The Pennsylvania State University
The Graduate School
Intercollege Graduate Program in Genetics

ROLE OF THE TMRNA PATHWAY IN BACTERIAL PHYSIOLOGY
AND ITS USE IN ANTIBIOTIC DEVELOPMENT

A Thesis in
Genetics
by
Lin Cheng

© 2008 Lin Cheng

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2008
The dissertation of Lin Cheng was reviewed and approved* by the following:

Kenneth C. Keiler  
Associate Professor of Biochemistry and Molecular Biology  
Dissertation Advisor

Paul Babitzke  
Professor of Biochemistry and Molecular Biology  
Chair of Committee

Stephen J. Benkovic  
Evan Pugh Professor and Eberly Chair in Chemistry

Douglas R. Cavener  
Professor and Head of Biology

Eric T. Harvill  
Associate Professor of Microbiology and Infectious Disease

Richard W. Ordway  
Associate Professor of Biology  
Chair, Intercollege Graduate Degree Program in Genetics

*Signatures are on file in the Graduate School
ABSTRACT

tmRNA is known to play a role in translational quality control in all bacteria, but the phenotypes of cells lacking tmRNA suggest that the tmRNA pathway is integrated into the global regulatory network. The tmRNA pathway is important for both bacterial physiology and bacterial pathogenesis. To understand the physiological role of tmRNA in bacteria, and to exploit tmRNA as a target for antibiotics, we investigated the activity of tmRNA in the model species Caulobacter crescentus. In C. crescentus lacking tmRNA activity, initiation of DNA replication is delayed. To understand the molecular basis of the tmRNA phenotype, the impacts of tmRNA activity on the expression of replication initiator protein DnaA, and the transcription from a strong promoter (P_s) within the origin were investigated. We demonstrated that transcription of dnaA and P_s are both delayed in cells lacking tmRNA activity. We identified a TCAA motif from the dnaA promoter that is important for the timing of dnaA transcription. A possible model is that tmRNA can regulate an unknown transcriptional regulator, and the unknown transcriptional regulator could bind to the TCAA motif in both the P_{dnaA} and P_s, thereby regulating the timing of transcription from both and ensuring the correct timing of DNA replication initiation. We also showed that the 5’-UTR of dnaA strongly represses dnaA expression. Although this repression is not tmRNA-dependent, it could be a conserved mechanism on dnaA regulation in other bacteria.

Because tmRNA is essential for viability or virulence in many species, inhibitors of the tmRNA pathway could have bactericidal activity. We developed genetic methods to screen libraries of cyclic peptides for inhibitors of individual steps in the tmRNA pathway, and had isolated cyclic peptides that inhibit the proteolysis of tmRNA-tagged proteins in E. coli. Inhibitors shared little sequence similarity and interfered with various steps in the ClpXP mechanism in vitro. Several of these inhibitors were bactericidal when added exogenously to
cultures of *C. crescentus*, which requires ClpXP activity for viability. This result suggests that the tmRNA pathway may be a useful target for antibiotic development.
TABLE OF CONTENTS

LIST OF FIGURES ................................................................................................................. vii

LIST OF TABLES ................................................................................................................... viii

ACKNOWLEDGEMENTS ..................................................................................................... ix

Chapter 1 General introduction: tmRNA and trans-translation in bacteria ................. 1

Abstract ............................................................................................................................ 2
I. Introduction................................................................................................................... 3
II. History and phylogeny of tmRNA .............................................................................. 4
III. The structure and processing of tmRNA ............................................................... 6
IV. The function of tmRNA ............................................................................................. 9
   IV a The trans-translation model ............................................................................... 9
   IV b The tagging signals ......................................................................................... 12
   IV c The fate of tagged proteins ........................................................................... 16
   IV d The fate of nonstop mRNAs ....................................................................... 17
   IV e Ribosome rescue ........................................................................................... 18
V. tmRNA associated protein factors ........................................................................... 19
   V a SmpB ................................................................................................................ 19
   V b Ala-RS ............................................................................................................. 20
   V c EF-Tu ................................................................................................................ 20
   V d Other protein factors ....................................................................................... 20
VI. Physiology of the tmRNA Pathway ........................................................................... 21
   VI a tmRNA pathway and bacteria pathogenesis .............................................. 21
   VI b tmRNA pathway and bacteria stress response ......................................... 22
   VI c tmRNA pathway and bacterial development ........................................... 22
VII. tmRNA pathway is a good target for novel antibiontics development ............... 23
VIII. Future directions ....................................................................................................... 24
References ........................................................................................................................ 25

