extreme C-terminal residues of the tag. Rather, multiple sequence elements within the extended *Mp*-tmRNA-tag provide critical signaling cues for recognition and selective proteolysis by *Mp*-Lon. Our findings thus support a model wherein *Mp*-Lon, as the sole soluble cytoplasmic AAA+ protease, has evolved to become a more robust and efficient enzyme for selective recognition and proteolytic turnover of tmRNA-tagged proteins. We propose that the extended *Mycoplasma* tmRNA tag has correspondingly evolved to present multiple recognition signals for Lon, and perhaps other protease.

A recent study showed that proteins carrying tmRNA-tag from *Mesoplasma florum (Mf)* are efficiently recognized and degraded by the *Mf*-Lon protease [186]. Through substitution analysis of the C-terminal end of the *Mf*-tmRNA tag, the authors concluded that the two aromatic residues (Y and F of the -YAFA-COO^- motif) were responsible for recognition of tag peptides by *Mf*-Lon. Their analysis was limited to the C-terminal part of the *Mf*-tag, particularly to the two aromatic residues, yet their conclusion agrees, in part, with our findings. Our analysis of the equivalent region of *Mp*-tmRNA tag demonstrates that all four residues of the -YAFA-COO^- motif, including the two alanines, are essential for recognition by *Mp*-Lon. Recognition of the C-terminal proximal alanine residues by *Mp*-Lon is thus reminiscent of the role played by the ultimate and penultimate alanines of the *E. coli* tag in recognition by the ClpXP system. The issue of the length of *Mycoplasma* tmRNA tag and its utility, i.e. whether additional recognition signals are present within the extended *Mf*-tag, were not directly addressed in the Gur and Sauer report [186]. Our analysis of the entire *Mp*-tag clearly demonstrates that additional distinct parts of the tag (Figure 3.2, 3.3 and 3.4) harbor signals that are critical for recognition by *Mp*-Lon, as changes in these determinants render tagged substrates highly resistant to proteolysis despite the presence of the intact C-terminal -YAFA-COO^- motif. These results suggest that interactions between Lon and tmRNA tag are multipartite in nature, and that the tag-recognition pocket of *Mp*-Lon is likely larger, encompassing both the C-terminal and internal signal elements of the tag. Therefore, these findings provide a more plausible explanation for the extended nature of the *Mycoplasma* tmRNA tag and hint at co-evolution of the tag sequence and Lon, the sole cytoplasmic AAA+ *Mycoplasma*. This conclusion is further supported by our analysis of the recognition by *Ec*-Lon of a protein carrying the *Mp*-tag. As noted

earlier, the *Mp*-tmRNA tag has large hydrophobic residues at its C-terminus (Figure 3.2A), which should make it an ideal substrate for recognition by *Ec*-Lon. Indeed, *Ec*-Lon does recognize and degrade an *Mp*-tagged substrate (Figure 3.2D). However, if the C-terminal –YAFA-COO^- determinant constituted the sole recognition signal for *Ec*-Lon then it should not have been affected by aspartic acid substitutions in regions 4 and 5 of the *Mp*-tag. The fact that *Ec*-Lon is affected by distal aspartic acid substitutions, despite the presence of the C-terminal recognition motif, suggests that the tag recognition pocket of *Ec*-Lon is correspondingly larger and encompasses more than the last four amino acids of the *Mp*-tag. We propose that *Ec*-Lon also recognizes additional parts of the *Mp*-tag. We submit that the reason for the substantial affects of aspartic acid substitutions in distal part of the *Mp*-tag is that *Ec*-Lon interactions with tmRNA tag are comparatively weaker and therefore more sensitive to disrupting alterations. Indeed, this conclusion is supported by our analysis of the steady-state parameters of the degradation of tmRNA tagged proteins by *Ec*-Lon and *Mp*-Lon (Figure 3.7 and Table 3.1).

The new picture that emerges from our analysis suggests that the weak interactions observed between *Ec*-Lon and *Ec*-tmRNA have co-evolved in *Mycoplasma* such that the *Mp*-tmRNA tag has become longer, to present more contact sites for an existing recognition pocket in the substrate binding domain of Lon, and *Mp*-Lon has co-evolved to recognize these signals better, through optimization of pre-existing low affinity sites (Figure 3.10A). Although we favor this simpler model, we cannot rule out the alternative possibility that Lon only recognizes the C-terminal –YAFA-COO^- motif, but binding of the –MLIA– signaling motif to a distal allosteric-site modulates its activity (Figure 3.10B). We also wish to make clear that the aforementioned analysis does not exclude the possible acquisition of new contact sites or an expansion of the binding pocket in *Mp*-Lon.

Although we have identified distinct regions of the *Mp*-tmRNA tag that are critical for Lon recognition, there are approximately 10 amino acids at the N-terminus of the tag that are not directly involved in recognition by Lon. It is plausible that these sequences serve as the binding sites for auxiliary cofactors in a manner analogous to SspB-ClpXP or ClpS-ClpAP interactions [102-104, 177, 179, 187-196]. To date, however, adaptor proteins have not been reported for Lon protease analogs. Nontheless, it is conceivable that Lon, like other AAA+ proteases, possesses a cofactor(s) that regulates its substrate range and specificity. Alternatively, the N-terminal part of the *Mp*-tmRNA tag might carry signals for recognition by the membrane-associated FtsH protease. Future studies are required to explore these possibilities.

3.5 Materials and Methods

3.5.1 Bacterial strains and plasmids

E. coli strains BL21 (DE3)/pLysS and CH1019 were used to express *Mp*-Lon and *Ec*-Lon proteases, respectively. All λ-cI-N reporter derivatives were expressed in the W3110 *clpX*$^-$*clpP*$^-$*lon*$^-$ protease deficient strain to minimize proteolysis or a periplasmic protease deficient strain *degP*$^-$*prc*$^-$*spr*$^-$ to prevent cleavage of the tag during purification. *In vivo* complementation assays were carried out in a *lon*$^-$ strain HDB98 or its otherwise isogenic HDB97 parental strain. The gene encoding *Ec*-Lon protease was cloned into pET21b vector. The gene encoding *Mp*-Lon protease was cloned into pET28b vector. An internal TGA codon in the *Mp-lon* gene, which encodes for tryptophan in *Mycoplasma*, was changed to TGG to prevent premature termination of translation during expression in *E. coli*. Both *Ec*-Lon and *Mp*-Lon genes were also cloned into pBAD18cm vector, respectively, to generate the complementation constructs for *in vivo* studies. The pPW500 plasmid, harboring the λ-cI-N-trpAt reporter construct, was modified to express the untagged λ-cI-N protein, the hard coded *E. coli* tmRNA tagged (λ-cI-N-ssrA$_{EC}$) protein, or the hard coded *M. pneumonia* tmRNA tagged λ-cI-N-ssrA$_{MP}$ protein. Reporter constructs expressing aspartic acids substitution tag variants of λ-cI-N-ssrA$_{MP}$ were generated by standard site directed mutagenesis, using *λ-cI-N-ssrA$_{MP}$* DNA as a template.

3.5.2 Protein Expression and Purification

To purify *Mp*-Lon, *E. coli* strain BL21(DE3)/pLysS harboring pET28b-*Mp*Lon-His6 was grown in 3 liters of LB supplemented with 100 μg/ml ampicillin and 30 μg/ml chloramphenicol. *Mp*-Lon expression was induced at OD_{600} ~ 0.7 by addition of 1 mM IPTG and continued for 3 hours at 30°C. Cells were harvested by centrifugation at 3,700 x g, and the resulting cell pellets were stored at -80 °C. The frozen cell pellets were resuspended in 100 ml of lysis buffer (50 mM $KHPO_4$ pH6.9, 1 mM EDTA, 1 mM DTT, and 10% glycerol) and lysed by sonication at 4˚C. Cellular debris was removed by centrifugation at 30,000 g for 2 hours. P11 cellulose phosphate resin (Whatman, GE Healthcare) was activated according to the manufacturer's instructions and pre-equilibrated in lysis buffer. The cleared cell lysate was loaded by gravity flow onto 5 grams of pre-equilibrated P11 resin and washed with 200 ml of lysis buffer. Bound proteins were eluted with 200ml of elution buffer (400 mM $kHPO_4$ pH6.9, 1 mM EDTA, 1 mM DTT, and 10% glycerol). The eluted proteins were concentrated down to 2ml and loaded onto an AKTA-FPLC Sephacryl S300 column (GE Healthcare), pre-equilibrated in Lon storage buffer (50 mM Tris-HCl pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT and 20% glycerol). Fractions containing purified *Mp*-Lon protease were pooled, concentrated, flash frozen in liquid nitrogen and stored at -80 °C. Purification of *Ec*-Lon, from *E. coli* strain CH1019 (DE3) harboring pET21b-*Ec*Lon, was carried out essentially as described above for *Mp*-Lon.

The λ-cI-N protein and it various tmRNA-tagged variants were expressed in *E. coli* strain W3110 *clpX*$^-$*clpP*$^-$*lon*$^-$. Expression of each reporter protein was induced at OD_{600} ~ 0.7 by addition of 1 mM IPTG and continued for 3 hours at 37°C. Cells were harvested by centrifugation at 3,700 x g, and the resulting cell pellets were stored at -80 °C. Frozen cell pellets were resuspended in a Ni^{2+}-NTA lysis buffer (50 mM KHPO4 pH7.0, 100 mM KCl, 1 mM EDTA, 2 mM β-ME, and 10 mM Imidazole). Cells were lysed by sonication at 4˚C and cell debris removed by centrifugation at 30,000 g for 2 hours. The cleared cell lysates were mixed with 2 ml of Ni^{2+}-NTA slurry equilibrated in the Ni^{2+}-NTA lysis buffer and rocked for 2 hours at 4 °C. The resin was washed with 200 ml of the lysis buffer and the bound protein was eluted with 30 ml of elution buffer (50 mM $KHPO_4$ pH7.0, 100 mM KCl, 1 mM EDTA, 20 mM β-ME, and 250 mM Imidazole). The eluted protein was concentrated to 5 ml, buffer exchanged into Q-Sepharose buffer A (50 mM Tris-HCl pH7.5, 50 mM KCl, 1 mM EDTA and 2 mM β-ME) and loaded onto a pre-equilibrated AKTA FPLC Q-Sepharose 10/10 column. The column was washed with 20 CV of buffer A and the bound protein was eluted by the application of a linear KCl gradient from 0% buffer B to 100% Buffer B (50 mM Tris-HCl pH7.5, 1 M KCl, 1 mM EDTA and 2 mM β-ME). Fractions containing the reporter protein were pooled, concentrated and dialyzed into storage buffer (50 mM Tris-HCl pH7.5, 50 mM KCl, 1 mM DTT and 10% glycerol). Protein aliquots were flash frozen in liquid nitrogen and stored at -80°C.

3.5.3 *In vitro* proteolysis assay

In vitro proteolysis assay was carried out in an 80 μl reaction mixture containing Lon activity buffer (50 mM Tris-HCl pH8.0, 10 mM $MgCl_2$, 1 mM DTT, and 10% glycerol), ATP regeneration system (50 mM creatine phosphate, 80 μg/ml creatine kinase, and 4 mM ATP), 10μM substrate, and Lon protease (either 100 nM *Mp*-Lon_6 or 400 nM *Ec*-Lon_6). The reaction was assembled and incubated at 30 °C. Aliquots were taken at designated time points and the reaction stopped by adding equal volume of 2XSDS-PAGE sample buffer. The reaction products were resolved by electrophoresis in 15% tris-tricine gel and quantified by Image-J software.

3.5.4 ^{35}S-labeled substrate preparation and proteolysis

For ^{35}S labeling of the chosen λ-cI-N reporters, the periplasmic protease deficient strain *degP*$^-$*prc*$^-$*spr*$^-$ was used for expression. Cultures of 50 mL were grown in LB broth at 37°C to an OD_{600} of between 0.8 and 1.0. Cells were harvested by centrifugation at 3,700 x g and the cell pellets were resuspended in 50 mL of M9 media supplemented with DOC mix lacking methionine and cysteine (US Biological). Cultures were induced with 1.0 mM IPTG and grown for an additional 20 minutes. A ^{35}S methionine/^{35}S-cysteine mixture (2.0 mCi, Perkin-

Elmer) was added to each culture and the cells were grown for an additional 2 hours at 37 °C. Cells were harvested and lysed in B-PER reagent (Pierce) containing 500 mM KCl, 2 mM β- ME, 0.1 mg/mL lysozyme, 5 U/mL DNaseI, and a protease inhibitor cocktail (Pierce). ^{35}S-labelled proteins were purified by chromatography on a Ni-NTA column, essentially as described above. The purified ^{35}S-labeled λ-cI-N-ssrA$_{MP}$ reporter proteins were buffer exchanged into storage buffer (50 mM Tris-HCl pH7.5, 50 mM KCl, 1 mM DTT and 10% glycerol) using an Econo-Pac-10DG desalting column (Biorad). Protein aliquots were flash frozen in liquid nitrogen and stored at -80° C.

Lon mediated proteolysis of ^{35}S-labeled λ-cI-N-ssrA$_{MP}$ and its tag peptide variants was performed at 30 °C using 100 nM of each Lon$_6$ in a degradation buffer (50 mM Tris-HCl pH8.0, 10 mM $MgCl_2$, 1 mM DTT, and 10% glycerol) that was supplemented with an ATP regeneration system (50 mM creatine phosphate, 80μg/mL creatine kinase, and 4 mM ATP). Proteolysis reaction samples, 10μl, were taken at indicated time points and quenched immediately by the addition of 20 μL of 20% trichloroacetic acid (TCA) and 10 μL of 10% BSA. The TCA-insoluble material was removed by centrifugation at 10,000 x g in a microcentrifuge. The amount of ^{35}S–labeled proteolytic peptides present in the supernatant was measured in a liquid scintillation counter. Initial rates of λ-cI-NssrA$_{MP}$ proteolysis by *Ec*-Lon or *Mp*-Lon were measured at a range of substrate concentrations. Each assay was repeated at least three times and the data were fitted to the Michaelis–Menten equation to determine K_M and V_{max} values.

3.5.5 *In vivo* MMS sensitivity assay

Strains that were assayed for MMS sensitivity were grown in LB to OD~1.5, respectively. Serial dilutions of each culture were made in LB and 10μl the diluted culture was spotted on appropriate plates either with or without 0.025% MMS. The plates were then incubated at 30 °C for 40 hours. The assay was performed with independent cultures for three times and one representative result was shown.

3.5.6 β-galactosidase assay

The production of β-galactosidase in each strain analyzed was quantified using a modified Miller Method [197]. The assay was performed independently for three times, each time in triplicates. The result from one representative experiment was shown.

3.6 Figures

Figure 3.1

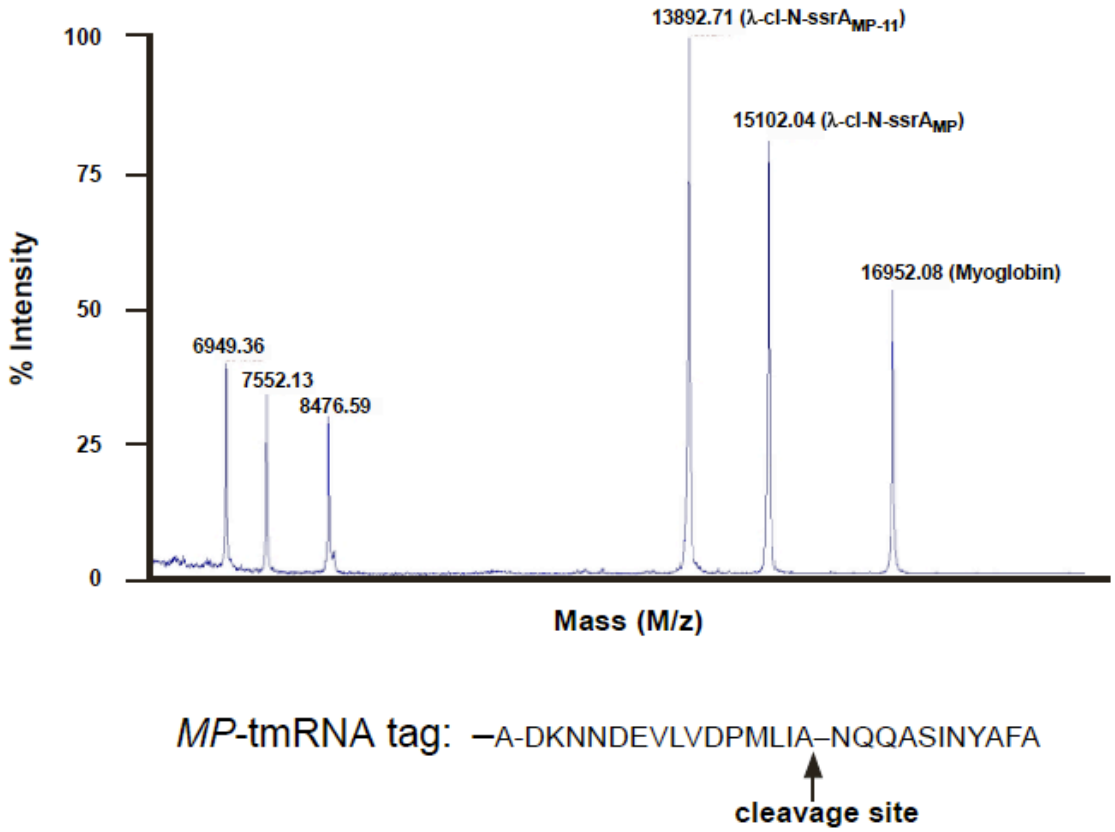


Figure 3.1: MALDI-TOF mass spectrometry assisted identification of the cleavage products of a reporter harboring *Mp*-tmRNA tag. The λ-cI-N-ssrA$_{MP}$ reporter protein was expressed in *E. coli* W3110 *clpX-clpP-lon-* strain, purified and subjected to MALDI-TOF MS analysis. The species with *m*/*z* =15,102 corresponds to the major product carrying the full *MP*-tmRNA tag (λ-cI-N30 ssrA$_{MP}$), whereas the peak with *m*/*z* =13,893 corresponds to the cleaved reporter, missing the last 11 amino acid residues of the *Mp*-tmRNA tag (λ-cI-N-$_{ssrAMP-11}$).