Chapter 2 tmRNA in the genetic network of cell cycle regulation ................................. 37

Abstract ............................................................................................................................ 38
Introduction...................................................................................................................... 39
Results .............................................................................................................................. 43
   The cell cycle regulation of dnaA transcription is disrupted in the ΔssrA strain ...... 43
   Both the synthesis and the accumulation of DnaA are disrupted by the ssrA deletion ............................................................................................................. 47
   tmRNA tags DnaA, but this tagging is not responsible for temporal transcription of dnaA ................................................................. 49
   tmRNA tags DnaA ........................................................................................... 49
   The degradation of DnaA is not affected by tmRNA ........................................ 51
   Mutations of the DnaA box in the 5'-UTR do not change the temporal transcription of dnaA in the ΔssrA strain ......................... 53
A motif in the dnaA promoter is responsible for the ssrA deletion phenotype ........55
The 5'-UTR of the dnaA gene represses dnaA expression..................................58
P, transcription is delayed by ssrA deletion .........................................................61
Discussion ..............................................................................................................63
Materials and methods .........................................................................................67
References ..............................................................................................................74

Chapter 3  Discovery of antibacterial cyclic peptides that inhibit the ClpXP protease ....82

Abstract ...................................................................................................................83
Introduction ..............................................................................................................84
Results .......................................................................................................................86
Screen for cyclic peptide inhibitors of ClpXP .........................................................86
Inhibition of ClpXP in vitro ......................................................................................89
Bactericidal activity of ClpXP inhibitors .................................................................94
Discussion ..............................................................................................................96
Materials and methods ..........................................................................................98
References ..............................................................................................................104
**LIST OF FIGURES**

Figure 1-1: Schematic structure of tmRNA in *E. coli* ................................................................. 7

Figure 1-2: Schematic representation of tmRNA processing in *C. crescentus* ......................... 8

Figure 1-3: The model of tmRNA function: *trans*-translation ....................................................... 11

Figure 2-1: The cell cycle regulated *dnaA* mRNA level is disrupted in ΔssrA cells ...................... 44

Figure 2-2: The transcription of *dnaA* gene is disrupted by the ssrA deletion ............................... 46

Figure 2-3: Both the synthesis and the accumulation of DnaA are disrupted by ΔssrA ............... 48

Figure 2-4: DnaA is tagged by tmRNA ............................................................................................ 50

Figure 2-5: The degradation of DnaA is not affected by the ssrA deletion .................................. 52

Figure 2-6: The DnaA box in the 5’-UTR of *dnaA* is not responsible for the ΔssrA phenotype ...................................................................................................................... 54

Figure 2-7: The TCAA motif is important for P*dnaA* activity and timing of *dnaA* transcription ........................................................................................................................ 57

Figure 2-8: C;71T mutation in P*dnaA* can partially rescue the ΔssrA phenotype ....................... 58

Figure 2-9: The leader sequence in *dnaA* can repress the expression of *dnaA* ......................... 60

Figure 2-10: The transcription from Ps was disrupted by the ssrA deletion ................................. 62

Figure 3-1: Selection for cyclic peptide inhibitors of ClpXP ......................................................... 88

Figure 3-2: Cyclic IXP1 inhibits ClpXP *in vitro* ........................................................................... 90

Figure 3-3: Interaction of IXP1 with ClpX and ClpP *in vitro* .................................................... 91

Figure 3-4: Cyclic IXP1 inhibits degradation of λ O by ClpXP ..................................................... 93
LIST OF TABLES

Table 2-1: Plasmids and strains used in this study................................................................. 68

Table 3-1: Cyclic peptides identified from in vivo screen. ..................................................... 89

Table 3-2: In vivo and in vitro properties of cyclic peptides................................................... 95
ACKNOWLEDGEMENTS

I would like to acknowledge my advisor Dr. Kenneth Keiler, for all his guidance and support. He gave me the best possible scientific training and transformed me from a lay person to a qualified Ph.D. His enthusiasm, motivation and persistence in research will always inspire and encourage me.

I am very thankful to my committee members, Dr. Paul Babitzke, Dr. Steve Benkovic, Dr. Doug Cavener, Dr. Eric Harvill, and Dr. Davis Ng for their helpful discussions and guidance. I am also very thankful to the chair of Genetics program Dr. Rick Ordway for his help and support.

The past and present Keiler lab members have given me tremendous help in both my research and in my daily life. Their names will be always remembered: Dr. Sue-Jean Hong, Dr. Faith Lessner, Ling Li, Elisabeth Mehan, Dennis Lee, Jay Russell, Xin Zhou, Anastasiya Yakhnin, Priscilla Tee, Tomoyo Takagi, and Jake Wesley. I am very lucky to have worked with these wonderful people.

I would also like to thank my collaborators in Benkovic lab, Dr. John Tomsho, Dr. Todd Naumann, and Dr. Alex Horswill for their help. It has been a lot of fun to work with them.

I am very grateful to all the friends I have known in State College. Their names have not been mentioned but will never be forgotten. They made my life in graduate school exciting.

Finally, my family has been a constant support to me both emotionally and financially. I am greatly indebted to my dad, my mom, my sister and my boyfriend for their tremendous love and encouragement. They are the source of my happiness and success.
Chapter 1

General introduction:

tmRNA and trans-translation in bacteria
Abstract

tmRNA is a unique RNA molecule found in all bacteria that has both a tRNA-like structure and an mRNA-like open reading frame (ORF). tmRNA binds to small protein B (SmpB), and is charged with alanine by alanyl-tRNA synthetase (AlaRS), then bound by EF-Tu-GTP. The current model of tmRNA function is *trans*-translation. During bacterial translation, ribosome stalling happens with such high frequency that there will be no free ribosomes left after even one cell generation unless there is a mechanism to release stalled ribosomes. The tmRNA complex can recognize and enter the A site of the stalled ribosome, at which point normal transpeptidation occurs and the nascent peptide chain is added to the alanine on tmRNA. Following transpeptidation, the open reading frame switches from the former mRNA template to the mRNA portion within tmRNA, and normal translation continues to the stop codon of the tmRNA ORF. The result of *trans*-translation is the generation of tagged proteins, the release of stalled ribosomes, and the degradation of nonstop mRNAs. Additionally, the tmRNA tagged proteins can be proteolysed by multiple intracellular proteases such as ClpXP. Therefore, the tmRNA pathway is important for bacterial protein quality control. Also, tmRNA is important for bacterial physiology, having roles in development, pathogenesis, and the stress response. Study of the tmRNA pathway will not only lead to a better understanding of bacterial physiology, but also provide new targets for novel antibiotic development.
I. Introduction

With the identification of small RNAs in both prokaryotic and eukaryotic organisms, biologists have started to realize the importance of small RNAs in gene regulation. Small RNAs function as modulators of gene expression in various genetic networks, to maintain the homeostasis in the cell, to respond to signals from the environment, or to make decisions on the entry into various developmental stages.

To date, three different functions have been identified for small RNAs. First, certain small RNAs, such as small interference RNAs (siRNAs) or microRNAs, can pair with their complementary sequences in mRNAs, which will block the mRNAs from translation, and/or lead to rapid degradation of the target mRNAs. This is the basis for RNA interference (RNAi) technology, and has been studied in several eukaryotic organisms including humans. One example is the lin4 RNA in Caenorhabditis elegans (the first identified temporal small RNA in a eukaryotic system [1].) The existence of lin4 RNA triggers the repression of other genes in the first larval stage. Second, some small RNAs can bind to protein factors and function together to increase or decrease the transcription or translation of their regulons. For instance, the 6S RNA in Escherichia coli binds to the σ^{70} holoenzyme and represses gene expression from σ^{70}-dependent promoters in stationary phase [2]. Small RNAs in CsrB family affect translation and/or mRNA stability by sequestering the RNA-binding protein CsrA [3]. Third, there is a special RNA molecule called tmRNA that functions by trans-translation. In bacteria, ribosomes stall on mRNAs without in frame stop codon (nonstop mRNAs) during translation. The tmRNA recognizes the stalled ribosome and adds a tag to the nascent peptide chain, which leads to rapid proteolysis of the tagged protein and the rescue of the stalled ribosome [4]. tmRNA has been found in all species of bacteria [5]. The tmRNA trans-translation function is important for protein quality control. Also, it is believed that tmRNA has an important regulatory role in bacterial
physiology, because mutation of tmRNA in different species of bacteria causes variant phenotypes. However, how tmRNA regulates gene expression is still unknown. Our study focuses on the role of tmRNA in the cell cycle regulation of *Caulobacter crescentus*, and the potential of the tmRNA pathway to serve as a target for the development of novel, broad-specificity antibiotics.