Figure 3.2

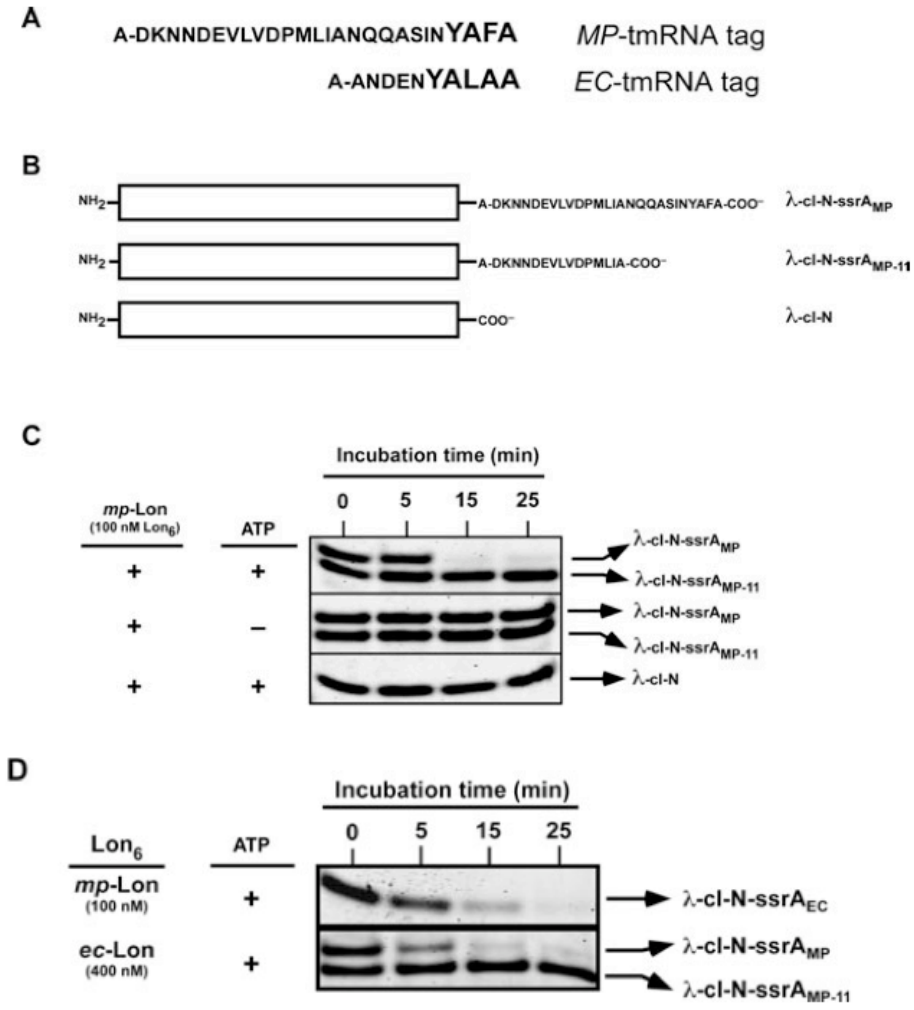


Figure 3.2: *Mp*-Lon selectively degrades a protein carrying the *Mp*-tmRNA tag. A) The amino acid sequences of the *Mp*-tmRNA and *Ec*-tmRNA tag are shown, with the C-terminal hydrophobic residues displayed in larger-size font for comparison. B) Schematic representation of the λ-cI-N-ssrA$_{MP}$ reporter, the λ-cI-N-ssrAMP-11 internal control and the λ-cI-N untagged reporter. C) *In vitro* proteolysis assays were carried out at 30°C in minimum Lon activity buffer, which contained 50 mM Tris-HCl pH8.0, 10 mM MgCl2, 1 mM DTT, and 10% glycerol. The reactions also contained 100 nM *Mp*-Lon$_6$, 10 μM substrate, 4 mM ATP (when indicated) and an ATP regeneration system. Aliquots were taken at designated time points, quenched with an equal volume of 2x SDS-sample buffer, resolved by electrophoresis on 15% Tris-tricine gel and stained with Coomassie Brilliant BlueR250. D) Reactions were carried out as described in C) with the designated Lon$_6$ concentration. *Mp*-Lon efficiently degrades λ-cI-N-ssrA$_{EC}$ at lower concentration (100 nM) while *Ec*-Lon requires higher concentration (400 nM) to effectively degrade λ-cI-N-ssrA$_{MP}$.

Figure 3.3

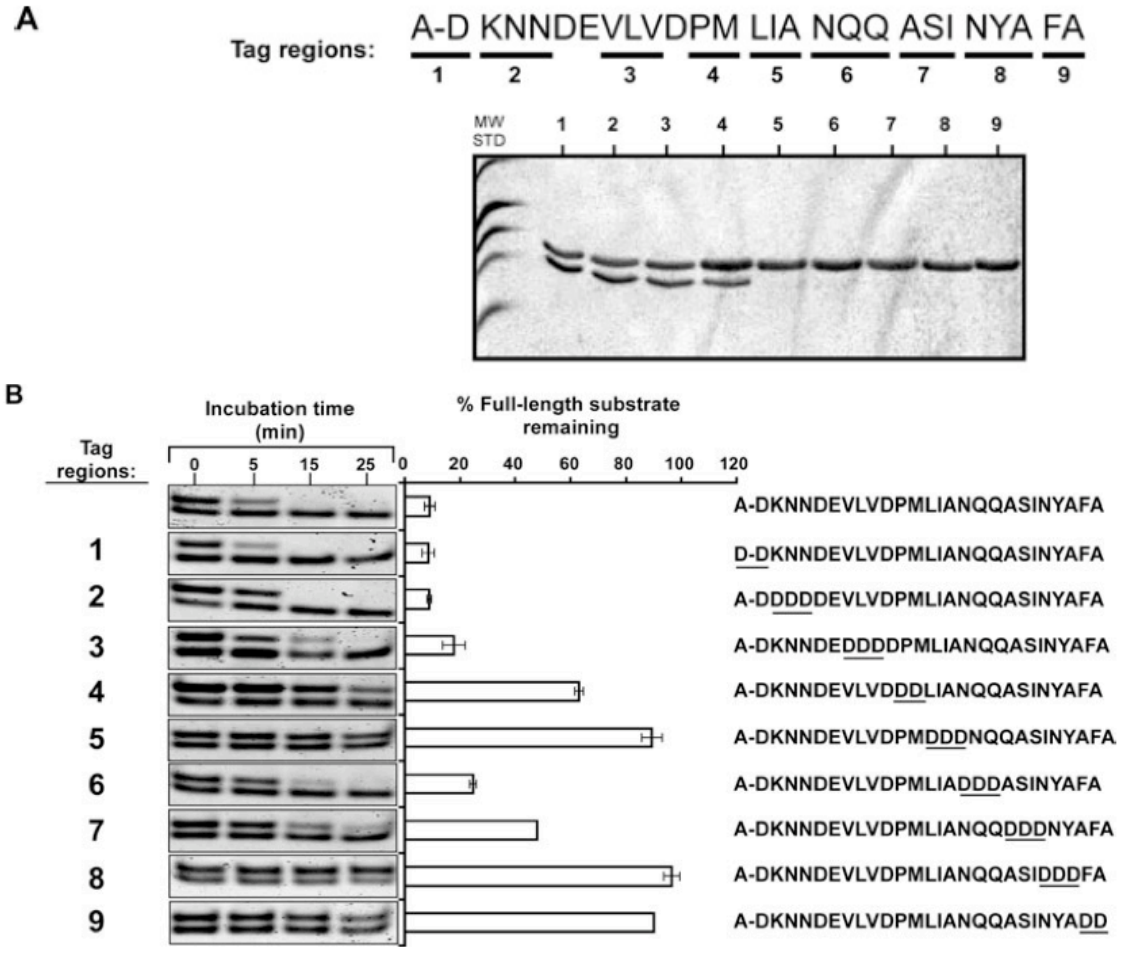


Figure 3.3: Aspartic acid substitutions in various regions of the *Mp*-tmRNA tag sequence result in alterations in its degradation pattern. A) The *Mp*-tmRNA tag sequence was divided into 9 regions. Amino acid residues within each region were substituted with aspartic acids. All λ-cI-N-ssrA$_{MP}$ reporter substrates, carrying *Mp*-tmRNA tag aspartic acid variants in each of the 9 regions, were individually purified, resolved by electrophoresis on a 15% Tris-tricine gel and stained with Coomassie Brilliant BlueR250. Aspartic acid substitutions in region 5 through 9 affect proteolytic cleavage of the λ-cI-N-ssrA$_{MP}$ reporter and production of the shorter λ-cI-N-ssrA$_{MP-11}$ fragment. B) The effect of aspartic acid substitutions, in each of the 9 regions of the *MP*-tmRNA tag, on degradation of λ-cI-N-ssrA$_{MP}$ reporter by *MP*-Lon$_6$ (100 nM) were assessed in an *in vitro* proteolysis assay as described in Figure 3.2. The amount of each full-length λ-cI-NssrA$_{MP}$ variant left at the 15 min time point, as an indicator of the efficiency of *MP*-Lon in degradation, was quantified using ImageJ and represented in bar graph format. Each experiment was repeated at least three times and the graph represents mean +/- standard deviation of three independent repeats.

Figure 3.4

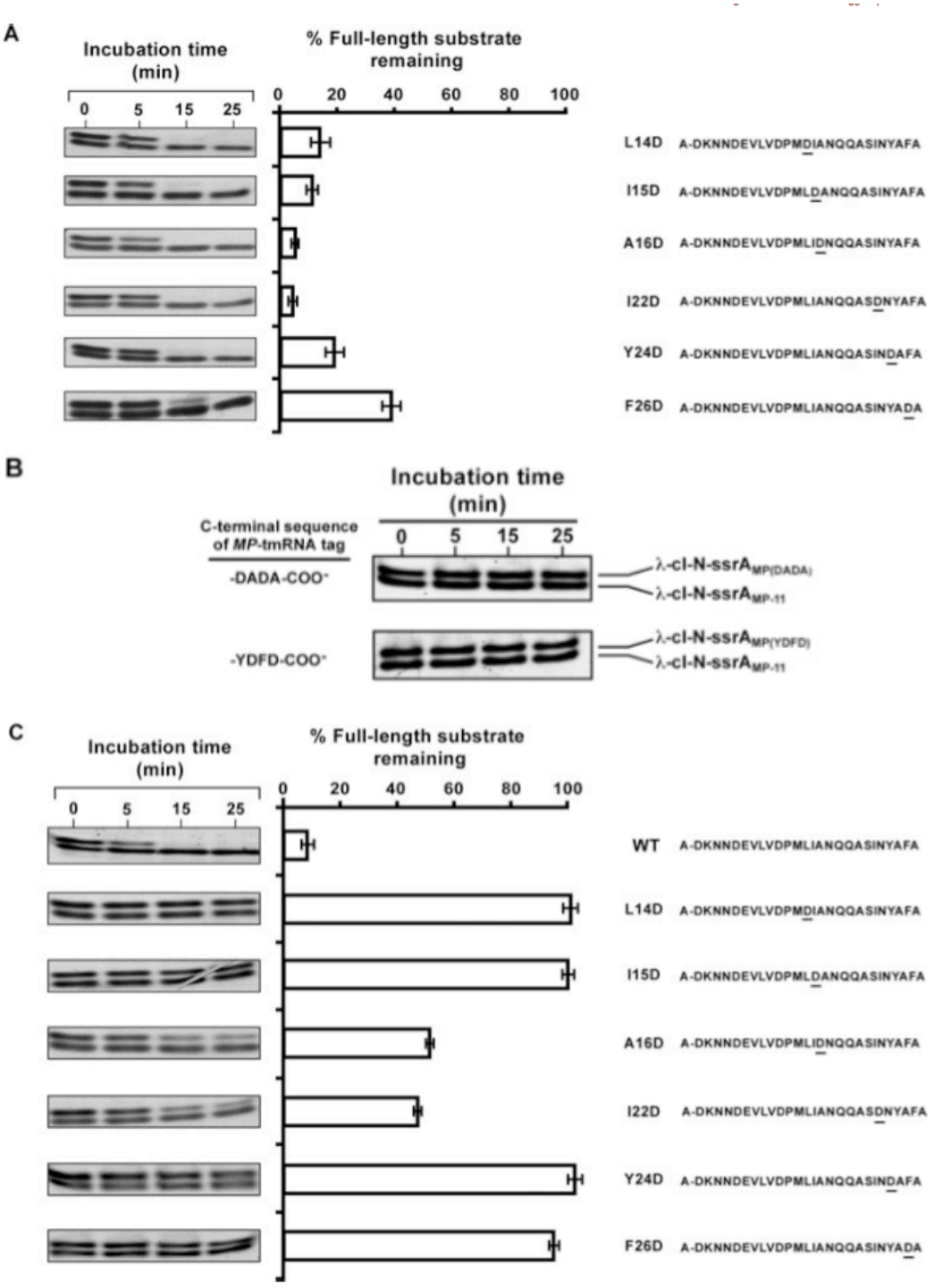


Figure 3.4: The contribution of amino acid residues in multiple regions of the *Mp*-tmRNA tag to degradation by Lon. A) Reporter substrates carrying single aspartic acid substitution in key signaling motifs in *Mp*-tmRNA tag were subjected to proteolysis by *Mp*-Lon_6 (100 nM) and analyzed as described in Figure 3.3. Each experiment was repeated at least three times and the graph represents mean +/- standard deviation of three independent repeats. B) The role of the C-terminal

four residues of *Mp*-tmRNA tag in degradation by *Mp*-Lon. λ-cI-NssrA_{MP} reporter variants, with the C-terminal residues of the *Mp*-tmRNA tag altered to either –DADA or –YDFD, were analyzed in the *in vitro* proteolysis assay by *Mp*-Lon_6 (100 nM), as described in Figure 3.2. Each experiment was repeated at least three times and the graph represents mean +/- standard deviation of three independent repeats. C) The effect of single aspartic acid substitutions in key signaling motifs of *Mp*-tmRNA tag on degradation of the λ-cI-N-ssrA_{MP} reporter by *Ec*-Lon. Reporter substrates carrying single aspartic acid substitution in key signaling motifs in *Mp*-tmRNA tag were subjected to proteolysis by *Ec*-Lon (400 nM) and analyzed as described in Figure 3.3. Each experiment was repeated at least three times and the graph represents mean +/- standard deviation of three independent repeats.

Figure 3.5

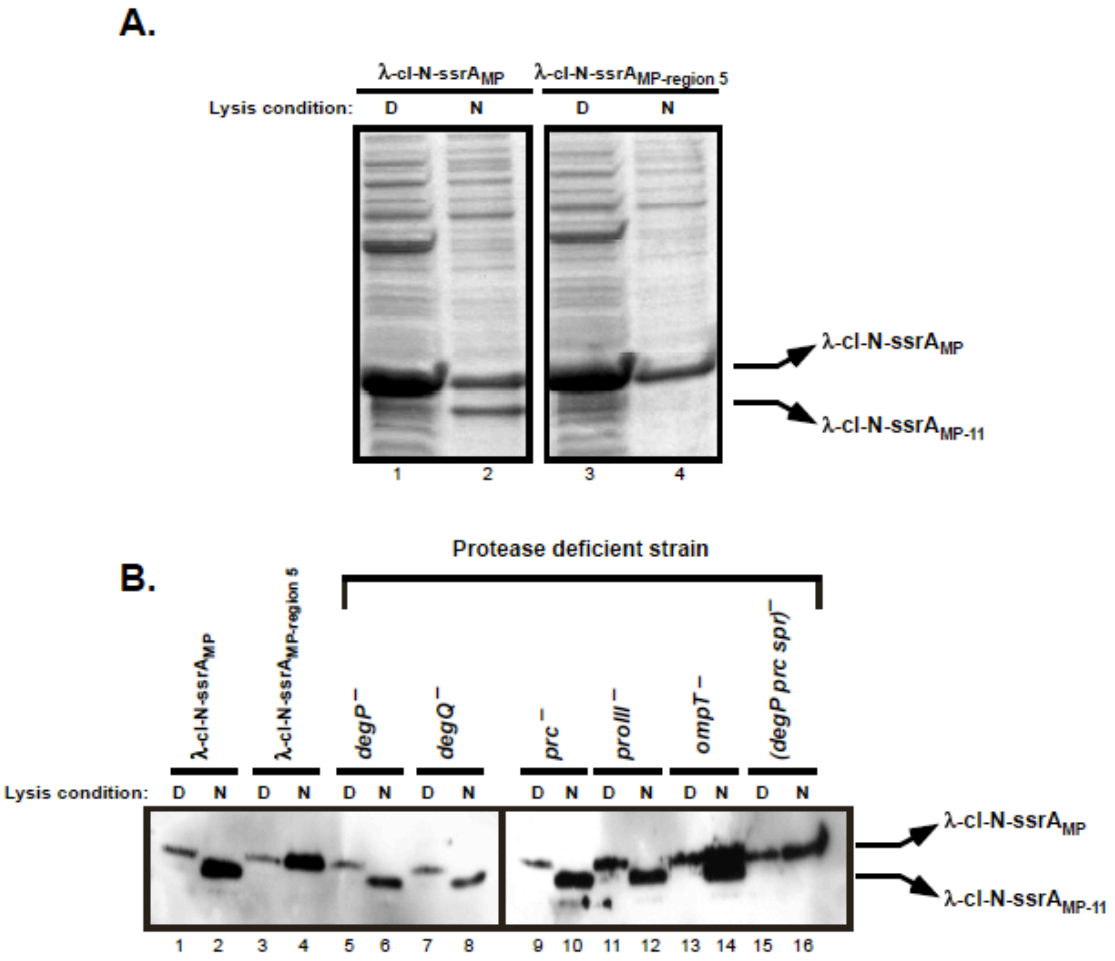


Figure 3.5: Periplasmic proteases are responsible for the cleavage of the reporter protein carrying *MP*-tmRNA tag. A) Either λ-cI-N-ssrA$_{MP}$ or its variant with region5 substituted to aspartic acids was expressed in a *clpX*$^-$*clpP*$^-$*lon*$^-$ strain. Cells were then lysed under either denaturing condition (D) or native condition (N). The cellular proteins were then resolved in 15% tris-tricine gel and visualized by coommassie staining. The cleavage product λ-cI-N-ssrA$_{MP-11}$ was observed only for λ-cI-N-ssrA$_{MP}$ under native lysis condition. B) *E. coli* strains lacking one or more periplasmic protease(s) were screened in the effort of finding the protease cleaving the *MP*-tmRNA tag. The λ-cI-N-ssrA$_{MP}$ reporter was expressed in these strains and cells were lysed both under native and denaturing conditions. The full-length λ-cI-N-ssrA$_{MP}$ and the truncated product was deteted by western blotting using anti-His6 antibody since the reporter carry an internal His6 epitope (Lane 5-16). The samples described in A) were also included in this analysis as controls to confirm that neither ClpXP nor Lon is responsible for the cleavage (Lane 1-4). Only in cells lacking both DegP and Tsp activity can the reporter remain intact under native lysis condition (Lane 15, 16), indicating that these two proteases are responsible for the cleavage of the λ-cI-N-ssrA$_{MP}$ reporter.

Figure 3.6

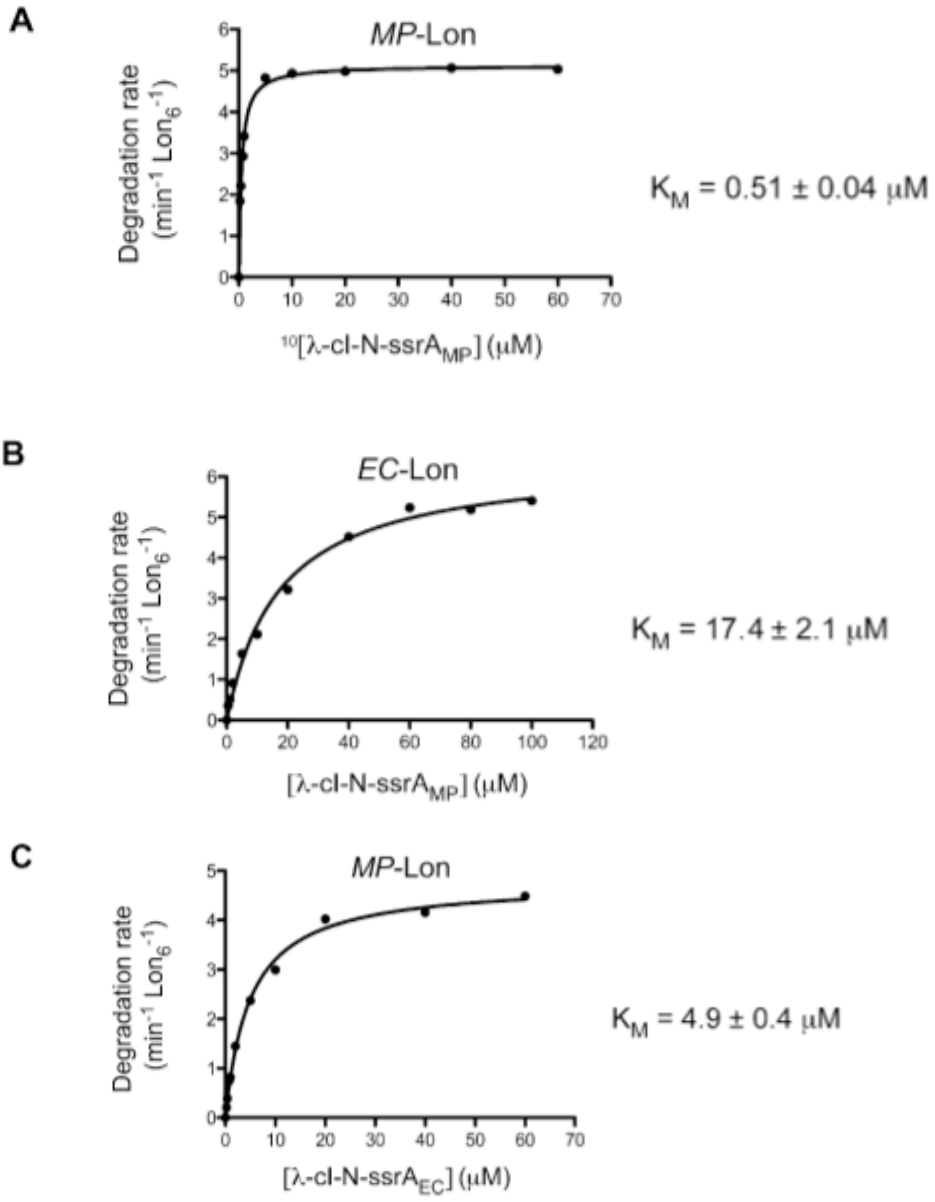


Figure 3.6: Steady-state kinetic analysis of the degradation of a [35S]-labeled tmRNA tagged protein by *Mp*-Lon and *Ec*-Lon. Degradation velocities were measured at a range of substrate concentrations and the data were fitted to the Michaelis-Menton equation. Each experiment was repeated at least three times. Shown are representative Michaelis-Menton plots for the degradation λ-cI-N-ssrA$_{MP}$ by A) 100 nM *Mp*-Lon$_6$ (K_M = 0.50 ± 0.04 μM and Vmax = 5.1 ± 0.1 min^{-1}Lon$_6^{-1}$), B) 100 nM *Ec*-Lon$_6$ (K_M = 17.4 ± 2.0 μM and Vmax = 6.4 ± 0.2 min^{-1}Lon$_6^{-1}$) and C) the degradation λ-cI-N-ssrA$_{EC}$ by 100 nM *Mp*-Lon$_6$ (K_M = 4.9 ± 0.4 μM and Vmax = 4.8 ± 0.1 min^{-1} Lon$_6^{-1}$).

Figure 3.7

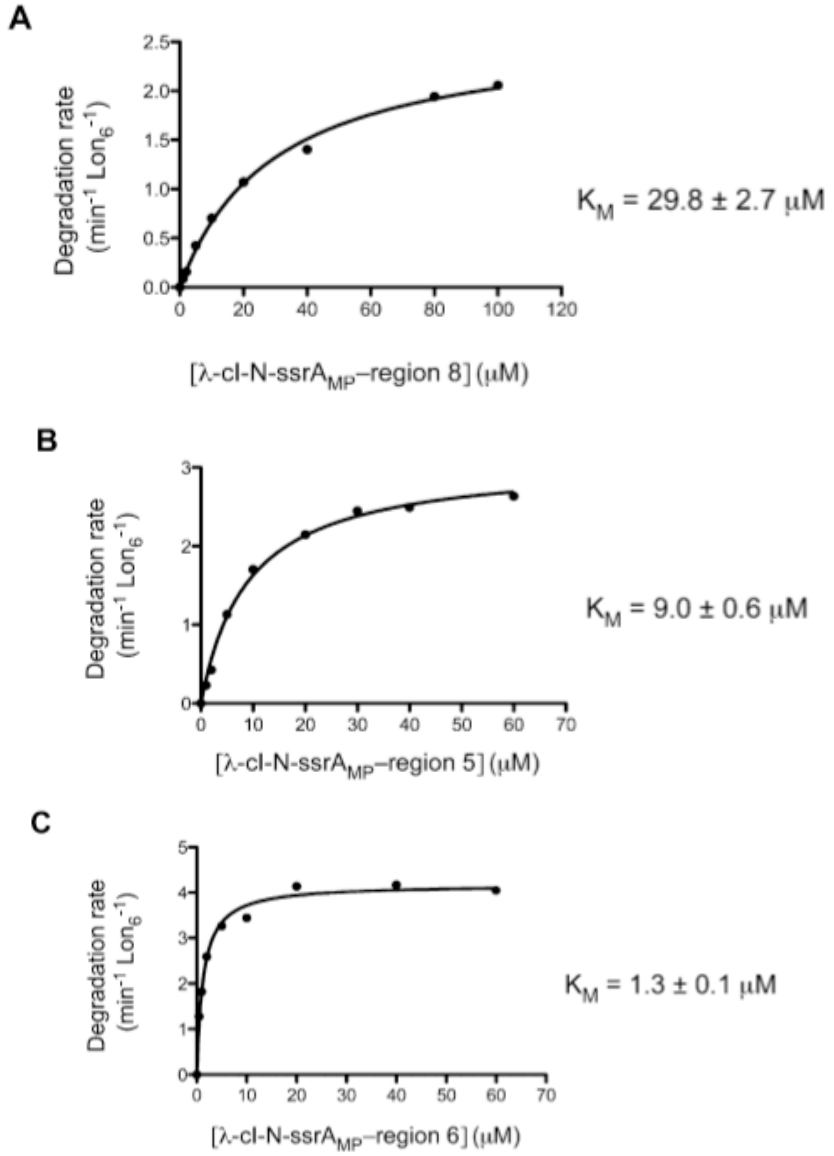


Figure 3.7: Steady-state kinetic analysis of the degradation of a [35S]-labeled protein carrying *Mp*-tmRNA tag sequence variants by *Mp*-Lon. Degradation velocities were measured at a range of substrate concentrations and the data were fitted to the Michaelis-Menton equation. Each experiment was repeated at least three times. Shown are representative Michaelis-Menton plots for degradation by 100 nM *Mp*-Lon_6 of λ-cI-N-$ssrA_{MP}$ tag sequence variants, carrying aspartic acid substitutions in A) region 8 (K_M = 29.8±2.7 mM and *V*max = 2.6±0.1 min^{-1} Lon_6^{-1}), B) region 5 (K_M = 9.0±0.6 mM and *V*max = 3.1±0.1 min^{-1} Lon_6^{-1}), and C) region 6 (K_M = 1.3±0.1 mM and *V*max = 4.2±0.1 min^{-1} Lon_6^{-1}) of the *Mp*-tmRNA tag.

Figure 3.8

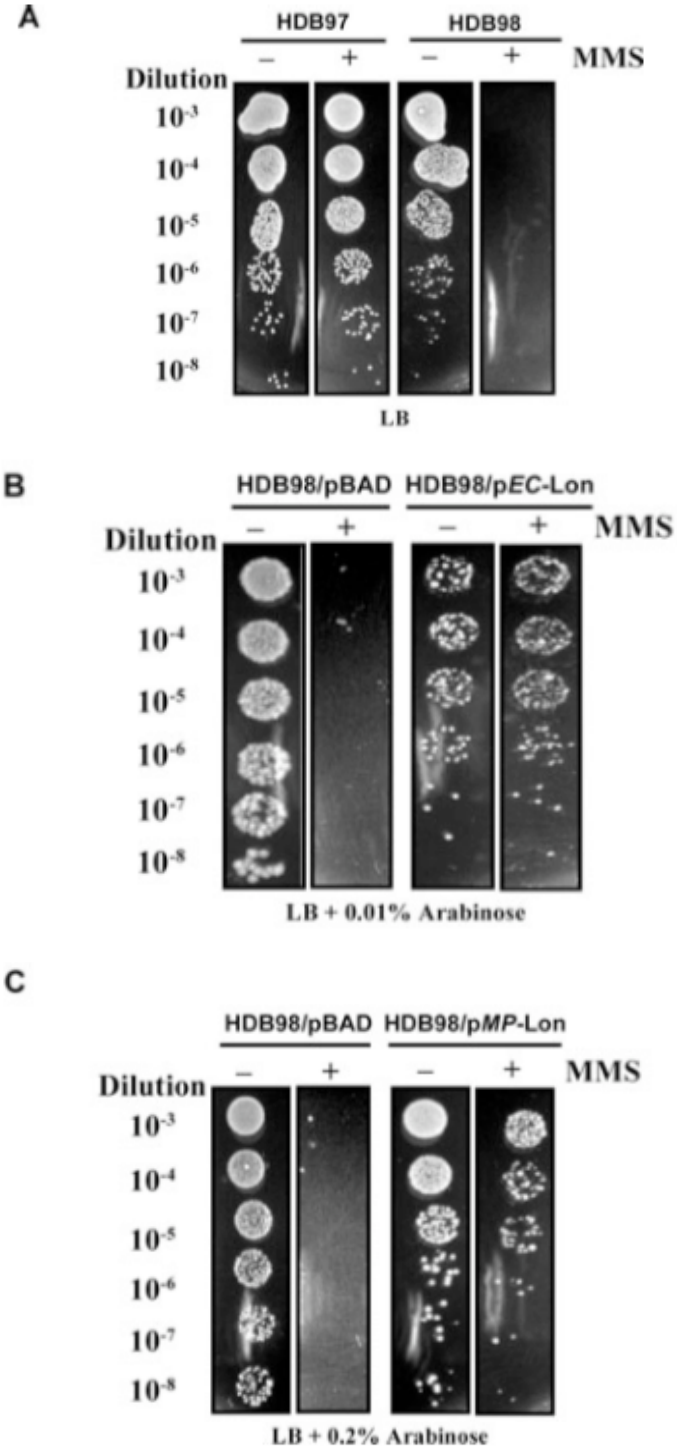


Figure 3.8: MMS sensitivity assay for complementation of the SulA phenotype of *E. coli lon*$^{-}$ cells by *Mp*-Lon. The ability of various *E. coli* strains to survive in the presence of 0.025% MMS was assessed by spotting 10μL of serially diluted cultures onto plates either in the presence or absence of MMS. Each complementation experiment was repeated at least three times, and a

representative example of each assay is shown. A) The abilities of the HDB97 parental strain and its isogenic HDB98 *lon*$^{-}$ mutant to grow in the presence of MMS. B) The HDB98 *lon*$^{-}$ strain was complemented with either pBAD vector or pBAD-*Ec*-Lon. The expression of plasmid borne *Ec*-Lon was induced by 0.01% arabinose. C) The HDB98 *lon*$^{-}$ strain was complemented with either the empty pBAD vector or pBAD-*Mp*-Lon. The expression of plasmid borne *Mp*-Lon was induced by 0.2% arabinose.

Figure 3.9

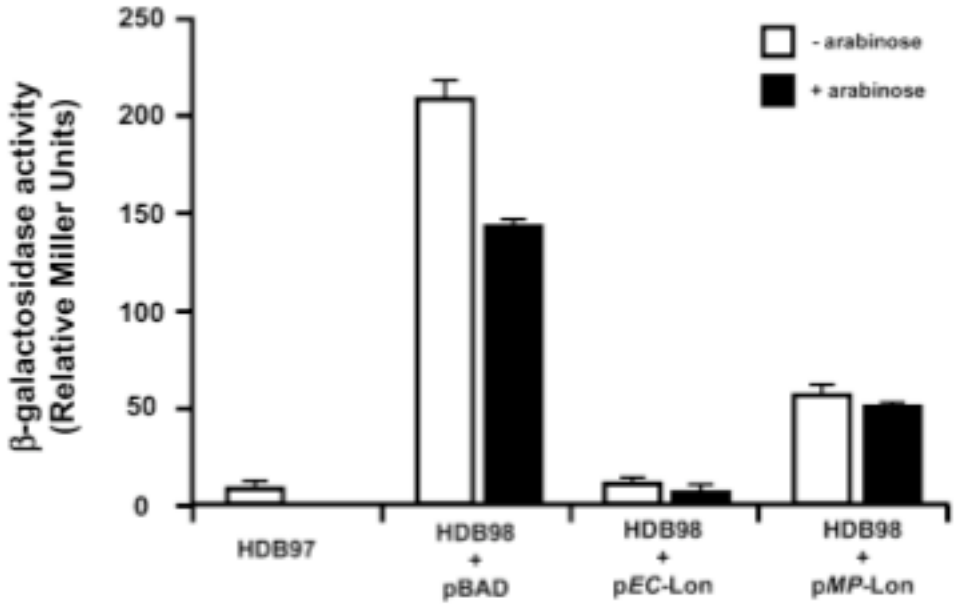


Figure 3.9: Beta-galactosidase assay for complementation of the RcsA phenotype of *E. coli* cells by *Mp*-Lon. Beta-galactosidase activity was assayed, in each of the indicated HDB97 parental and HDB98 *lon*$^{-}$ strains, using the ONPG substrate, the hydrolysis of which gives rise to an absorbance signal at 420 nm. Strains harboring the complementation plasmids were grown either in the presence or absence of the inducer (0.1% arabinose) to affect the expression of the plasmid borne *Ec*-Lon or *Mp*-Lon proteases. All beta-galactosidase activities reported are the means of three or more assays and the graph represents mean +/- standard deviation of three independent repeats.

Figure 3.10

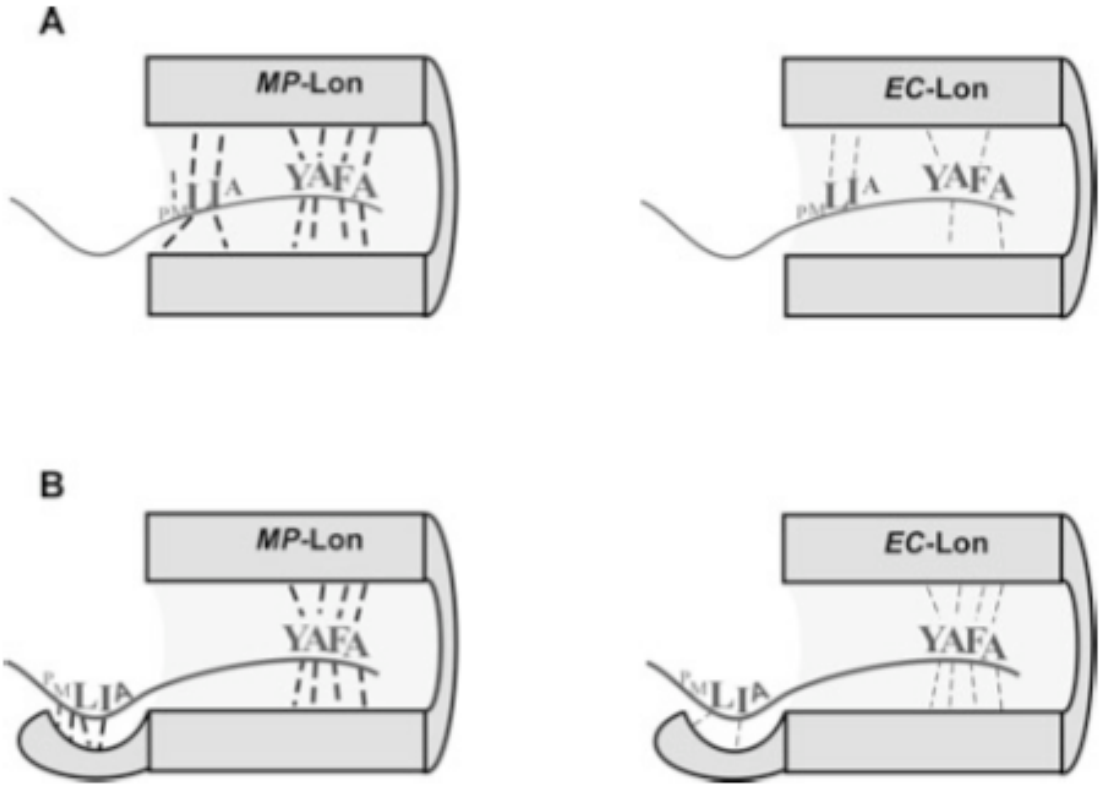


Figure 3.10: Proposed model for co-evolution of multipartite interactions between Lon and tmRNA tag. Both *Mp*-Lon and *Ec*-Lon proteases contact multiple discrete signaling elements within the *Mp*-tmRNA tag sequence. We propose that pre-existing multipartite weak interactions (thin dashed lines) between Lon and tmRNA tag have co-evolved in *Mycoplasma* to facilitate strong (bold dashed lines) and selective recognition of tagged proteins. To accommodate these multipartite interactions, either: A) the binding pocket of Lon protease is large enough to encompass the last 16 residues of the *Mp*-tmRNA tag, or B) the binding pocket of Lon encompasses only the C-terminal 4 residues of the tag while the internal signaling motif interacts with an allosterically regulated site on the periphery of the peptide-binding pocket of Lon. See Discussion for details.

Table 3.1

Substrate	K_M (μM)	V_{max} (Substrate degraded $min^{-1}Lon_6^{-1}$)
λ-cI-N-ssrA$_{MP}$	**0.51±0.04**	**5.1±0.1**
λ-cI-N-ssrA$_{MP\text{-}region5}$	**9.0±0.6**	**3.1±0.1**
λ-cI-N-ssrA$_{MP\text{-}region6}$	**1.3±0.1**	**4.2±0.1**
λ-cI-N-ssrA$_{MP\text{-}region8}$	**29.8±2.7**	**2.6±0.1**
λ-cI-N-ssrA$_{EC}$	**4.9±0.4**	**4.8±0.1**

Table 3.1: Steady-state kinetic parameters for degradation by *Mp*-Lon of a protein carrying either the wild-type *Mp*-tmRNA tag, representative aspartic acid substitution variants of the *Mp*-tmRNA tag, or the wild-type *Ec*-tmRNA tag.

Chapter 4 Lon substrates in *Yersinia pestis*: Proteomic profiling of the substrate identities and the cleavage sites

4.1 Summary

Pathogens like *Yersinia pestis* have complicated mechanisms of regulating their gene expression, especially for genes that contribute to virulence. Proteolysis of key virulence gene regulators is of vital importance in maintaining the appropriate level of the specific regulator. Lon protease has been implicated in regulating the pathogenesis of *Yersinia pestis* by degrading YmoA, the negative regulator of its Type III secretion system (TTSS) components, upon activation of the TTSS. In addition, Lon is also known to degrade the thermal sensor RovA in *Yersinia*. However, the native substrates of Lon protease in *Yerisnia* have not been systematically identified. In this study, we utilized an inactive mutant of Lon protease to trap the substrates of *Yersinia pestis in vivo*. The substrates were then identified by mass spectrometry. Over 100 proteins were identified as potential Lon substrates, including several master transcriptional regulators, and many metabolic proteins, which can provide great insight into how Lon systematically regulates gene expression in *Yersinia*. In addition, we took a further step and mapped the cleavage pattern of Lon protease when degrading these different substrates. Finally, we compared the cleavage pattern of Lon proteases from different bacterial species.