II. History and phylogeny of tmRNA

In 1978, Lee et al. discovered a novel band of small stable RNAs in total *E. coli* total RNA on a 12% polyacrylamide gel. This RNA was named 10S RNA according to its estimated density [6]. Later, the 10S RNA was identified as a mixture. It has both 10Sa RNA which is tmRNA, and 10Sb RNA which is the ribozyme RNase P [7, 8]. The tmRNA gene is called *ssrA* for small stable RNA A.

The precursor of tmRNA was identified in 1989 [9], and the open reading frame within tmRNA was first identified the same year [10]. Tyagi et al identified the tmRNA gene in *Mycobacterium tuberculosis* in 1992, and noted that the 3’ end of tmRNA has high homology with the pseudouridine arm of tRNA [11]. The tRNA-like structure of tmRNA was subsequently proposed and verified by two independent groups [12, 13], but its function remained a mystery.

In 1995, Tu et al attempted to over express murine interleukin 6 (IL-6) in *E. coli*, and found that all truncated variants of IL-6 recovered from inclusion bodies had an identical C-terminal tag which does not exist in the full length IL-6 [14]. However, this tag could not be detected when IL-6 was overexpressed in *E. coli* without the *ssrA* gene. Therefore, tmRNA is responsible for the addition of the C-terminal tag, but the mechanism was still unknown.

In 1996, Keiler et al noticed that tagged IL-6 proteins have a C-terminal sequence identical to that of the Tsp protease recognition site. This finding led to the assumption that
tmRNA can add a tag containing proteolysis determinants to the nascent proteins, resulting in the rapid degradation of tagged proteins. Subsequently, they designed reporter genes with a transcriptional terminator but without a stop codon to generate nonstop mRNAs, and found that proteins expressed from nonstop mRNAs are much more stable in strains lacking tmRNA activity. The tagging was also confirmed by mass spectrometry in the Tsp⁻ strain [4]. This was the first time that the \textit{trans}-translation model of tmRNA function had been proposed.

The \textit{trans}-translation model has now been verified and modified with further understanding of the tmRNA pathway. For example, we now know that tmRNA must also bind with protein factors SmpB (small protein B) and EF-Tu·GTP [15, 16]. The details of the tmRNA pathway will be discussed in later sections.

The tmRNA pathway is highly conserved. Both tmRNA and SmpB have been identified in more than 500 species of bacteria [17]. In most bacteria, tmRNA is a single RNA molecule. But in \(\alpha\)-proteobacteria [18], cyanobacteria [19] and one of the \(\beta\)-proteobacteria (\textit{D. aromatica} RCB) [20], there is a circular permutation in the \textit{ssrA} gene. The intervening sequence is processed from the tmRNA precursor, resulting in a mature tmRNA consisting of two RNA molecules. However, the two-piece tmRNAs have the same tRNA like structure and the same tagging activity as the one-piece tmRNA. The circular permutations in 3 different lineages of bacteria could have happened independently during evolution, which suggests that there may be a certain selective advantage. Or could have originated once, followed by recombination.

The tmRNA gene is also found in bacteriophages such as mycobacteriophage [21] and bacillushage G [17]. Some chloroplasts [22] have both tmRNA and SmpB. In the mitochondria of the jakobid \textit{Reclinomonas americana}, partial tmRNA genes lacking an mRNA-like region were found [23]. Nevertheless, no tmRNA gene has been identified in archaea or eukaryotic nuclear genomes, supporting that they have totally different systems to release stalled ribosomes containing nonstop mRNAs.
III. The structure and processing of tmRNA

III a. tmRNA structure

As figure 1-1 shows, the 5’ end and 3’ end of mature tmRNA fold into a tRNA like structure including the D-arm, T-arm and acceptor arm. This structure was first identified in *E. coli*, *Mycoplasma capricolum* and *Bacillus subtilis* [12, 13], and has now been confirmed in all species of bacteria investigated. The acceptor arm of tmRNA has conserved characteristics of tRNAAla. For example, the third base pair of the acceptor arm is the same G:U wobble base pair and the fourth base from the 3’ end is also adenine [13]. With its conserved acceptor arm and T-arm, tmRNA can be recognized and charged by alanyl-tRNA synthetase (AlaRS) both *in vitro* and *in vivo* [12]. Only the charged tmRNA has tagging activity [24]. The folding of tmRNA occurs during transcription. It has been reported that pausing of the RNA polymerase at certain locations of the ssrA gene can facilitate the folding of tmRNA [25].

From its 5’ to 3’ end, tmRNA forms three or four pseudoknots depending on the particular species [17, 26, 27] and the mRNA-like domain usually lies between pseudoknot 1 and 2. The structure of the pseudoknots is more important than the sequence [28]. Pseudoknots are believed to help tmRNA folding, which is important both for the stability of tmRNA and the binding of tmRNA-associated protein factors [29].

The mRNA-like domain contains a short open reading frame (ORF). The encoded peptide length ranges from 8 to 35 residues in different species of bacteria [28]. Comparison of 610 tmRNA sequences from different bacteria indicates the bias of amino acid composition in the ORFs: Ala and Asp are the most commonly used amino acids while Trp and Cys are very rare. This preference may ensure that tmRNA can function normally even under stringent conditions
such as oxidation, because Trp and Cys are sensitive to oxidation [28]. Recently, it also has been reported that the 3’ end of the ORF is required for the degradation of nonstop mRNAs [30].

![Schematic structure of tmRNA in E. coli](image)

**Figure 1-1**: Schematic structure of tmRNA in *E. coli*

The green line is the tRNA like domain, which includes the D-arm, T-arm and acceptor arm. The purple bar is the mRNA like domain, which encodes a tag sequence containing a determinant for proteolysis.

### III b. The processing of tmRNA

The tmRNA precursor is transcribed from the *ssrA* gene (small stable RNA A). It must be processed to mature tmRNA and become charged to be active.

In *E. coli*, the 457 nucleotide tmRNA precursor is transcribed from a single gene located at 56.5 min on the chromosome [31]. The ribozyme RNase P removes 7 bases from its 5’ end [12], just as it does to tRNA precursors. The extra nucleotides on the 3’ end are first cleaved by RNase III in the presence of Mn$^{2+}$ [32], and then by RNase T and RNase PH [33]. The mature *E. coli* tmRNA is a one-piece RNA containing 363 nucleotides, and its 3’ end -CCA tail comes from transcription, unlike *Bacillus subtilis* tmRNA, whose -CCA tail is added by alanyl-tRNA nucleotidyltransferase [13].
In α-proteobacteria, cyanobacteria and β-proteobacterium *D. aromatica* RCB, however, there is a circular permutation in the *ssrA* gene. Therefore, the transcript is a one-piece precursor and must be processed into the mature two-piece tmRNA. For example in *C. crescentus* the tmRNA precursor has 327 nucleotides. From the 5’ to 3’ end, nucleotides 1-83 of the precursor form the acceptor RNA, nucleotides 113-327 form the coding RNA, and the 30 nucleotides in between form a closed loop between the two pieces that is removed as Fig 1-2 shows [18]. RNase P or other endoribonucleases are thought to cut the intervening loop, allowing other enzymes to process the -CCA tail and charge it with alanine [18].

![Figure 1-2: Schematic representation of tmRNA processing in *C. crescentus*](image)

The tmRNA precursor in *C. crescentus* has an intervening loop (yellow) because of a circular permutation in the *ssrA* gene. Processing of this loop is required to generate the mature tmRNA. The mature tmRNA has a 214 nt coding RNA, and an 83 nt acceptor RNA. The 5’ end of the coding RNA and the 3’ end of the acceptor RNA form the tRNA like domain (green). The purple bar is the ORF encoding tag sequence. The two-piece tmRNA has similar structure as the one piece tmRNA.
IV. The function of tmRNA

IV a. The *trans*-translation model

In bacteria, normal translation will end when the ribosome reaches the stop codon on the mRNA template. Release factors (RF) enter the A site of the ribosome, and trigger the hydrolysis of the ester bond in peptidyl-tRNA. The nascent peptide chain is then released, and ribosome subunits are dissociated by ribosome recycling factor (RRF) [34]. However, when the mRNA template lacks a stop codon, the ribosome will stall at the 3’ end of the nonstop mRNA. This event occurs at a high frequency about 13,000 times per cell per generation [35]. The stalled ribosome is stable *in vitro* [36]. Therefore, a special mechanism is needed to release stalled ribosomes and maintain a normal ribosome pool *in vivo*. This mechanism is the *trans*-translation process, or tmRNA pathway.