4.2 Introduction

Pathogenic bacteria are equipped with very complex virulence mechanisms against the host cell. Many bacterial species adopt one or more types of secretion apparatus for the delivery of effector proteins into host cells. In addition, when the bacteria are engulfed into the host cell, such as macrophages, they are subjected to very adverse conditions such as change in growh temperature, the acidic phagosomal pH, and oxidative stress. To adjust to the environmental change and overcome the defense mechanisms of the host cell, bacteria rely heavily on proteolysis not only for the removal of damaged or unfolded proteins resulting from adverse host environments, but also for the timely removal of regulatory proteins in response to the environmental change. Energy dependent proteases such as ClpXP, ClpAP, and Lon are of particular importance in this process, as they are the house-keeping proteases that contribute significantly in post-translational protein quality control.

In gram-negative pathogens, both the Clp family proteases and Lon protease are conserved. They all belong to the AAA+ (ATPase associated with various cellular activities) superfamily. The Clp family proteases consist of two heptameric rings of the protease subunit and a hexameric ring of the ATPase subunit attached non-covalently at one or both ends of the protease ring. The hexameric ATPase compartment is responsible for substrate recognition,

unfolding, translocation, and ATP hydrolysis. The protease compartment contains the His-Asp-Ser catalytic triad and is responsible for peptide bond cleavage. Lon protease is composed of six identical subunits, forming a hexameric rings. Each subunit is a single peptide chains carrying an N-terminal domain that is indicated to be involved in substrate recognition, an ATPase domain responsible for ATP hydrolysis and substrate unfolding and translocation, and a C-terminal protease domain that contains the catalytic Ser-Lys dyad. The involvement of the Clp family proteases and Lon protease in the regulation of Type III secretion system (TTSS) in several bacteria species has been extensively studied.

There are cases when the lack of the AAA+ protease activity in certain bacterial species would lead to enhanced or constitutive expression of the TTSS. For instance, in the plant pathogen *Pseudomonas syringae*, the TTSS is encoded on the hypersensitive response and pathogenicity island (*hrp*). Lon protease degrades the positive regulator of *hrp*, HrpR. Removal of Lon activity from *Pseudomonas syringae* leads to constitutive expression of the TTSS genes and the accumulation of effector proteins [146, 198]. In *Salmonella enterica*, lack of Lon protease activity results in the accumulation of both HilC and HilD [199], which in turn upregulates HilA. HilA is the activator of *Salmonella* TTSS that is encoded on the *Salmonella* pathogenicity island 1 (SPI1). The upregulation of HilA results in enhanced expression of SPI1 genes [200]. As a result, a *Salmonella lon*$^-$ mutant exhibits enhanced invasion of epithelial cells and induces massive apoptosis in macrophages [201, 202]. However, the virulence of *Salmonella lon*$^-$ mutant in animal models is attenuated despite the enhanced SPI1 gene expression, probably because the cells are not able to overcome the oxidative stress and the acidic pH without the activity of Lon protease [142]. In *Salmonella*, ClpXP mainly targets the FlhD/FlhC master regulator (60). In the *Salmonella clpXP*$^-$ mutant, the accumulation of FlhD/FlhC master regulator results in the upregulation of FliZ, which modulates the activator of TTSS HilA [203], and eventually leads to the increased expressed SPI1 expression [201].

In other cases, the AAA+ proteases degrade the negative regulators of the TTSS and lead to the expression of TTSS genes. For example, in enterohemorrhagic *E. coli* (EHEC), the TTSS is encoded on the locus of enterocyte affacement (LEE) under the regulation of a negative regulator, GrlR. ClpXP targets GrlR and upregulates the expression of TTSS expression [204]. A similar case is observed in *Yersinia pestis*. *Yersinia pestis* is the causative agent of plague, a disease that is often fatal to the infected host. *Yersinia pestis* uses a virulence system encoded on a 70kb plasmid, named pCD1, to escape host immune defenses and achieve intracellular replication and dissemination [205, 206]. The pCD1 plasmid encodes for a set of secreted effector proteins termed Yersinia outer proteins (Yops) and a delivery system classified as a type III secretion system (TTSS) [207, 208]. Upon contact of *Yersinia* and the host cells, the virulence system encoded on pCD1 is induced and assembled into a needle like structure on the surface of the bacteria, which then enables the direct

injection of Yops into the cytoplasm of targeted host cells [209]. *In vitro*, Yop expression is repressed at the growth temperature of 26°C. At 37°C, Yops expression is induced but the secretion is blocked unless the extracellular calcium in the growth media is removed [210, 211]. The expression of genes encoded on pCD1 is tightly regulated. One way that *Yersinia* has adopted to regulate its virulence genes is through the negative regulator YmoA. The histone-like protein has been shown to repress the expression of genes on pCD1 when *Yersinia* is not in contact with any host cells or when it is grown at 26°C *in vitro* [212, 213]. However, upon contact with the host cells, or a growth temperature switch from 26°C to 37°C, the repression effect of YmoA is alleviated, evidently through the rapid proteolysis and removal of YmoA by a number of energy-dependent proteases, especially Lon protease [141]. ClpXP also has the ability to degrade YmoA but play a relatively minor role in regulation of *Yersinia* TTSS compared to Lon protease.

In this study, to further investigate the role of Lon protease in *Yersinia* physiology and pathogenesis, we performed a proteome-wide profiling of Lon substrates in *Yersinia pestis*. We identify over 100 proteins as potential Lon substrates, including *Yersinia* transcriptional regulators, virulence factors, and metabolic enzymes. We then took a step further to investigate the common elements in these substrates, such as the cleavage sites within different substrate by Lon protease.

4.3 Results

4.3.1 Construction of *Yp*-Lon^{trap} to capture native substrates of Lon protease in *Yersinia pestis*

To isolate the native substrates of Lon protease in *Yersinia pestis*, we utilized an inactive mutant of Lon protease with an S679A mutation (hereinafter designated as *Yp*-Lon^{trap}), which renders the protease incompetent in cleaving the peptide bond. However, the protease's ability in recognizing, unfolding, and translocating substrates, as well as its ability to hydrolyze ATP, is intact. To facilitate the purification of the *Yp*-Lon^{trap}, together with its substrates, one His6 tag and three tandem Myc tags were added to the C-terminus of the protease (Figure 4.1). We obtained *Y. pestis* KIM5-3001 parent strain and its isogenic *lon*$^{-}$ mutant from Dr. Plano from University of Florida. The construct expressing *Yp*-Lon^{trap} was introduced into *Y. pestis* KIM5-3001 *lon*$^{-}$ and the parental strain. The expression level of *Yp*-Lon^{trap} can be regulated by the final concentration of IPTG added to the growth media. We determined that when the final concentration of IPTG is 0.5mM, *Yp*-Lon^{trap} expresses at near endogenous level. The expressed *Yp*-Lon^{trap} can be detected by α-*Yp*-Lon (gift from Dr. Plano), α-His6 (GE healthcare), as well as α-Myc, suggesting the successful incorporation of the tandem affinity tags into the *Yp*-Lon ORF.

4.3.2 Trapping and identification of native *Yp*-Lon substrates *in vivo*

Yersinia pestis has two different growth conditions *in vitro*. When the bacteria are grown in high calcium media at 26°C, the expression of the TTSS genes are repressed. In contrast, when the bacteria are grown in low calcium media at 37°C, the repression of the TTSS genes is alleviated and those genes are expressed. We hypothesized that Lon may have different degradation profiles under the two growth conditions. As a result, we performed the trapping assay under both growth conditions. To obtain the cellular substrates recognized by *Yp*-Lon^{trap}, protease-substrate complexes were isolated from *Y. pestis* KIM5-3001 *lon*$^{-}$ complemented by the plasmid borne *Yp*-Lon^{trap} (Figure 4.2). *Y. pestis* KIM5-3001 *lon*$^{-}$ harboring an empty pMMB67EH vector was also included in the trapping assay as a control for any proteins that may be non-specifically pulled down during the tandem affinity purification. Proteins that co-purified with *Yp*-Lon^{trap} were resolved in 15% tris-tricine gel and stained with coomassie blue for visualization of the captured proteins. The co-purified proteins have a wide molecular weight range, ranging from 10kDa to 90kDa. The protein mixture was then analyzed directly by LC-MS-MS and approximately 100 proteins were identified (Table4.1). Among the identified proteins, there are many metabolic proteins, transcriptional regulators, translational factors, as well as several proteins that have been indicated to be *Yerisnia* virulence factors. Only a few proteins were identified by mass spectrometry in samples generated from *Yersinia* harboring the empty vector, suggesting that the tandem affinity purification is specific enough to isolate *Yp*-Lon^{trap}, as well as its associated proteins.

4.3.3 Confirmation of potential substrates *in vitro*

To further confirm that the proteins identified in mass spectrometry are indeed *Yp*-Lon substrates, we cloned the following genes: y2853, HspQ, NusG, Fur, y0390, RsuA, CRP, FtsZ, and TrmJ, into an expression vector. We then over-expressed and purified these proteins. We also overexpressed and purified *Yp*-Lon protease. Finally we tested the ability of *Yp*-Lon protease to degrade these potential substrates *in vitro*. The TTSS negative regulator YmoA was reported previously as a Lon substrate under TTSS inducing conditions [141]. Although it was not identified as a potential *Yp*-Lon substrate in our mass spectrometry analysis, I cloned, expressed, purified, and tested it in an *in vitro* proteolysis assay as a positive control for degradation by Lon. It has also been reported previously that certain proteins, especially transcriptional regulators such as YmoA, undergo a conformational change upon the temperature shift from 30°C to 37°C, which in turn exposes certain sequence elements that are otherwise concealed in the folded protein, and thus allows proteases to bind and degrade the protein. As a result, we performed the proteolysis assay at both 30°C and 37°C, in the attempt to distinguish if Lon protease exhibits different activity against a particular substrate in response to a temperature change.

Results showed that 8 out of the 10 proteins tested are efficiently degraded by *Yp*-Lon protease *in vitro*, with the exception of TrmJ, which showed minimal degradation at either 30°C or 37°C in 90 minutes (Figure 4.3). The proteins that showed efficient degradation also exhibited different degradation rates, suggesting that *Yp*-Lon probably has different affinities and degradation rates for different substrates. For instance, y2853 is completely degraded within 15 minutes, while there is still a small amount of CRP remaining intact (~10%) after 90 minutes incubation. Another observation that we made was that several proteins showed higher degradation rates when incubated at 37°C than 30°C, such as HspQ, NusG, y0390, and FtsZ. This is consistent with our hypothesis that at a higher temperature, certain proteins may be more susceptible to proteolysis by Lon. Whether this increased susceptibility is caused by a conformational change in the substrate or an increased Lon activity at a higher temperature remains to be elucidated.

Since the majority of the potential *Yp*-Lon substrates we tested can indeed be degraded by *Yp*-Lon *in vitro*, we conclude that *Yp*-Lon^{trap} is successfully translocating its cellular substrates and retaining them inside the protease chamber, and allowing the co-purification of these proteins. In addition, the wide range of substrates identified in this study also suggests that Lon protease is involved in many important biological pathways and cellular functions of *Yersinia*. Furthermore, the identification of several master transcriptional regulators such as CRP, as well as *Yersinia* virulence factors, such as Ymt and Ail, suggests that Lon plays an important role in *Yersinia* physiology and pathogenesis. The biological importance of the turnover of these proteins by *Yp*-Lon protease is to be further explored.

4.3.4 Lon protease prefers to cleave after hydrophobic residues

Aside from the consensus recognition motif that Lon protease recognizes, we are also interested in the cleavage sites of Lon protease. We set out to identify the cleavage sites in different substrates by *Yp*-Lon. In order to analyze the cleavage products, we performed *in vitro* proteolysis assay and let the reaction continue until all the substrates were completely degraded (indicated by the absolute disappearance of the band corresponding to the substrate). We then TCA precipitated the intact proteins and collected the supernatant from the TCA precipitation. We then analyzed the supernatant, which contains the short peptides generated from Lon cleavage, by mass spectrometry. Analysis of the peptides permitted the unambiguous identification of peptide sequence. The C-terminal residue of each peptide corresponds to amino acid residues that Lon protease cleaves after. In this analysis, we used the best *Yp*-Lon substrates (RsuA, y2853, CRP, NusG, YmoA, Fur, HspQ, and y0390), from our *in vitro* degradation studies. Analysis of the short peptides generated from the digestion of these proteins by *Yp*-Lon revealed that *Yp*-Lon primarily cleavages after hydrophobic residues such as Leu, Ala, Val, and Phe (Figure 4.4A). Figure 4A

summarizes the primary cleavage site (Postiion7), 6 amino acids upstream of the cleavage site (Position1-6), and 6 amino acids downstream of the cleavage site (Position8-13). Amino acids that are identified at Position7 showed a high degree of conservation, with Ala and Leu being the most frequently identified amino acids, while residues at all other positions showed little or no conservation.

Previous studies on Lon protease in our lab revealed the similarities in the substrate specificity between Lon proteases from *E. coli* and *Mycoplasma pneumoniae* [214]. We hypothesized that the similarity may also be extended to the preferred cleavage sites. In order to test this theory, we utilized the model substrate for Lon protease, the λ-cI-N-ssrA$_{MP}$, and we also purified two more known *Ec*-Lon substrates S2 and L9 [215] to try to analyze the cleavage sites within these substrates. The three substrates were digested by either *Ec*-Lon or *Mp*-Lon and the digestion products were analyzed by LC-MS-MS. As we expected, both *Ec*-Lon and *Mp*-Lon exhibited very similar preference in the cleavage sites as *Yp*-Lon. The hydrophobic residues Ala, Leu, Val, and Phe are still among the most frequently identified cleavage sites (Figure 4.4B, C). As a result, we analyzed peptides generated from different substrates by all three proteases, and propose that Lon protease, in general, prefers to cleave after hydrophobic residues such as Leu, Ala, Phe, and Val (Figure 4.4D).

4.3.5 Cleavage by Lon protease results in short peptides ranging from 8-17 amino acids in length

The analysis of the preferred cleavage sites by Lon protease led us to postulate that when the unfolded substrate is translocated through the central pore and reaches the protease chamber, the proteolytic active sites scans for its preferred cleavage sites within the polypeptide and cleaves the peptide bond when a site is found with little preference for the surrounding amino acids. It is also important to note that Lon does not cleave after every preferred residue presented in the entire substrate. It remains unclear how Lon chooses between 2 identical residues at different locations of the substrate. For one substrate molecule, if two of the most preffered cleavage sites are located very close to each other, it is most likely that Lon cleavage happens only after one of these two residues.

To further understand the step size of Lon protease in peptide bond cleavage, we also determined the length of the peptides generated from Lon digestion. The peptides range from 6 amino acids up to 35 amino acids. However, most of the peptides are of 8-17 amino acid long (Figure 5). Although more mechanistic evidence is needed to further understand the pattern of peptide bond cleavage by Lon protease, it is conceivable that the step size of peptide bond cleavage is linked to the step size of translocation. The length of the substrate that is translocated into the protease chamber by each stroke of translocation, together with the presence of any preferred cleavage sites within the sequence

newly translocated, as well as the positioning of the cleavage site to make it an optimal target for the catalytic dyad of Lon, may be the determinants for where Lon cleaves.

4.4 Discussion

4.4.1 Consensus recognition sites and cofactors for Lon protease

Although over the years many substrates of Lon protease (especially *Ec*-Lon protease) have been identified and studied individually, and the recognition sites in these individual proteins have been identified, there is no systematic study of a consensus recognition motif for Lon protease. It is also known that Lon prefers hydrophobic residues in its recognition motifs. With the large pool of *Yp*-Lon substrates we have identified, we are interested in searching for the recognition sequence within different substrates and hopefully finding the consensus recognition sequence of Lon protease.
For proteins identified by mass spectrometry that were not degraded *in vitro* by *Yp*-Lon, it is possible that some proteins co-purified non-specifically with *Yp*-Lon^{trap}. Another more intriguing possibility is that the *in vitro* condition is not comparable to the physiological condition where a certain protein is subjected to proteolysis by Lon protease. For instance, a protein may become a Lon substrate *in vivo* when it is incorrectly folded, when the change in the environment induces a conformational change that exposes the recognition motif for Lon, or when it is delivered to Lon protease via the facilitation of another factor. In a limited *in vitro* system, when these conditions are not met, the protein then becomes resistant to Lon protease.

4.4.2 Deciphering the role Lon protease is playing in *Yersinia* physiology and pathogenesis

Protein turnover is no less important that its synthesis. Lon protease is playing a very important role in many biological processes. It exerts its housekeeping function by degrading misfolded proteins and tmRNA tagged proteins in certain bacterial species. It also regulates the expression of many genes by degrading small regulatory proteins, which, for instance, then allows the bacteria to respond to environmental stimulants in a precise and timely manner. We have long lasting interest in understanding how the pathogenesis of *Yersinia pestis* is regulated. The identification of Lon substrates in *Yersinia* may reveal the role of proteolysis in regulating the expression or suppression of its virulence genes. Another interesting approach we can take is to compare the substrates of Lon under TTSS permissive or non-permissive conditions. Many key regulators of *Yersinia* virulence may become Lon substrate in one condition but remain resistant in the other. Or the amount of a certain protein degraded by Lon may fluctuate under different conditions. The trapping experiment performed

in this study is more qualitative than quantitative and therefore is not revealing enough in terms of the difference in the amount of substrates under TTSS permissive or non-permissive conditions. Quantitative mass spectrometry can be used to obtain further information in this aspect.

In addition to the apparent virulence factors and transcriptional regulators identified, there are also many metabolic enzymes, which include key enzymes in various metabolic pathways that are essential for the survival and fitness of the bacteria. In recent years, more and more metabolic pathways have been shown to be involved in bacterial pathogenesis. Further investigation is needed to elucidate this indirect regulatory role of Lon in *Yersinia* pathogenesis.

4.5 Materials and Methods

4.5.1 Bacterial strains and media

All *Yersinia pestis* strains used in this study are *pgm*$^-$. The *Yersinia pestis* KIM5-3001 *lon*$^-$ strain, as well as its otherwise isogenic parental strain, are a generous gift from Dr. Plano. These *Y. pestis* strains were routinely grown in HI (Heart Infusion) media. To mimic the physiological conditions where the *Yersinia* TTSS is either suppressed or activated, different components were added to the HI broth. For TTSS non-permissive growth condition, the HI broth was supplemented with 2.5 mM calcium chloride. For TTSS permissive growth condition, the HI broth was supplemented with 20 mM sodium oxalate and 20 mM magnesium chloride. *E. coli* strain XL-1 blue was used for routine cloning. *E. coli* strain BL21(DE3) was used for the overexpression of *Yersinia* Lon protease. *E. coli* strain S17(λ-Pir) was used for conjugation. All *E. coli* strains were grown in standard LB broth. Antibiotics were supplemented into the media when needed at a final concentration of: 100 μg/ml for Ampicillin, 50 μg/ml for Kanamycin, and 30 μg/ml for chloramphenicol.