The stalled ribosome at the 3’-end of a nonstop mRNA is the trigger for *trans*-translation. The tmRNA pathway does not interfere with ribosomes during the normal translation processes [35]. A possible explanation is that SmpB has an interaction with the ribosome that imitates the anticodon in tRNAs and facilitates the entrance of the tmRNA complex into the A site [37-41]. But this interaction may not be as strong as the codon-anticodon pairing. Also, the ribosome normally has its mRNA entrance tunnel closed during translation elongation [42], and the tmRNA forms a complex with SmpB and EF-Tu·GTP [43, 44] which is too large to access the A site of the ribosome. When the ribosome is stalled at or near the 3’ end of nonstop mRNAs, it may have the tunnel open [28]. This conformational change could provide an opportunity for the tmRNA complex to enter the stalled ribosome.

After the tmRNA complex enters the A site of the stalled ribosome, as figure 1-3 shows, the normal transpeptidation happens, and the nascent peptide chain is added to the alanine on
tmRNA [44]. Following this event, the tRNA like region of the tmRNA moves to the ribosome P site, and the mRNA portion within tmRNA occupies the decoding center of the ribosome. The open reading frame will then switch from the nonstop mRNA template to the tmRNA ORF. The translation of the tmRNA ORF could start with a frame shift when sequence upstream of the ORF is mutated [45-47]. Normal translation continues on the tmRNA ORF until it encounters the stop codon, then RF functions to terminate translation [4]. The ribosome will release even in the absence of RRF, which demonstrates the ribosome recycling activity of tmRNA [48].

The results of trans-translation are: 1) The tagged protein is recognized and degraded rapidly by multiple cellular proteases. 2) The stalled ribosome is rescued and recycled. 3) The nonstop mRNA is degraded by ribonucleases.

The trans-translation model has been supported by many experiments. For example, tmRNA is only associated with 70S ribosome particles but not 30S or 50S subunits [13, 49]. The cryo-EM structure of the tmRNA complex with stalled ribosomes has been established [37, 44, 50]. Also, the first amino acid in the tag sequence is neither from the nonstop mRNA nor the tmRNA ORF, it comes from the alanine charged to tmRNA [51]. When tmRNA is mutagenized and charged by histidine instead of alanine, the first amino acid in the tag sequence switches to histidine too [51]. Moreover, the nascent peptide translated from a nonstop mRNA can be tagged by tmRNA both in vitro and in vivo [4, 52, 53]. And the tagged proteins will be degraded rapidly by proteases such as ClpXP, ClpAP, FtsH and Tsp [4, 54-57] [58]. All the data show that the tmRNA pathway is very important for protein quality control in all bacteria. It ensures ribosome recycling and avoids the accumulation of abnormal protein products.
Figure 1-3: The model of tmRNA function: *trans*-translation

The ribosome (grey) stalls at the 3’-end of nonstop mRNA (green line). The charged tmRNA can recognize and enter the A site of the stalled ribosome. Then normal transpeptidation happens, and the nascent peptide chain (green boxes) will be added to the alanine (yellow box) charged to the tmRNA. After that, the open reading frame switches from the former mRNA template (green line) to the tmRNA ORF (purple), and the released nonstop mRNA will be degraded. Normal translation continues on the tmRNA ORF until the ribosome reaches the stop codon. Then the stalled ribosome and the tagged protein are released. The tag sequence (purple bar) has multiple determinants for proteolysis and will be degraded rapidly.
IV b. The tagging signals

Several different tagging signals have been identified. These signals can be at the DNA level, the RNA level, or the protein level and can result in nonstop mRNAs, which will then stall ribosomes and trigger the tmRNA pathway.

**Damaged mRNAs**

Damaged mRNAs can lose their in-frame stop codon by physical or chemical damage, or nucleolytic degradation [59]. In eukaryotic systems, transcription occurs in the nucleus, while translation takes place in the cytoplasm. The truncated mRNAs cannot be exported to the cytoplasm, and therefore will not occupy the translation apparatus. However, in bacteria, translation is coupled with transcription, and so the ribosome cannot distinguish damaged mRNAs from normal mRNAs until it reaches the 3'-end and becomes stalled. The tmRNA pathway not only rescues the ribosome stalled at the damaged mRNA, but also leads to the degradation of damaged mRNA templates [60].

**Cleavage of mRNA by toxins**

Toxin-antitoxin (TA) systems were first discovered in the study of plasmid maintenance. The toxins are stable proteins. The antitoxins are not unstable, they could be RNA (type I) or protein (type II), and they can neutralize the toxins by inhibiting translation of toxin (type I) or sequestering the binding of toxin (type II) [61, 62]. Therefore, the host cell must maintain the plasmids to avoid the accumulation of toxins. There are also chromosomally encoded TA systems, which are very important for bacterial stress response [63, 64]. The connection between TA systems and the tmRNA pathway is that some toxins, for example the ribonuclease RelE in E.
coli, can cleave mRNA at the A site in paused ribosomes [65], and activate the *trans*-translation process. Under stress such as nutrient shortage, the toxins are induced and stop nonessential transcriptions by cutting the mRNAs at the ribosomal A site, ensuring that all available resources will be saved only for essential protein synthesis [66, 67]. In such cases, the tmRNA pathway is coordinated to clean up the burst of nonstop mRNAs and release stalled ribosomes.

**Transcription roadblock**

The binding of a transcriptional regulator onto the DNA sequence, or the existence of a transcriptional terminator, will block transcription and consequently block the coupled translation process, resulting in a stalled ribosome. For example in *E. coli*, high concentration of LacI protein can bind to the lac operator at both the O1 and O3 sites [68]. The O3 site is within the *lacI* coding sequence [69]. Binding of LacI to O3 will block transcription and generate truncated *lacI* mRNA, which will then be translated to truncated LacI and be tagged by tmRNA for rapid degradation [69]. This negative autoregulation ensures the LacI concentration is autogenously controlled.

**Rare codons**

During translation, when there are consecutive rare codons and the cell is short of the cognate tRNA, the ribosome will stall at the rare codons, irrespective of whether they are near the end of the coding sequence [53, 70]. The mRNA will then be cleaved within or near the ribosome A site [71], and generate a truncated mRNA, which will then activate the tmRNA pathway. This reaction happens when bacteria are under stress response, such as during amino acid starvation [72]. Stalled ribosomes are also observed in strains lacking certain tRNA genes or release factors [53]. When the gene has a rare codon adjacent to an inefficient translational termination codons,
such as an opal (UGA) stop codon in the ribokinase gene \textit{rbsK} in \textit{E. coli}, tagging will be enhanced [70].

\textit{Protein C terminus sequence}

The C-terminal sequence of some proteins is important for tmRNA tagging. For example, the \textit{E. coli} protein YbeL can be tagged by tmRNA as a full length protein. A single proline as the ultimate amino acid is sufficient for tmRNA tagging, and the penultimate amino acid can increase the tagging efficiency if it is Asp, Glu, Ile, Pro or Val. Also, an inefficient stop codon leads to a higher tagging level, and the tagging is decreased by overexpression of release factors [73]. Another example is the \textit{E. coli} LacIΔC1 protein that lacking of the last C-terminal amino acid. Addition of LacIΔC1 C-terminal Leu-Glu-Ser-Gly sequence to the unrelated CRP protein is sufficient to trigger tmRNA tagging of the chimeric protein [74]. This tagging is not dependent on translational readthrough. The possible explanations are: the special C-terminal sequence may interact with ribosomes and cause pausing; it may decrease the affinity of release factors; or it may increase the affinity of tmRNA complex binding to ribosomes.

\textit{Translational readthrough}

The inefficient opal stop codon (UGA) can lead to translational readthrough, and enhance tmRNA tagging [74]. The presence of ochre or amber suppressor tRNA will also result in translational readthrough and increased tmRNA tagging. Strains with suppressor tRNAs but lacking tmRNA activity have a growth defect [75]. A possible explanation is that the translational readthrough may increase the longevity of ribosome pausing at the 3’ end of transcripts, instead
of releasing at the normal stop codon. If the tmRNA pathway cannot function to remove the abnormal transcription and translation products, the cell will be affected [75].

**DNA Motif**

In *C. crescentus*, there is a 16-mer conserved motif CGACAAGATCGTCGTG upstream of the tagging site in 66% of tmRNA substrates. Mutations within this motif decrease the tagging efficiency [76]. The mechanism for this tagging signal is still unknown. It is unlikely that the motif provides a ribonuclease cleavage site because it is upstream of the tagging site, but it is possible that the mRNA of this motif can stop the translation by interacting with ribosome after the ribosome goes through it.