4.5.2 Plasmids and constructs

The gene encoding *Yp*-Lon protease was cloned into pMMB67EH vector under an IPTG inducible promoter. The Myc_3-His_6 affinity tag was introduced on an oligonucleotide cassette to the C-terminus of *Yp*-Lon. The C-terminal affinity tag sequence is: HHHHHHSGEQKLISEEDLVDNGEQKLISEEDLNGEQKLISEEDLN. The active site mutation S679A was introduced by site directed mutagenesis. The gene encoding *Yp*-Lon protease was also cloned into pET28b vector for overexpression and purification. The plasmids over-expressing *Ec*-Lon, *Mp*-Lon and λ-cI-N-$ssrA_{MP}$ were described in Ge *et al*. (2009). All potential *Yp*-Lon substrates were cloned into pDEST17 (Invitrogen) under an IPTG inducible

promoter for over-expression. All potential *Ec*-Lon substrates were from the ASKA library [216].

4.5.3 Lon substrates trapping *in vivo* and MS identification of trapped proteins

Y. pestis strains harboring the plasmid pMMB67EH-*Yp*-Lon^{trap} were grown at 26°C in 750ml of either TTSS permissive or non-permissive media to A_{260} of 0.7. Cells grown in TTSS permissive media were then switched to 37°C to allow the activation of TTSS components. Growth was continued for another hour for all samples. The expression of *Yp*-Lon^{trap} was induced by including 0.5 mM IPTG in the media from the beginning of growth. The cells were then harvested by centrifugation at 3,700 x g for one hour. The cell pellets were stored at -80°C for future use.

Each cell pellet was resuspended in 10 ml of Buffer A containing: 20 mM Tris-HCl pH 8.0, 300 mM K-Glutamate, 10 mM $MgCl_2$, and 1 mM β-ME. A cocktail of protease inhibitors (pierce) was added to the cell suspension prior to cell lysis. Cells were gently lysed by passing through a French Press. The crude cell lysates were spun at 10,000 x g for 30 minutes to remove cell debris. The cleared cell lysates were mixed with 1ml Ni-NTA slurry (sigma) pre-equilibrated in Buffer A, and incubated at 4°C for 2 hours to allow binding of *Yp*-Lon to the resin. The resin was then washed with Buffer A and the bound proteins were eluted in 3-6 ml of Buffer B containing: 20 mM Tris-HCl pH 8.0, 300 mM K-Glutamate, 10 mM MgCl2, 1 mM β-ME and 200 mM imidazole. The Myc antibody affinity resin was generated by cross-linking anti-myc antibody to Affi-gel Hz hydrazide gel (Biorad). The eluent from Ni-NTA resin was buffer exchanged into Buffer C containing: 20 mM Tris-HCl pH 8.0, 300 mM K-Glutamate, and 10 mM $MgCl_2$, and mixed with the Myc affinity resin pre-equilibrated in Buffer C. The mixture was incubated at 4°C for 2 hours to allow binding of *Yp*-Lon to the resin. The resin was then washed with Buffer C and the bound proteins were eluted in 3-6ml of 200 mM glycine pH 2.5. The eluent was then dialyzed into PBS and stored at -20°C for future use.

4.5.4 Protein expression and purification

E. coli strain BL21(DE3) cells, harboring the plasmid pET28b-*Yp*-Lon, were grown in LB broth at 37°C to A_{260} of 0.7. The expression of *Yp*-Lon proteases was induced by the addition of IPTG to a final concentration of 1mM. Cells were allowed to grow for an additional 16 hours at 16°C. Cells were harvested by centrifugation at 3,700 x g for one hour. The cell pellets were stored at -80°C for future use. Cell pellets containing *Yp*-Lon protease were resuspended in Buffer D containing: 50 mM $KHPO_4$ pH 6.9, 10% glycerol, and 1 mM DTT. Cells were lysed by sonication. The crudes cell lysates were then spun

at 10,000 x g for 30 minutes to remove cell debris. 5 g of the P11-cellulose resin was activated according to the instruction manual, equilibrated in Buffer D, and loosely packed into a gravity-flow column. The cleared cell lysate was then passed through the column twice by gravity to allow binding of *Yp*-Lon protease to the resin. The column was washed with Buffer D to remove unbound proteins. The bound *Yp*-Lon was eluted in 10 ml of Buffer E containing: 250 mM $KHPO_4$ pH 6.9, 10% glycerol, and 1 mM DTT. The resulted protein was exchanged into buffer E containing: 20 mM $KHPO_4$ pH 6.9, 50 mM KCl, 10% glycerol, and 1 mM DTT. The Mono-Q HR 10/10 column (GE healthcare) was equilibrated in Buffer E on the AKTA FPLC system (GE healthcare) and the partially purified *Yp*-Lon protease was loaded onto the column. The column was washed with 20 CV of Buffer E and the bound protein was eluted by the application of a linear KCl gradient from 0% buffer F to 100% Buffer F, containing 20 mM $KHPO_4$ pH 6.9, 1 M KCl, 10% glycerol, and 1 mM DTT. Fractions containing *Yp*-Lon protease were pooled, concentrated, aliquoted, and stored at -80°C for future use. The expression and purification procedure for *Ec*-Lon and *Mp*-Lon is as described in Ge *et al.* (2009).

E. coli strain BL21(DE3) harboring the plasmids expressing each of the potential *Yp*-Lon substrate was grown in 6L of LB broth at 37°C to A_{260} of 0.7. The expression of the target protein was induced by the addition of IPTG to a final concentration of 1 mM. Cells were allowed to grow for an additional 16 hours at 16°C. Cells were harvested by centrifugation at 3,700 x g for one hour. The cell pellets were stored at -80°C for future use. Cell pellets containing each of the potential *Yp*-Lon substrates were resuspended in Buffer F containing: 50 mM $KHPO_4$ pH6.9, 500 mM KCl, 1 mM EDTA, and 2 mM β-ME. Cells were lysed by sonication. The crudes cell lysate were then spun at 10,000 x g for 30 minutes to remove cell debris. The cleared cell lysates were mixed with 2 ml Ni-NTA slurry (sigma) pre-equilibrated in Buffer F, and incubated at 4°C for 2 hours to allow binding of the targeted protein to the resin. The resin was then washed with Buffer A and the bound proteins were eluted in 25 ml of Buffer G containing: 50 mM $KHPO_4$ pH6.9, 500 mM KCl, 1 mM EDTA, 2 mM β-ME, and 200 mM imidazole. The eluted protein was concentrated and buffer changed into Buffer E. Either the Mono-Q HR 10/10 column (GE healthcare) or the Mono-S HR 10/10 column (GE healthcare) column was used, depending on the pI of the particular protein being purified. The Mono-Q/Mono-S HR 10/10 column (GE healthcare) was equilibrated in Buffer E on the AKTA FPLC system (GE healthcare) and the partially purified protein was loaded onto the column. The column was washed with 20 CV of Buffer E and the bound protein was eluted by the application of a linear KCl gradient from 0% buffer F to 100% Buffer F. Fractions containing pure target protein were pooled, concentrated, aliquoted, and stored at -80°C for future use.

For the overexpression of *E. coli* S2 and L9, the clone containing the expression construct of the target protein was selected from the ASKA library [216] and was grown in 6 L of LB broth at 37°C to A_{260} of 0.7. The expression of

the target protein was induced by the addition of IPTG to a final concentration of 1 mM. Cells were allowed to grow for an additional 16 hours at 16°C. Cells were harvested by centrifugation at 3,700 x g for one hour. The cell pellets were stored at -80°C for future use. Cell pellets containing either S2 or L9 were resuspended in Buffer G containing: 50 mM $KHPO_4$ pH6.9, 500 mM KCl, 1 mM EDTA, 4 M urea, and 2 mM β-ME. Cells were lysed by sonication. The crudes cell lysate were then spun at 10,000 x g for 30 minutes to remove cell debris. The cleared cell lysates were mixed with 2 ml Ni-NTA slurry (sigma) pre-equilibrated in Buffer G and incubated at 4°C for 2 hours to allow binding of the targeted protein to the resin. The resin was then washed with Buffer G and the bound proteins were eluted in 25 ml of Buffer I containing: 50 mM $KHPO_4$ pH6.9, 500 mM KCl, 1 mM EDTA, 4 M urea, 2 mM β-ME, and 200 mM imidazole. The eluted protein was concentrated and buffer changed into Buffer E. The Mono-Q HR 10/10 column (GE healthcare) was equilibrated in Buffer E on the AKTA FPLC system (GE healthcare) and the partially purified protein was loaded onto the column. The column was washed with 20 CV of Buffer E and the bound protein was eluted by the application of a linear KCl gradient from 0% buffer F to 100% Buffer F. Fractions containing pure target protein were pooled, concentrated, aliquoted, and stored at -80°C for future use.

4.5.5 *In vitro* proteolysis assay

In vitro proteolysis assay was carried out in an 80 µl reaction mixture containing Lon activity buffer (50 mM Tris-HCl pH8.0, 10 mM $MgCl_2$, 1 mM DTT, and 10% glycerol), ATP regeneration system (50 mM creatine phosphate, 80 µg/ml creatine kinase, and 4 mM ATP), 5 µM substrate, and Lon protease (either 100 nM *Mp*-Lon_6, or 200 nM *Yp*-Lon_6, or 400 nM *Ec*-Lon_6). The reaction was assembled and incubated at 30°C or 37°C. Aliquots were taken at designated time points and the reaction stopped by adding equal volume of 2XSDS-PAGE sample buffer.

4.5.6 Sample preparation for Mass Spectrometry analysis

To prepare the MS samples for the identification of the trapped proteins, 2 mls of the protein solution containing a mixture of *Yp*-Lon^{trap} and its co-purified proteins were subjected to TCA precipitation in order to concentrate the proteins. 400µl of PBS was added to the protein mixture, followed by the addition of 600µl of 100% TCA (w/v). The mixture was then incubated on ice for 30min to allow the intact proteins to precipitate. The mixture was spun at 10,000 x g for 15 minutes to separate the precipitated proteins and the supernatant. The supernatant was discarded and the resulting protein pellet was washed with ice-cold acetone twice. The pellet was then dried and sent for mass spectrometry analysis.

To prepare the MS samples for the analysis of *Yp*-Lon degradation products, 100 µl *in vitro* proteolysis reaction mixture was assembled as described above and incubated at 37 °C for a period of time that is sufficient for the reaction to go to completion. Upon completion of the reaction, 20 µl of 10 mg/ml BSA was added to the reaction, followed by the addition of 30 µl of 100% TCA (w/v). The mixture was then incubated on ice for 30min to allow the intact proteins to precipitate. The mixture was spun at 10,000 x g for 15 minutes to separate the precipitated proteins and the supernatant. The resulting supernatant was collected and stored at -20°C until subjected to MS analysis.

4.6 Figures

Figure 4.1

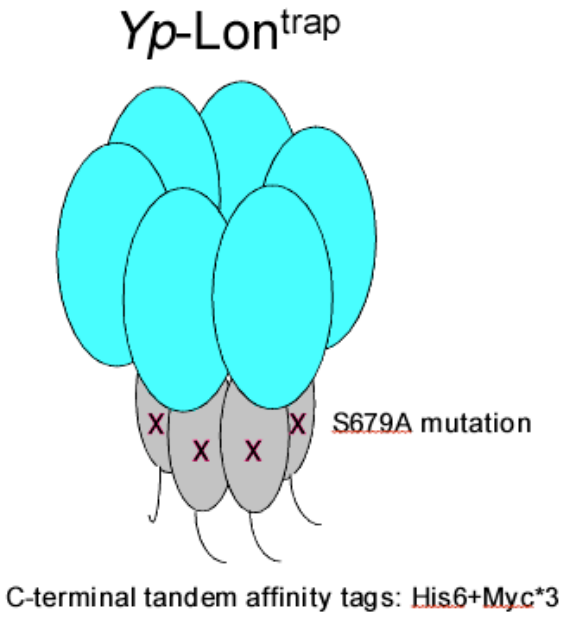
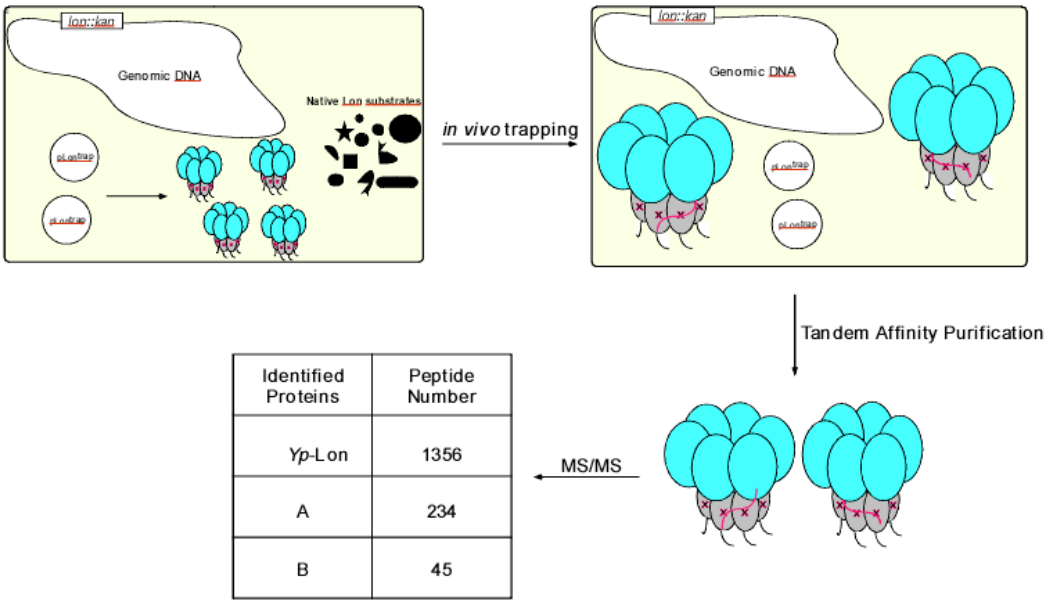


Figure 4.1: Designing of *Yp*-Lon^{trap}. The S679A mutation renders the protease inactive in peptide bond cleavage. However, the abilities of the protease to oligomerize, to recognize, bind, unfold, and translocate its native substrates, and to hydrolyze ATP, are intact. To facilitate purification, the *Yp*-Lon^{trap} also carries a tandem affinity tag at its C-terminus, consisting of a His6 tag and three myc tags.

Figure 4.2



Identified Proteins	Peptide Number
Yp-Lon	1356
A	234
B	45

Figure 4.2: *In vivo* trapping of the native substrates of Lon protease in *Yersinia pestis*. The expression of *Yp*-Lon^{trap} is induced to near endogenous level in a *Yersinia pestis* KIM5 *lon*$^{-}$ strain. The *Yp*-Lon^{trap} recognizes and traps its native substrates during cell growth. The protease-substrate complexes are then isolated using tandem affinity purification. Finally the substrates that co-purify with *Yp*-Lon^{trap} are identified by LC-MS-MS.

Figure 4.3

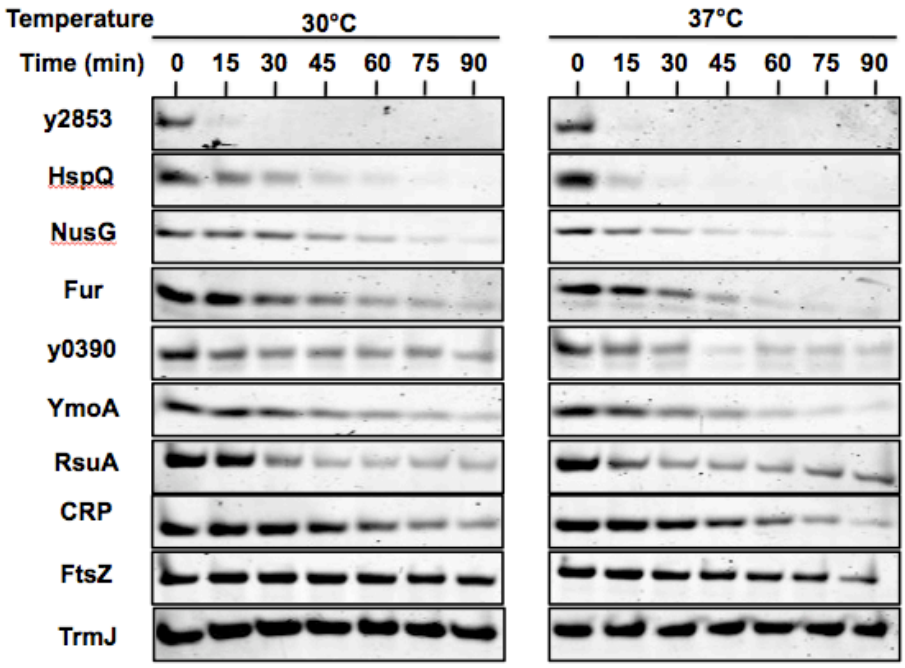


Figure 4.3: *In vitro* proteolysis of potential *Yp*-Lon substrates. Substrates included in this assay are y2853, HspQ, NusG, Fur, y0390, YmoA, RsuA, CRP, FtsA, and TrmJ. *In vitro* proteolysis assays were carried out at 30°C or 37°C in minimum Lon activity buffer containing 50 mM Tris-HCl pH8.0, 10 mM $MgCl_2$, 1 mM DTT, and 10% glycerol. The reactions also contain 200 nM *Yp*-Lon_6, 5 μM substrate, and an ATP regeneration system. Aliquots were taken at designated time points, quenched with an equal volume of 2x SDS-sample buffer, resolved by electrophoresis on 15% Tris-tricine gel and stained with Coomassie Brilliant BlueR250.

Figure 4.4

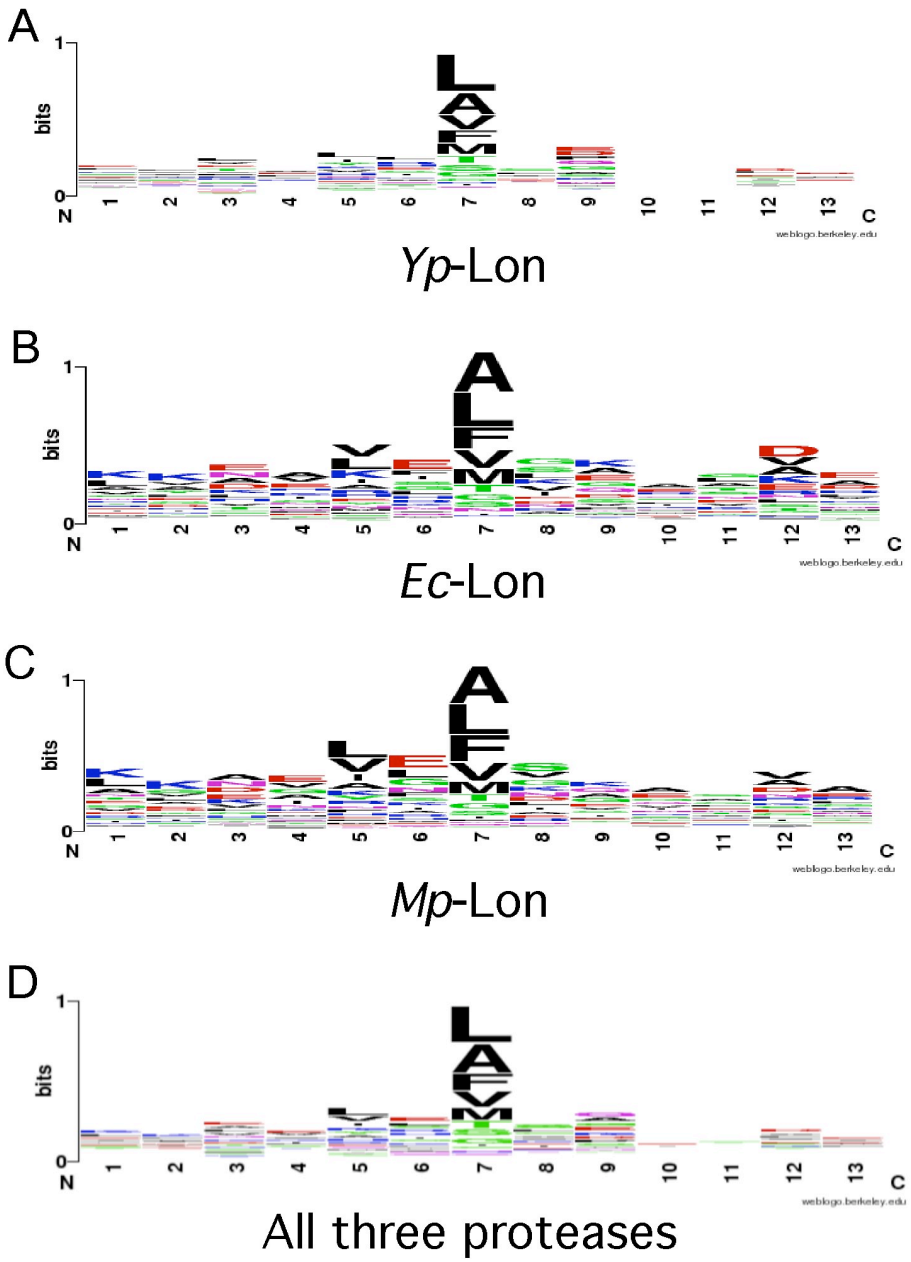
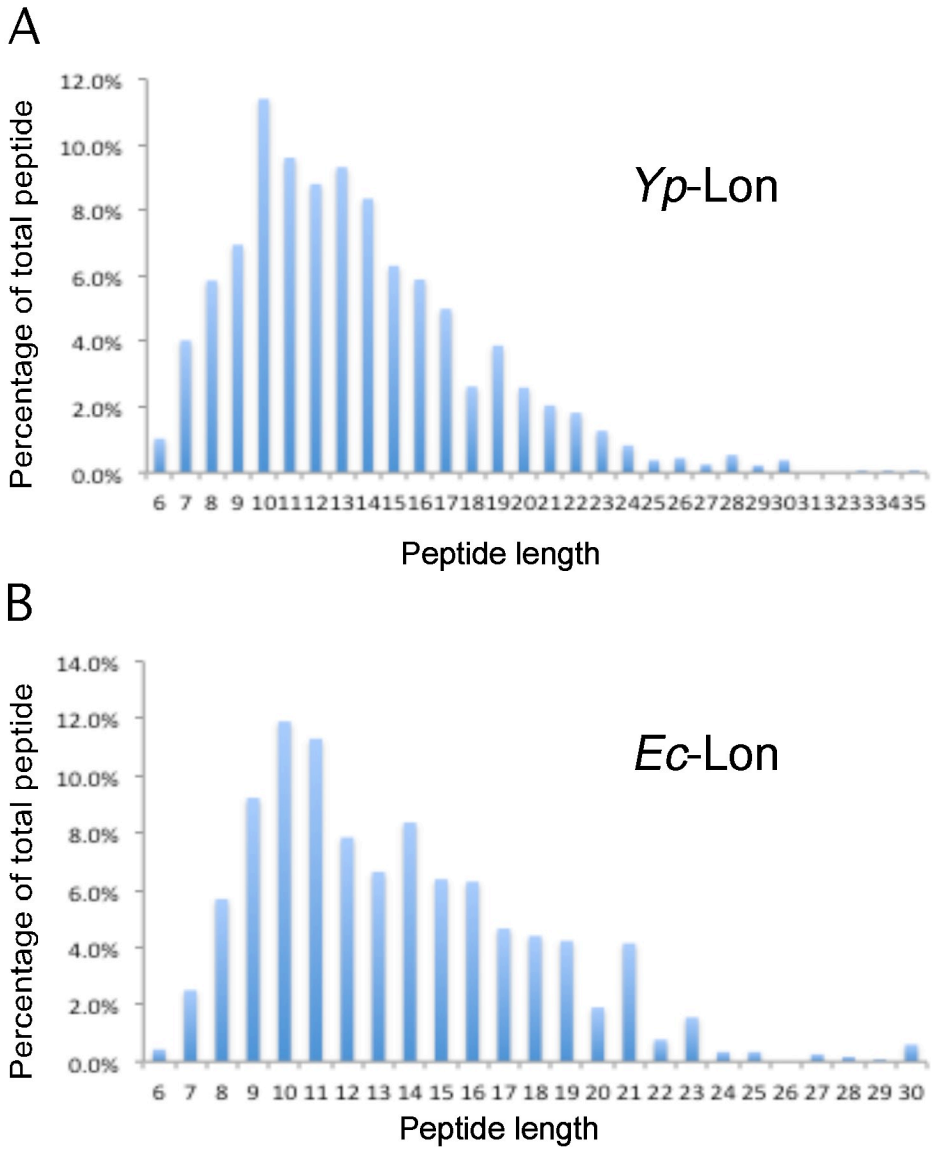


Figure 4.4: Preferred cleavage sites by Lon proteases. The peptides generated from digestion by Lon protease were identified by mass spectrometry. The last amino acid of the identified peptide is the amino acid that Lon cleaves after. After determining the cleavage site, sequences of 6 amino acids both upstream and downstream of the cleavage sites were extracted. A database was constructed from the sequences, as well as the number that a certain spectrum was identified from each position. The database was then analyzed at http://weblogo.berkeley.edu/logo.cgi for the degree of conservation at the cleavage site by A) *Yp*-Lon, B) *Ec*-Lon, C) *Mp*-Lon, and D) All three proteases.

Figure 4.5


A
Percentage of total peptide
12.0%
10.0%
8.0%
6.0%
4.0%
2.0%
0.0%
Yp-Lon
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35
Peptide length
B
Percentage of total peptide
14.0%
12.0%
10.0%
8.0%
6.0%
4.0%
2.0%
0.0%
Ec-Lon
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Peptide length

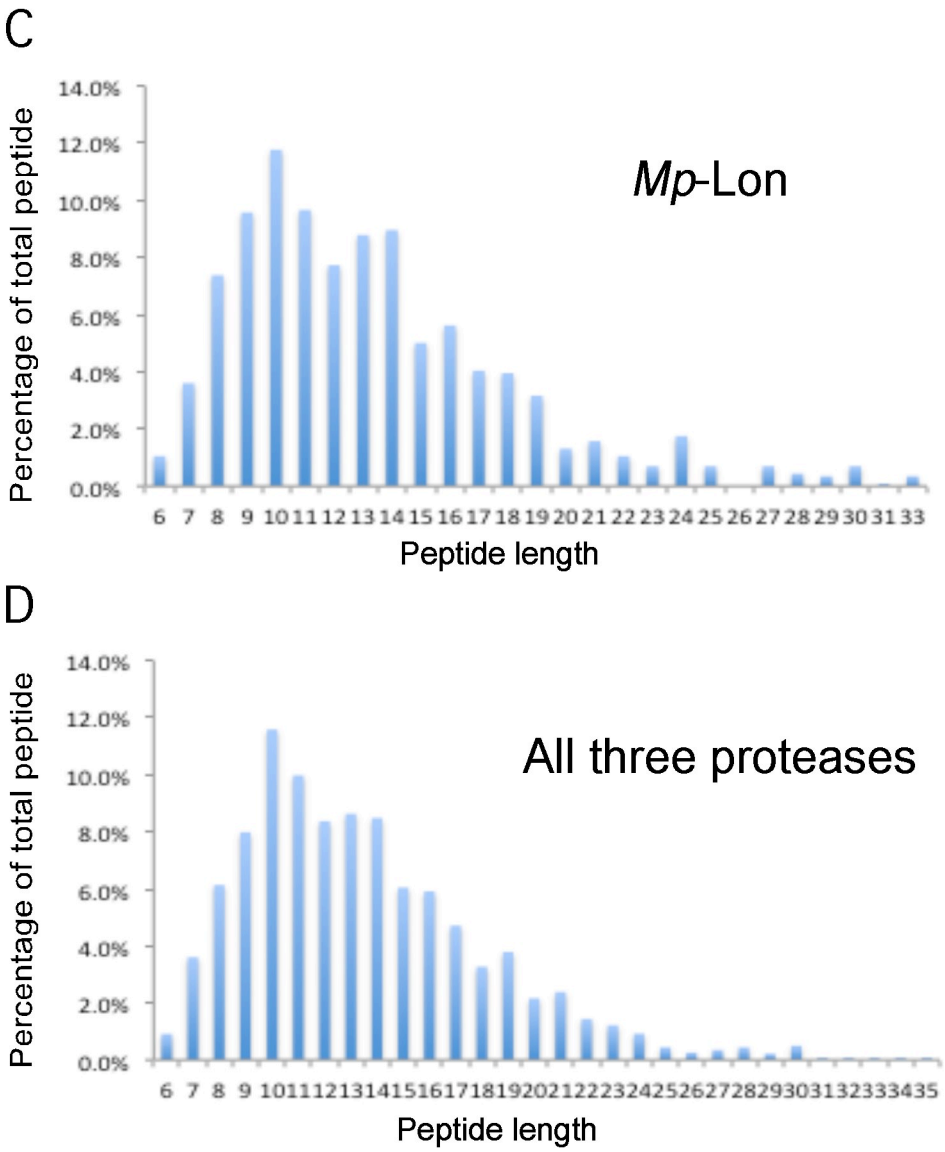


Figure 4.5: Peptide length distribution of Lon degradation products: A) *Yp*-Lon, B) *Ec*-Lon, C) Mp-Lon, and D) All three proteases. Mass spectrometry identified peptides ranging from a length of 6 amino acids up to 35 amino acids. The distribution of the peptide lengths was analyzed by calculating the percentage of the peptide of a certain length in total peptide. Results showed that most peptides are in 8-17 amino acid range.

Table 4.1

y2853	**Putative sensory transduction regulator**
y3621	**Tubulin-like GTP-binding protein and GTPase (Cell division protein FtsZ)**
y2917	**Ribosomal small subunit pseudouridine synthase A**
y3956	**cyclic AMP receptor protein**
y0479	**NusG**
y1208	**Fur, ferric uptake regulator**
y1332	**TrmJ, tRNA (cytidine/uridine-2'-O-)-methyltransferase**
y2726	**HspQ**
y0390	**transcriptional regulator**
YPMT1.74	***Yersinia* murine toxin (*ymt*)**
YPCD1.09c	**uncharacterized protein**
y0965	**Sigma factor-binding protein crl**
y2287	**Putative transcriptional regulator**
y4025	**Smg**
y3999	**RplN, 50S ribosomal protein L14**
y3622	**ATP-binding cell division protein (Cell division protein FtsA)**
y3249	**Predicted transcriptional regulators**
y1324	**Attachment invasion locus protein (*ail*)**
YPKp07	**Plasminogen activator protease (*pla*)**
Y3668	**Hypothetical protein vgrG5**
y3674	**Hypothetical protein**

Table 4.1: Example list of *Yp*-Lon substrates identified by LC-MS-MS.

Chapter 5 Concluding Remarks

5.1 Summary

This dissertation has provided new insights in two major aspects of the *trans*-translation process: the *trans*-translation mediated non-stop mRNA decay, and the recognition motifs in the tmRNA tag for Lon protease. In addition, the study on Lon protease has also been extended to the general substrate recognition and degradation mechanism of this protease, as well as a comprehensive profiling of its native substrates in *Yersinia pestis*.

To be more specific, in Chapter 2 I studied the structure-function correlation of RNase R. I determined that the C-terminal Lysine-rich region of RNase R is required for the non-stop mRNA decay activity, and this activity is also mediated by the tmRNA-SmpB machinery. The results suggest that during *trans*-translation, when the ribosome disengages the non-stop mRNA and resumes translation on the tmRNA ORF, the non-stop mRNA is not released entirely to the cytosol before it is bound and degraded by RNase R. To the contrary, RNase R is brought to the vicinity of the non-stop mRNA through the interaction of the C-terminal domain of RNase R and the *trans*-translation machinery and is likely to be loaded onto the non-stop mRNA before it is released from the ribosome. This also suggests that the ribosome, in addition to being the factory for protein synthesis, can also serve as the platform of other important biological processes.

In Chapter 3, we explored a unique case in *Mycoplasma* where the Clp family proteases are absent while the tmRNA-SmpB function is not only conserved but also essential for the survival for the bacteria. In the Mycoplasma species, Lon becomes the primary protease that is responsible for degrading the tmRNA tagged proteins. In order to fulfill its function, the *Mp*-Lon and *Mp*-tmRNA tag co-evolved. *Mp*-Lon acquired more robust activity against the *Mp*-tmRNA tagged proteins, meanwhile the *Mp*-tmRNA tag acquired additional sequence feature that are proven to enhance Mp-Lon binding in order to allow more efficient degradation of the tagged proteins. We also found that although *Mp*-Lon and *Ec*-Lon share a lot of similarities in their substrate specificity, there are also a lot of differences that probably occurred during evolution. For instance, both Lon proteases selectively degrade tmRNA tagged proteins, as well as *Ec*-Lon native substrates SulA and RcsA. *Mp*-Lon exhibited much lower affinity towards SulA and RcsA than *Ec*-Lon, probably because SulA and RcsA are both absent in *Mycoplasma*. *Mp*-Lon is much more efficient in degrading the tmRNA tagged proteins because the tmRNA tagged proteins are mainly degraded by Lon in *Mycoplasma*.

In Chapter 4, we used a combination of biochemical and proteomic techniques to profile the native substrates of Lon protease in *Yersinia pestis*. Many novel substrates were identified, ranging from transcriptional regulators,

cell division factors, metabolic enzymes, as well as several *Yersinia* virulence factors. This study provided insights into the role of Lon protease in *Yersinia* physiology and pathogenesis. In addition, the identification of a large number of native substrates allowed us to do analysis, which identified the preferred cleavages sites of Lon protease.

5.2 Remaining questions

5.2.1 RNase R

The current model suggests that the initiation of non-stop mRNA decay occurs on the ribosome. It seems that before the stalled ribosome resumes translation on the tmRNA ORF and releases the non-stop mRNA, RNase R is recruited to the stalled ribosome through the activity of tmRNA-SmpB system, so that it can be in close vicinity of the non-stop mRNA to initiate the decay process. The detailed mechanism of how RNase R is loaded onto the non-stop mRNA, as well as where RNase R is located on the stalled ribosome during this process remains unknown. However, it is conceivable that RNase R may be located close to the A site of the ribosome since that is where the tmRNA-SmpB complex is accommodated.

In addition, we have solid evidence showing that the recruitment of RNase R to stalled ribosomes requires active *trans*-translation, since in the absence of tmRNA-SmpB RNase R suffers a defect in the recruitment. However, it is still unclear which factor(s) in the *trans*-translation machinery is responsible for the recruitment of RNase R. It has been reported that RNase R interacts with SmpB via the C-terminal lysine rich domain *in vitro*, which may indicate that it is SmpB that recruits RNase R during *trans*-translation. There is also evidence that certain mutations in the SmpB C-terminal domain, as well as in the tmRNA ORF would abolish the non-stop mRNA decay activity of RNase R (108), which may indicate that these mutations somehow affect the interaction between tmRNA-SmpB and RNase R, thus resulting the inability of RNase R to be recruited to stalled ribosomes. So far we have not ruled out the possibility that factors other than tmRNA and SmpB may be responsible for recruiting RNase R during *trans*-translation.

5.2.2 Lon protease

Although we have identified distinct regions of the *Mp*-tmRNA tag that are critical for Lon recognition, there are approximately 10 amino acids at the N-terminus of the tag that are not directly involved in recognition by Lon. It is plausible that these sequences serve as the binding sites for auxiliary cofactors in a manner analogous to SspB-ClpXP or ClpS-ClpAP interactions [102-104, 177, 179, 187-196]. To date, however, adaptor proteins have not been reported

for Lon protease analogs. Nontheless, it is conceivable that Lon, like other AAA+ proteases, possesses a cofactor(s) that regulates its substrate range and specificity. Alternatively, the N-terminal part of the *Mp*-tmRNA tag might carry signals for recognition by the membrane-associated FtsH protease since FtsH is also conserved in *Mycoplasma*. Future studies are required to explore these possibilities.

We are also interested in investigating the mechanistic features of Lon protease, such as identifying the residues in Lon that are recognize and bind the substrates, studying the interaction and communication between the ATPase domain and the protease domain, as well as calculating the amount of energy needed for Lon to degrade its substrates. The crystal structure of Lon protease with a translocating substrate bound would provide significant insight into the mechanical movement of the ATPase domain, as well as its interaction with the N-terminal or protease domain.

In addition, with the large pool of substrates we have identified in *Yersinia pestis*, we can analyze the primary sequence of the substrates and hopefully identify the consensus recognition sequence of Lon protease. The detailed investigation of a certain substrate may reveal the biological importance of the turnover of the protein, and thus illustrate the function of Lon protease in maintaining the fitness of the bacteria.

Bibliography

1. Steitz, T.A., *A structural understanding of the dynamic ribosome machine.* Nat Rev Mol Cell Biol, 2008. **9**(3): p. 242-53.
2. Ramakrishnan, V., *Ribosome structure and the mechanism of translation.* Cell, 2002. **108**(4): p. 557-72.
3. Bashan, A., et al., *Ribosomal crystallography: peptide bond formation and its inhibition.* Biopolymers, 2003. **70**(1): p. 19-41.
4. Zucker, F.H. and J.W. Hershey, *Binding of Escherichia coli protein synthesis initiation factor IF1 to 30S ribosomal subunits measured by fluorescence polarization.* Biochemistry, 1986. **25**(12): p. 3682-90.
5. Gualerzi, C.O., et al., *Initiation factors in the early events of mRNA translation in bacteria.* Cold Spring Harb Symp Quant Biol, 2001. **66**: p. 363-76.
6. Karimi, R., et al., *Initiation factors IF1 and IF2 synergistically remove peptidyl-tRNAs with short polypeptides from the P-site of translating Escherichia coli ribosomes.* J Mol Biol, 1998. **281**(2): p. 241-52.
7. Laursen, B.S., et al., *Initiation of protein synthesis in bacteria.* Microbiol Mol Biol Rev, 2005. **69**(1): p. 101-23.
8. Petrelli, D., et al., *Mapping the active sites of bacterial translation initiation factor IF3.* J Mol Biol, 2003. **331**(3): p. 541-56.
9. Petrelli, D., et al., *Translation initiation factor IF3: two domains, five functions, one mechanism?* EMBO J, 2001. **20**(16): p. 4560-9.
10. Hartz, D., et al., *Domains of initiator tRNA and initiation codon crucial for initiator tRNA selection by Escherichia coli IF3.* Genes Dev, 1990. **4**(10): p. 1790-800.
11. Gualerzi, C.O. and C.L. Pon, *Initiation of mRNA translation in prokaryotes.* Biochemistry, 1990. **29**(25): p. 5881-9.
12. Green, R. and H.F. Noller, *Ribosomes and translation.* Annu Rev Biochem, 1997. **66**: p. 679-716.
13. Valle, M., et al., *Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy.* Nat Struct Biol, 2003. **10**(11): p. 899-906.
14. Schmeing, T.M., et al., *An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA.* Nature, 2005. **438**(7067): p. 520-4.
15. Selmer, M., et al., *Crystallization and preliminary X-ray analysis of Thermotoga maritima ribosome recycling factor.* Acta Crystallogr D Biol Crystallogr, 1999. **55**(Pt 12): p. 2049-50.
16. Schuwirth, B.S., et al., *Structures of the bacterial ribosome at 3.5 A resolution.* Science, 2005. **310**(5749): p. 827-34.
17. Rodnina, M.V., et al., *Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome.* Nature, 1997. **385**(6611): p. 37-41.
18. Ito, K., M. Uno, and Y. Nakamura, *A tripeptide 'anticodon' deciphers stop codons in messenger RNA.* Nature, 2000. **403**(6770): p. 680-4.
19. Kisselev, L.L. and R.H. Buckingham, *Translational termination comes of age.* Trends Biochem Sci, 2000. **25**(11): p. 561-6.

20. Zavialov, A.V., R.H. Buckingham, and M. Ehrenberg, *A posttermination ribosomal complex is the guanine nucleotide exchange factor for peptide release factor RF3.* Cell, 2001. **107**(1): p. 115-24.
21. Gong, M., L.R. Cruz-Vera, and C. Yanofsky, *Ribosome recycling factor and release factor 3 action promotes TnaC-peptidyl-tRNA Dropoff and relieves ribosome stalling during tryptophan induction of tna operon expression in Escherichia coli.* J Bacteriol, 2007. **189**(8): p. 3147-55.
22. Janosi, L., et al., *Ribosome recycling by ribosome recycling factor (RRF)--an important but overlooked step of protein biosynthesis.* Adv Biophys, 1996. **32**: p. 121-201.
23. Janosi, L., R. Ricker, and A. Kaji, *Dual functions of ribosome recycling factor in protein biosynthesis: disassembling the termination complex and preventing translational errors.* Biochimie, 1996. **78**(11-12): p. 959-69.
24. Gao, N., et al., *Mechanism for the disassembly of the posttermination complex inferred from cryo-EM studies.* Mol Cell, 2005. **18**(6): p. 663-74.
25. Himeno, H., et al., *Escherichia coli tmRNA (10Sa RNA) in trans-translation.* Nucleic Acids Symp Ser, 1997(37): p. 185-6.
26. Himeno, H., et al., *In vitro trans translation mediated by alanine-charged 10Sa RNA.* J Mol Biol, 1997. **268**(5): p. 803-8.
27. Komine, Y., M. Kitabatake, and H. Inokuchi, *10Sa RNA is associated with 70S ribosome particles in Escherichia coli.* J Biochem, 1996. **119**(3): p. 463-7.
28. Komine, Y., et al., *A tRNA-like structure is present in 10Sa RNA, a small stable RNA from Escherichia coli.* Proc Natl Acad Sci U S A, 1994. **91**(20): p. 9223-7.
29. Muto, A., et al., *Structure and function of 10Sa RNA: trans-translation system.* Biochimie, 1996. **78**(11-12): p. 985-91.
30. Nameki, N., et al., *Amino acid acceptor identity switch of Escherichia coli tmRNA from alanine to histidine in vitro.* J Mol Biol, 1999. **289**(1): p. 1-7.
31. Nameki, N., et al., *An NMR and mutational analysis of an RNA pseudoknot of Escherichia coli tmRNA involved in trans-translation.* Nucleic Acids Res, 1999. **27**(18): p. 3667-75.
32. Nameki, N., et al., *Functional and structural analysis of a pseudoknot upstream of the tag-encoded sequence in E. coli tmRNA.* J Mol Biol, 1999. **286**(3): p. 733-44.
33. Nameki, N., et al., *Three of four pseudoknots in tmRNA are interchangeable and are substitutable with single-stranded RNAs.* FEBS Lett, 2000. **470**(3): p. 345-9.
34. Nonin-Lecomte, S., B. Felden, and F. Dardel, *NMR structure of the Aquifex aeolicus tmRNA pseudoknot PK1: new insights into the recoding event of the ribosomal trans-translation.* Nucleic Acids Res, 2006. **34**(6): p. 1847-53.
35. Tanner, D.R., et al., *Genetic analysis of the structure and function of transfer messenger RNA pseudoknot 1.* J Biol Chem, 2006. **281**(15): p. 10561-6.
36. Dulebohn, D.P., H.J. Cho, and A.W. Karzai, *Role of conserved surface amino acids in binding of SmpB protein to SsrA RNA.* J Biol Chem, 2006. **281**(39): p. 28536-45.
37. Karzai, A.W., M.M. Susskind, and R.T. Sauer, *SmpB, a unique RNA-binding protein essential for the peptide-tagging activity of SsrA (tmRNA).* EMBO J, 1999. **18**(13): p. 3793-9.

38. Sundermeier, T.R. and A.W. Karzai, *Functional SmpB-ribosome interactions require tmRNA.* J Biol Chem, 2007. **282**(48): p. 34779-86.
39. Dong, G., J. Nowakowski, and D.W. Hoffman, *Structure of small protein B: the protein component of the tmRNA-SmpB system for ribosome rescue.* EMBO J, 2002. **21**(7): p. 1845-54.
40. Gutmann, S., et al., *Crystal structure of the transfer-RNA domain of transfer-messenger RNA in complex with SmpB.* Nature, 2003. **424**(6949): p. 699-703.
41. Bessho, Y., et al., *Structural basis for functional mimicry of long-variable-arm tRNA by transfer-messenger RNA.* Proc Natl Acad Sci U S A, 2007. **104**(20): p. 8293-8.
42. Cheng, K., et al., *tmRNA.SmpB complex mimics native aminoacyl-tRNAs in the A site of stalled ribosomes.* J Struct Biol, 2010. **169**(3): p. 342-8.
43. Sundermeier, T.R., et al., *A previously uncharacterized role for small protein B (SmpB) in transfer messenger RNA-mediated trans-translation.* Proc Natl Acad Sci U S A, 2005. **102**(7): p. 2316-21.
44. Barends, S., et al., *Simultaneous and functional binding of SmpB and EF-Tu-TP to the alanyl acceptor arm of tmRNA.* J Mol Biol, 2001. **314**(1): p. 9-21.
45. Shimizu, Y. and T. Ueda, *SmpB triggers GTP hydrolysis of elongation factor Tu on ribosomes by compensating for the lack of codon-anticodon interaction during trans-translation initiation.* J Biol Chem, 2006. **281**(23): p. 15987-96.
46. Williams, K.P. and D.P. Bartel, *Phylogenetic analysis of tmRNA secondary structure.* RNA, 1996. **2**(12): p. 1306-10.
47. Keiler, K.C., L. Shapiro, and K.P. Williams, *tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: A two-piece tmRNA functions in Caulobacter.* Proc Natl Acad Sci U S A, 2000. **97**(14): p. 7778-83.
48. Kovacs, L., et al., *Cloning, expression and purification of SmpB from Mycobacterium tuberculosis.* Acta Microbiol Immunol Hung, 2004. **51**(3): p. 297-302.
49. Sharkady, S.M. and K.P. Williams, *A third lineage with two-piece tmRNA.* Nucleic Acids Res, 2004. **32**(15): p. 4531-8.
50. Huang, C., et al., *Charged tmRNA but not tmRNA-mediated proteolysis is essential for Neisseria gonorrhoeae viability.* EMBO J, 2000. **19**(5): p. 1098-107.
51. de la Cruz, J. and A. Vioque, *Increased sensitivity to protein synthesis inhibitors in cells lacking tmRNA.* RNA, 2001. **7**(12): p. 1708-16.
52. Vioque, A. and J. de la Cruz, *Trans-translation and protein synthesis inhibitors.* FEMS Microbiol Lett, 2003. **218**(1): p. 9-14.
53. Konno, T., et al., *Initiation-shift of trans-translation by aminoglycosides.* Nucleic Acids Symp Ser (Oxf), 2004(48): p. 299-300.
54. Luidalepp, H., et al., *tmRNA decreases the bactericidal activity of aminoglycosides and the susceptibility to inhibitors of cell wall synthesis.* RNA Biol, 2005. **2**(2): p. 70-4.
55. Okan, N.A., J.B. Bliska, and A.W. Karzai, *A Role for the SmpB-SsrA system in Yersinia pseudotuberculosis pathogenesis.* PLoS Pathog, 2006. **2**(1): p. e6.

56. Okan, N.A., et al., *The smpB-ssrA mutant of Yersinia pestis functions as a live attenuated vaccine to protect mice against pulmonary plague infection.* Infect Immun, 2010. **78**(3): p. 1284-93.
57. Withey, J.H. and D.I. Friedman, *The biological roles of trans-translation.* Curr Opin Microbiol, 2002. **5**(2): p. 154-9.
58. Karzai, A.W., E.D. Roche, and R.T. Sauer, *The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue.* Nat Struct Biol, 2000. **7**(6): p. 449-55.
59. Kapust, R.B., K.M. Routzahn, and D.S. Waugh, *Processive degradation of nascent polypeptides, triggered by tandem AGA codons, limits the accumulation of recombinant tobacco etch virus protease in Escherichia coli BL21(DE3).* Protein Expr Purif, 2002. **24**(1): p. 61-70.
60. Roche, E.D. and R.T. Sauer, *Identification of endogenous SsrA-tagged proteins reveals tagging at positions corresponding to stop codons.* J Biol Chem, 2001. **276**(30): p. 28509-15.
61. Abo, T., et al., *SsrA-mediated tagging and proteolysis of LacI and its role in the regulation of lac operon.* EMBO J, 2000. **19**(14): p. 3762-9.
62. Kushner, S.R., *mRNA decay in prokaryotes and eukaryotes: different approaches to a similar problem.* IUBMB Life, 2004. **56**(10): p. 585-94.
63. Deutscher, M.P., *Degradation of RNA in bacteria: comparison of mRNA and stable RNA.* Nucleic Acids Res, 2006. **34**(2): p. 659-66.
64. Mathy, N., et al., *5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA.* Cell, 2007. **129**(4): p. 681-92.
65. Kushner, S.R., *mRNA decay in Escherichia coli comes of age.* J Bacteriol, 2002. **184**(17): p. 4658-65; discussion 4657.
66. Mudd, E.A., H.M. Krisch, and C.F. Higgins, *RNase E, an endoribonuclease, has a general role in the chemical decay of Escherichia coli mRNA: evidence that rne and ams are the same genetic locus.* Mol Microbiol, 1990. **4**(12): p. 2127-35.
67. Babitzke, P. and S.R. Kushner, *The Ams (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of Escherichia coli.* Proc Natl Acad Sci U S A, 1991. **88**(1): p. 1-5.
68. Taraseviciene, L., A. Miczak, and D. Apirion, *The gene specifying RNase E (rne) and a gene affecting mRNA stability (ams) are the same gene.* Mol Microbiol, 1991. **5**(4): p. 851-5.
69. Schmeissner, U., et al., *Removal of a terminator structure by RNA processing regulates int gene expression.* J Mol Biol, 1984. **176**(1): p. 39-53.
70. Ow, M.C., T. Perwez, and S.R. Kushner, *RNase G of Escherichia coli exhibits only limited functional overlap with its essential homologue, RNase E.* Mol Microbiol, 2003. **49**(3): p. 607-22.
71. Deana, A. and J.G. Belasco, *The function of RNase G in Escherichia coli is constrained by its amino and carboxyl termini.* Mol Microbiol, 2004. **51**(4): p. 1205-17.
72. Perwez, T. and S.R. Kushner, *RNase Z in Escherichia coli plays a significant role in mRNA decay.* Mol Microbiol, 2006. **60**(3): p. 723-37.

73. Arnold, T.E., J. Yu, and J.G. Belasco, *mRNA stabilization by the ompA 5′ untranslated region: two protective elements hinder distinct pathways for mRNA degradation.* RNA, 1998. **4**(3): p. 319-30.
74. Mackie, G.A., *Ribonuclease E is a 5′-end-dependent endonuclease.* Nature, 1998. **395**(6703): p. 720-3.
75. Celesnik, H., A. Deana, and J.G. Belasco, *Initiation of RNA decay in Escherichia coli by 5′ pyrophosphate removal.* Mol Cell, 2007. **27**(1): p. 79-90.
76. Deana, A., H. Celesnik, and J.G. Belasco, *The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal.* Nature, 2008. **451**(7176): p. 355-8.
77. Deutscher, M.P., *Ribonuclease multiplicity, diversity, and complexity.* J Biol Chem, 1993. **268**(18): p. 13011-4.
78. Cannistraro, V.J. and D. Kennell, *The 5′ ends of RNA oligonucleotides in Escherichia coli and mRNA degradation.* Eur J Biochem, 1993. **213**(1): p. 285-93.
79. Coburn, G.A. and G.A. Mackie, *Differential sensitivities of portions of the mRNA for ribosomal protein S20 to 3′-exonucleases dependent on oligoadenylation and RNA secondary structure.* J Biol Chem, 1996. **271**(26): p. 15776-81.
80. McLaren, R.S., et al., *mRNA degradation by processive 3′-5′ exoribonucleases in vitro and the implications for prokaryotic mRNA decay in vivo.* J Mol Biol, 1991. **221**(1): p. 81-95.
81. Cheng, Z.F. and M.P. Deutscher, *An important role for RNase R in mRNA decay.* Mol Cell, 2005. **17**(2): p. 313-8.
82. Chen, C. and M.P. Deutscher, *Elevation of RNase R in response to multiple stress conditions.* J Biol Chem, 2005. **280**(41): p. 34393-6.
83. Purusharth, R.I., et al., *Exoribonuclease R interacts with endoribonuclease E and an RNA helicase in the psychrotrophic bacterium Pseudomonas syringae Lz4W.* J Biol Chem, 2005. **280**(15): p. 14572-8.
84. Yamamoto, Y., et al., *SsrA-mediated trans-translation plays a role in mRNA quality control by facilitating degradation of truncated mRNAs.* RNA, 2003. **9**(4): p. 408-18.
85. Mehta, P., J. Richards, and A.W. Karzai, *tmRNA determinants required for facilitating nonstop mRNA decay.* RNA, 2006. **12**(12): p. 2187-98.
86. Richards, J., P. Mehta, and A.W. Karzai, *RNase R degrades non-stop mRNAs selectively in an SmpB-tmRNA-dependent manner.* Mol Microbiol, 2006. **62**(6): p. 1700-12.
87. Moore, S.D. and R.T. Sauer, *The tmRNA system for translational surveillance and ribosome rescue.* Annu Rev Biochem, 2007. **76**: p. 101-24.
88. Choy, J.S., L.L. Aung, and A.W. Karzai, *Lon protease degrades transfer-messenger RNA-tagged proteins.* J Bacteriol, 2007. **189**(18): p. 6564-71.
89. Hanson, P.I. and S.W. Whiteheart, *AAA+ proteins: have engine, will work.* Nat Rev Mol Cell Biol, 2005. **6**(7): p. 519-29.
90. Flynn, J.M., et al., *Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals.* Mol Cell, 2003. **11**(3): p. 671-83.

91. Flynn, J.M., et al., *Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10584-9.
92. Kim, Y.I., et al., *Dynamics of substrate denaturation and translocation by the ClpXP degradation machine.* Mol Cell, 2000. **5**(4): p. 639-48.
93. Levchenko, I., et al., *Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag.* Mol Cell, 2003. **12**(2): p. 365-72.
94. Wah, D.A., et al., *Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the AAA+ ClpXP protease.* Mol Cell, 2003. **12**(2): p. 355-63.
95. Mettert, E.L. and P.J. Kiley, *ClpXP-dependent proteolysis of FNR upon loss of its O2-sensing [4Fe-4S] cluster.* J Mol Biol, 2005. **354**(2): p. 220-32.
96. Kuroda, A., et al., *Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in E. coli.* Science, 2001. **293**(5530): p. 705-8.
97. Kowit, J.D. and A.L. Goldberg, *Intermediate steps in the degradation of a specific abnormal protein in Escherichia coli.* J Biol Chem, 1977. **252**(23): p. 8350-7.
98. Gonzalez, M., et al., *Lon-mediated proteolysis of the Escherichia coli UmuD mutagenesis protein: in vitro degradation and identification of residues required for proteolysis.* Genes Dev, 1998. **12**(24): p. 3889-99.
99. Gur, E. and R.T. Sauer, *Recognition of misfolded proteins by Lon, a AAA(+) protease.* Genes Dev, 2008. **22**(16): p. 2267-77.
100. Flynn, J.M., et al., *Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation.* Genes Dev, 2004. **18**(18): p. 2292-301.
101. Neher, S.B., et al., *Latent ClpX-recognition signals ensure LexA destruction after DNA damage.* Genes Dev, 2003. **17**(9): p. 1084-9.
102. Erbse, A., et al., *ClpS is an essential component of the N-end rule pathway in Escherichia coli.* Nature, 2006. **439**(7077): p. 753-6.
103. Wang, K.H., et al., *Tuning the strength of a bacterial N-end rule degradation signal.* J Biol Chem, 2008. **283**(36): p. 24600-7.
104. Wang, K.H., et al., *The molecular basis of N-end rule recognition.* Mol Cell, 2008. **32**(3): p. 406-14.
105. Sauer, R.T., et al., *Sculpting the proteome with AAA(+) proteases and disassembly machines.* Cell, 2004. **119**(1): p. 9-18.
106. Baker, T.A. and R.T. Sauer, *ATP-dependent proteases of bacteria: recognition logic and operating principles.* Trends Biochem Sci, 2006. **31**(12): p. 647-53.
107. Ito, K. and Y. Akiyama, *Cellular functions, mechanism of action, and regulation of FtsH protease.* Annu Rev Microbiol, 2005. **59**: p. 211-31.
108. Lee, C., et al., *ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal.* Mol Cell, 2001. **7**(3): p. 627-37.
109. Kenniston, J.A., et al., *Effects of local protein stability and the geometric position of the substrate degradation tag on the efficiency of ClpXP denaturation and degradation.* J Struct Biol, 2004. **146**(1-2): p. 130-40.

110. Koodathingal, P., et al., *ATP-dependent proteases differ substantially in their ability to unfold globular proteins.* J Biol Chem, 2009. **284**(28): p. 18674-84.
111. Kenniston, J.A., et al., *Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine.* Cell, 2003. **114**(4): p. 511-20.
112. Gur, E. and R.T. Sauer, *Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine.* Proc Natl Acad Sci U S A, 2009. **106**(44): p. 18503-8.
113. Song, H.K., et al., *Mutational studies on HslU and its docking mode with HslV.* Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14103-8.
114. Yamada-Inagawa, T., et al., *Conserved pore residues in the AAA protease FtsH are important for proteolysis and its coupling to ATP hydrolysis.* J Biol Chem, 2003. **278**(50): p. 50182-7.
115. Siddiqui, S.M., R.T. Sauer, and T.A. Baker, *Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates.* Genes Dev, 2004. **18**(4): p. 369-74.
116. Hinnerwisch, J., et al., *Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation.* Cell, 2005. **121**(7): p. 1029-41.
117. Park, E., et al., *Role of the GYVG pore motif of HslU ATPase in protein unfolding and translocation for degradation by HslV peptidase.* J Biol Chem, 2005. **280**(24): p. 22892-8.
118. Graef, M. and T. Langer, *Substrate specific consequences of central pore mutations in the i-AAA protease Yme1 on substrate engagement.* J Struct Biol, 2006. **156**(1): p. 101-8.
119. Zhang, F., et al., *Mechanism of substrate unfolding and translocation by the regulatory particle of the proteasome from Methanocaldococcus jannaschii.* Mol Cell, 2009. **34**(4): p. 485-96.
120. Martin, A., T.A. Baker, and R.T. Sauer, *Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding.* Nat Struct Mol Biol, 2008. **15**(11): p. 1147-51.
121. Botos, I., et al., *The catalytic domain of Escherichia coli Lon protease has a unique fold and a Ser-Lys dyad in the active site.* J Biol Chem, 2004. **279**(9): p. 8140-8.
122. Swamy, K.H. and A.L. Goldberg, *E. coli contains eight soluble proteolytic activities, one being ATP dependent.* Nature, 1981. **292**(5824): p. 652-4.
123. Melnikov, E.E., et al., *Limited proteolysis of E. coli ATP-dependent protease Lon - a unified view of the subunit architecture and characterization of isolated enzyme fragments.* Acta Biochim Pol, 2008. **55**(2): p. 281-96.
124. Park, S.C., et al., *Oligomeric structure of the ATP-dependent protease La (Lon) of Escherichia coli.* Mol Cells, 2006. **21**(1): p. 129-34.
125. Botos, I., et al., *Crystal structure of the AAA+ alpha domain of E. coli Lon protease at 1.9A resolution.* J Struct Biol, 2004. **146**(1-2): p. 113-22.
126. Charette, M.F., et al., *DNA-stimulated ATPase activity on the lon (CapR) protein.* J Bacteriol, 1984. **158**(1): p. 195-201.

127. Charette, M.F., G.W. Henderson, and A. Markovitz, *ATP hydrolysis-dependent protease activity of the lon (capR) protein of Escherichia coli K-12.* Proc Natl Acad Sci U S A, 1981. **78**(8): p. 4728-32.
128. Gottesman, S., *Proteolysis in bacterial regulatory circuits.* Annu Rev Cell Dev Biol, 2003. **19**: p. 565-87.
129. Majdalani, N. and S. Gottesman, *Genetic dissection of signaling through the Rcs phosphorelay.* Methods Enzymol, 2007. **423**: p. 349-62.
130. Fernandez De Henestrosa, A.R., et al., *Identification of additional genes belonging to the LexA regulon in Escherichia coli.* Mol Microbiol, 2000. **35**(6): p. 1560-72.
131. Gonzalez, M. and R. Woodgate, *The "tale" of UmuD and its role in SOS mutagenesis.* Bioessays, 2002. **24**(2): p. 141-8.
132. Goodman, M.F., *Coping with replication 'train wrecks' in Escherichia coli using Pol V, Pol II and RecA proteins.* Trends Biochem Sci, 2000. **25**(4): p. 189-95.
133. Bertani, I., et al., *The Pseudomonas putida Lon protease is involved in N-acyl homoserine lactone quorum sensing regulation.* BMC Microbiol, 2007. **7**: p. 71.
134. Takaya, A., et al., *Negative regulation of quorum-sensing systems in Pseudomonas aeruginosa by ATP-dependent Lon protease.* J Bacteriol, 2008. **190**(12): p. 4181-8.
135. Cohn, M.T., et al., *Contribution of conserved ATP-dependent proteases of Campylobacter jejuni to stress tolerance and virulence.* Appl Environ Microbiol, 2007. **73**(24): p. 7803-13.
136. Gottesman, S., E. Halpern, and P. Trisler, *Role of sulA and sulB in filamentation by lon mutants of Escherichia coli K-12.* J Bacteriol, 1981. **148**(1): p. 265-73.
137. Stewart, B.J., J.L. Enos-Berlage, and L.L. McCarter, *The lonS gene regulates swarmer cell differentiation of Vibrio parahaemolyticus.* J Bacteriol, 1997. **179**(1): p. 107-14.
138. Su, S., et al., *Lon protease of the alpha-proteobacterium Agrobacterium tumefaciens is required for normal growth, cellular morphology and full virulence.* Microbiology, 2006. **152**(Pt 4): p. 1197-207.
139. Claret, L. and C. Hughes, *Rapid turnover of FlhD and FlhC, the flagellar regulon transcriptional activator proteins, during Proteus swarming.* J Bacteriol, 2000. **182**(3): p. 833-6.
140. Tomoyasu, T., et al., *Turnover of FlhD and FlhC, master regulator proteins for Salmonella flagellum biogenesis, by the ATP-dependent ClpXP protease.* Mol Microbiol, 2003. **48**(2): p. 443-52.
141. Jackson, M.W., E. Silva-Herzog, and G.V. Plano, *The ATP-dependent ClpXP and Lon proteases regulate expression of the Yersinia pestis type III secretion system via regulated proteolysis of YmoA, a small histone-like protein.* Mol Microbiol, 2004. **54**(5): p. 1364-78.
142. Takaya, A., et al., *Lon, a stress-induced ATP-dependent protease, is critically important for systemic Salmonella enterica serovar typhimurium infection of mice.* Infect Immun, 2003. **71**(2): p. 690-6.
143. Robertson, G.T., et al., *The Brucella abortus Lon functions as a generalized stress response protease and is required for wild-type virulence in BALB/c mice.* Mol Microbiol, 2000. **35**(3): p. 577-88.

144. Lan, L., et al., *Mutation of Lon protease differentially affects the expression of Pseudomonas syringae type III secretion system genes in rich and minimal media and reduces pathogenicity.* Mol Plant Microbe Interact, 2007. **20**(6): p. 682-96.
145. Herbst, K., et al., *Intrinsic thermal sensing controls proteolysis of Yersinia virulence regulator RovA.* PLoS Pathog, 2009. **5**(5): p. e1000435.
146. Losada, L.C. and S.W. Hutcheson, *Type III secretion chaperones of Pseudomonas syringae protect effectors from Lon-associated degradation.* Mol Microbiol, 2005. **55**(3): p. 941-53.
147. Garcia-Angulo, V.A., et al., *Regulation of expression and secretion of NleH, a new non-locus of enterocyte effacement-encoded effector in Citrobacter rodentium.* J Bacteriol, 2008. **190**(7): p. 2388-99.
148. Richards, J., P. Mehta, and A.W. Karzai, *RNase R degrades non-stop mRNAs selectively in an SmpB-tmRNA-dependent manner.* Mol Microbiol, 2006. **62**(62): p. 1700-12.
149. Keiler, K.C., P.R. Waller, and R.T. Sauer, *Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA.* Science, 1996. **271**(5251): p. 990-3.
150. Gillet, R. and B. Felden, *Emerging views on tmRNA-mediated protein tagging and ribosome rescue.* Mol Microbiol, 2001. **42**(Part 4): p. 879-886.
151. Atkins, J.F. and R.F. Gesteland, *A case for trans translation.* Nature, 1996. **379**(6568): p. 769-71.
152. Grzymski, E.C., *tmRNA to the rescue.* Nat Struct Biol, 2003. **10**(5): p. 321.
153. Haebel, P.W., S. Gutmann, and N. Ban, *Dial tm for rescue: tmRNA engages ribosomes stalled on defective mRNAs.* Current opinion in structural biology, 2004. **14**(1): p. 58-65.
154. Muto, A., C. Ushida, and H. Himeno, *A bacterial RNA that functions as both a tRNA and an mRNA.* Trends Biochem Sci, 1998. **23**(1): p. 25-9.
155. Withey, J.H. and D.I. Friedman, *A salvage pathway for protein structures: tmRNA and trans-translation.* Annu Rev Microbiol, 2003. **57**: p. 101-23.
156. Dulebohn, D., et al., *Trans-translation: the tmRNA-mediated surveillance mechanism for ribosome rescue, directed protein degradation, and nonstop mRNA decay.* Biochemistry, 2007. **46**(16): p. 4681-93.
157. Frazao, C., et al., *Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex.* Nature, 2006. **443**(7107): p. 110-4.
158. Zuo, Y., et al., *Structural basis for processivity and single-strand specificity of RNase II.* Mol Cell, 2006. **24**(1): p. 149-56.
159. Sundermeier, T., et al., *Studying tmRNA-mediated surveillance and nonstop mRNA decay.* Methods Enzymol, 2008. **447**: p. 329-58.
160. Matos, R.G., A. Barbas, and C.M. Arraiano, *RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation.* Biochem J, 2009. **423**(2): p. 291-301.
161. Vincent, H.A. and M.P. Deutscher, *Insights into how RNase R degrades structured RNA: analysis of the nuclease domain.* J Mol Biol, 2009. **387**(3): p. 570-83.

162. Vincent, H.A. and M.P. Deutscher, *The roles of individual domains of RNase R in substrate binding and exoribonuclease activity. The nuclease domain is sufficient for digestion of structured RNA.* J Biol Chem, 2009. **284**(1): p. 486-94.
163. Kramer, G., et al., *The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins.* Nat Struct Mol Biol, 2009. **16**(6): p. 589-97.
164. Gottesman, S., *Proteases and their targets in Escherichia coli.* Annu Rev Genet, 1996. **30**: p. 465-506.
165. Karzai, A.W. and R.T. Sauer, *Protein factors associated with the SsrA.SmpB tagging and ribosome rescue complex.* Proc Natl Acad Sci U S A, 2001. **98**(6): p. 3040-4.
166. Keiler, K.C., *Biology of trans-Translation.* Annu Rev Microbiol, 2008. **62**: p. 133-51.
167. Keiler, K.C., et al., *C-terminal specific protein degradation: activity and substrate specificity of the Tsp protease.* Protein Sci, 1995. **4**(8): p. 1507-15.
168. Parsell, D.A., K.R. Silber, and R.T. Sauer, *Carboxy-terminal determinants of intracellular protein degradation.* Genes Dev, 1990. **4**(2): p. 277-86.
169. Silber, K.R., K.C. Keiler, and R.T. Sauer, *Tsp: a tail-specific protease that selectively degrades proteins with nonpolar C termini.* Proc Natl Acad Sci U S A, 1992. **89**(1): p. 295-9.
170. Silber, K.R. and R.T. Sauer, *Deletion of the prc (tsp) gene provides evidence for additional tail-specific proteolytic activity in Escherichia coli K-12.* Mol Gen Genet, 1994. **242**(2): p. 237-40.
171. Gottesman, S., et al., *The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system.* Genes Dev, 1998. **12**(9): p. 1338-47.
172. Herman, C., et al., *Degradation of carboxy-terminal-tagged cytoplasmic proteins by the Escherichia coli protease HflB (FtsH).* Genes Dev, 1998. **12**(9): p. 1348-55.
173. Herman, C., et al., *Lack of a robust unfoldase activity confers a unique level of substrate specificity to the universal AAA protease FtsH.* Mol Cell, 2003. **11**(3): p. 659-69.
174. Kihara, A., Y. Akiyama, and K. Ito, *FtsH is required for proteolytic elimination of uncomplexed forms of SecY, an essential protein translocase subunit.* Proc Natl Acad Sci U S A, 1995. **92**(10): p. 4532-6.
175. Kihara, A., Y. Akiyama, and K. Ito, *Dislocation of membrane proteins in FtsH-mediated proteolysis.* Embo J, 1999. **18**(11): p. 2970-81.
176. Farrell, C.M., T.A. Baker, and R.T. Sauer, *Altered specificity of a AAA+ protease.* Mol Cell, 2007. **25**(1): p. 161-166.
177. Levchenko, I., et al., *A specificity-enhancing factor for the ClpXP degradation machine.* Science, 2000. **289**(5488): p. 2354-6.
178. Martin, A., T.A. Baker, and R.T. Sauer, *Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates.* Mol Cell, 2008. **29**(4): p. 441-50.

179. McGinness, K.E., et al., *Altered tethering of the SspB adaptor to the ClpXP protease causes changes in substrate delivery.* J Biol Chem, 2007. **282**(15): p. 11465-73.
180. Roche, E.D. and R.T. Sauer, *SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity.* EMBO J, 1999. **18**(16): p. 4579-89.
181. Howard-Flanders, P., E. Simson, and L. Theriot, *A Locus That Controls Filament Formation and Sensitivity to Radiation in Escherichia Coli K-12.* Genetics, 1964. **49**: p. 237-46.
182. Markovitz, A., *Regulatory Mechanisms for Synthesis of Capsular Polysaccharide in Mucoid Mutants of Escherichia Coli K12.* Proc Natl Acad Sci U S A, 1964. **51**: p. 239-46.
183. Bernstein, H.D. and J.B. Hyndman, *Physiological basis for conservation of the signal recognition particle targeting pathway in Escherichia coli.* J Bacteriol, 2001. **183**(7): p. 2187-97.
184. Hutchison, C.A., et al., *Global transposon mutagenesis and a minimal Mycoplasma genome.* Science, 1999. **286**(5447): p. 2165-9.
185. Glass, J.I., et al., *Essential genes of a minimal bacterium.* Proc Natl Acad Sci U S A, 2006. **103**(2): p. 425-30.
186. Gur, E. and R.T. Sauer, *Evolution of the ssrA degradation tag in Mycoplasma: specificity switch to a different protease.* Proc Natl Acad Sci U S A, 2008. **105**(42): p. 16113-8.
187. Bolon, D.N., et al., *Nucleotide-dependent substrate handoff from the SspB adaptor to the AAA+ ClpXP protease.* Mol Cell, 2004. **16**(3): p. 343-50.
188. Chien, P., et al., *Structure and substrate specificity of an SspB ortholog: design implications for AAA+ adaptors.* Structure, 2007. **15**(10): p. 1296-305.
189. Chien, P., et al., *Direct and adaptor-mediated substrate recognition by an essential AAA+ protease.* Proc Natl Acad Sci U S A, 2007. **104**(16): p. 6590-5.
190. Farrell, C.M., A.D. Grossman, and R.T. Sauer, *Cytoplasmic degradation of ssrA-tagged proteins.* Mol Microbiol, 2005. **57**(6): p. 1750-61.
191. Levchenko, I., et al., *Versatile modes of peptide recognition by the AAA+ adaptor protein SspB.* Nat Struct Mol Biol, 2005. **12**(6): p. 520-5.
192. Lies, M. and M.R. Maurizi, *Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in Escherichia coli.* J Biol Chem, 2008. **283**(34): p. 22918-29.
193. Park, E.Y., et al., *Structural basis of SspB-tail recognition by the zinc binding domain of ClpX.* J Mol Biol, 2007. **367**(2): p. 514-26.
194. Thibault, G., et al., *Specificity in substrate and cofactor recognition by the N-terminal domain of the chaperone ClpX.* Proc Natl Acad Sci U S A, 2006. **103**(47): p. 17724-9.
195. Hou, J.Y., R.T. Sauer, and T.A. Baker, *Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP.* Nat Struct Mol Biol, 2008. **15**(3): p. 288-94.
196. Mogk, A., R. Schmidt, and B. Bukau, *The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies.* Trends Cell Biol, 2007. **17**(4): p. 165-72.

197. Griffith, K.L. and R.E. Wolf, Jr., *Measuring beta-galactosidase activity in bacteria: cell growth, permeabilization, and enzyme assays in 96-well arrays.* Biochem Biophys Res Commun, 2002. **290**(1): p. 397-402.
198. Bretz, J., et al., *Lon protease functions as a negative regulator of type III protein secretion in Pseudomonas syringae.* Mol Microbiol, 2002. **45**(2): p. 397-409.
199. Schechter, L.M. and C.A. Lee, *AraC/XylS family members, HilC and HilD, directly bind and derepress the Salmonella typhimurium hilA promoter.* Mol Microbiol, 2001. **40**(6): p. 1289-99.
200. Takaya, A., et al., *Degradation of the HilC and HilD regulator proteins by ATP-dependent Lon protease leads to downregulation of Salmonella pathogenicity island 1 gene expression.* Mol Microbiol, 2005. **55**(3): p. 839-52.
201. Takaya, A., et al., *The ATP-dependent lon protease of Salmonella enterica serovar Typhimurium regulates invasion and expression of genes carried on Salmonella pathogenicity island 1.* J Bacteriol, 2002. **184**(1): p. 224-32.
202. Boddicker, J.D. and B.D. Jones, *Lon protease activity causes down-regulation of Salmonella pathogenicity island 1 invasion gene expression after infection of epithelial cells.* Infect Immun, 2004. **72**(4): p. 2002-13.
203. Lucas, R.L., et al., *Multiple factors independently regulate hilA and invasion gene expression in Salmonella enterica serovar typhimurium.* J Bacteriol, 2000. **182**(7): p. 1872-82.
204. Iyoda, S. and H. Watanabe, *ClpXP protease controls expression of the type III protein secretion system through regulation of RpoS and GrlR levels in enterohemorrhagic Escherichia coli.* J Bacteriol, 2005. **187**(12): p. 4086-94.
205. Cornelis, G.R., *Molecular and cell biology aspects of plague.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 8778-83.
206. Perry, R.D., et al., *DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of Yersinia pestis KIM5.* Infect Immun, 1998. **66**(10): p. 4611-23.
207. Ben-Gurion, R. and A. Shafferman, *Essential virulence determinants of different Yersinia species are carried on a common plasmid.* Plasmid, 1981. **5**(2): p. 183-7.
208. Hueck, C.J., *Type III protein secretion systems in bacterial pathogens of animals and plants.* Microbiol Mol Biol Rev, 1998. **62**(2): p. 379-433.
209. Rosqvist, R., K.E. Magnusson, and H. Wolf-Watz, *Target cell contact triggers expression and polarized transfer of Yersinia YopE cytotoxin into mammalian cells.* EMBO J, 1994. **13**(4): p. 964-72.
210. Michiels, T., et al., *Secretion of Yop proteins by Yersiniae.* Infect Immun, 1990. **58**(9): p. 2840-9.
211. Straley, S.C., et al., *Regulation by Ca2+ in the Yersinia low-Ca2+ response.* Mol Microbiol, 1993. **8**(6): p. 1005-10.
212. Cornelis, G.R., et al., *ymoA, a Yersinia enterocolitica chromosomal gene modulating the expression of virulence functions.* Mol Microbiol, 1991. **5**(5): p. 1023-34.
213. Cornelis, G.R., *Role of the transcription activator virF and the histone-like protein YmoA in the thermoregulation of virulence functions in yersiniae.* Zentralbl Bakteriol, 1993. **278**(2-3): p. 149-64.

214. Ge, Z. and A.W. Karzai, *Co-evolution of multipartite interactions between an extended tmRNA tag and a robust Lon protease in Mycoplasma.* Mol Microbiol, 2009. **74**(5): p. 1083-99.
215. Tsilibaris, V., G. Maenhaut-Michel, and L. Van Melderen, *Biological roles of the Lon ATP-dependent protease.* Res Microbiol, 2006. **157**(8): p. 701-13.
216. Kitagawa, M., et al., *Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research.* DNA Res, 2005. **12**(5): p. 291-9.