All the above signals can generate either nonstop mRNAs, but the tmRNA tagging efficiency is related to the length of mRNA downstream of the ribosome P site. The efficiency is high when the length is \( \leq 6 \) nt, and there will be no tagging if the length is more than 15 nt [77]. Therefore, no matter how the signal was originally generated, the tmRNA pathway only functions when the ribosome is stalled near the end of the mRNA, or when RNA cleavage makes the length [71, 78, 79]. There are at least two different mechanisms for mRNA cleavage, ribosome A site cleavage and the edge cleavage. A site cleavage could be caused by the intrinsic cleavage activity of paused ribosomes [78, 80], as well as by TA systems. Edge cleavage happens when the mRNA downstream of the A site is cleaved at a position near the tunnel that mRNA emerges from the ribosome. It is not clear how the decision of which cleavage is made, but the cleavage is not tmRNA or SmpB dependent [78]. Also, the A-site cleavage is avoided or inhibited for regulated ribosome stalling, such as SecM mediated ribosome arrest [79]. In general, the translational pause on an mRNA shorter than 15 nt downstream of the A site is necessary for tagging.
IV c. The fate of tagged proteins

The tag sequence encoded by tmRNA contains determinants for multiple intracellular proteases and proteolytic adaptors, so that the tagged proteins will be recognized and degraded rapidly. This degradation is important because usually the tagged proteins are truncated products and accumulation of them may be harmful to the cell.

There are several proteases and adaptors responsible for the degradation of tmRNA tagged proteins. In *E. coli*, ClpXP, ClpAP, and Lon are cytoplasmic proteases and will degrade the tagged proteins in the cytoplasm [54-57]; FtsH is on the inner side of cytoplasmic membrane and will degrade tagged integral membrane proteins [58]; Tsp will degrade the tagged proteins in the periplasm [4]. ClpXP and ClpAP are both enzyme complexes. The hexamers of ClpX or ClpA have ATPase activity. They recognize the C terminus of the tmRNA tag, unfold the tagged proteins, and translocate the unfolded substrates to ClpP which has the peptidase activity. ClpXP is preferred and is responsible for the degradation of the majority of the tagged proteins [81], while ClpAP and Lon work mostly when ClpX is not available. ClpXP is also highly conserved in other bacterial species such as *C. crescentus* and *B. subtilis* [82, 83]. SspB (stringent starvation protein B) is the adaptor that can bind to both the N terminus of the tag sequence and ClpX. In *E. coli*, binding of the SspB dimer can tether the tagged proteins to the ClpXP complex, and enhance the degradation both *in vitro* and *in vivo* [84]. The tmRNA tagged proteins will still be degraded by ClpXP without SspB, but at a slower rate [84]. SspB in *C. crescentus* shares limited sequence similarity with *E. coli* SspB, but it has the same function as a proteolysis adaptor [85-87]. There is another adaptor named ClpS in *E. coli*, which can bind to the N terminus of ClpA and prevent ClpAP from degrading the tagged protein [88].

The tag sequence is highly conserved throughout the bacterial kingdom. It always starts with alanine, which is the amino acid charged to tmRNA. In most species, the tag ends with two
alanines, or at least nonpolar residues that are good targets for proteolysis. Addition of the tmRNA tag is sufficient for exogenous proteins, such as GFP, to be degraded both in vitro and in vivo [84, 87, 89]. When the last two codons in the tmRNA ORF were switched from alanine (tmRNA-AA) to aspartate codons (tmRNA-DD) [4], or the last 6 amino acids codons were replaced by 6 histidine codons (tmRNA-His6) [90], the variant tmRNAs can still tag their substrates, but most of the tagged proteins will be stable and cannot be degraded. There are two exceptions that have been noted so far. One is that Lon can degrade some tmRNA-His6 tagged proteins [55]. The other is that the tmRNA-DD tagged bacteriophage Mu repressor can still be degraded rapidly, because the tagging changes the protein conformation and exposes other ClpXP recognition sites [91].

IV d. The fate of nonstop mRNAs

Nonstop mRNAs occupy the translation apparatus. Although some of these nonstop mRNAs are important for gene regulation, such as the previously mentioned lacI regulation, most of them must be degraded to maintain homeostasis in the cell. tmRNA can specifically accelerate the degradation of nonstop mRNAs [92]. In the E. coli ssrA deletion strain, the full length mRNA and the nonstop mRNA transcribed from the same gene had the same half life, which was about 2-3 min. While in strains with ssrA, nonstop mRNA had a half life of about 0.6 min, and the half life of full length mRNA was not affected [60]. A possible reason is that during translation, binding of ribosomes will protect the mRNAs from degradation. Once the tmRNA pathway releases the stalled ribosomes and exposes the nonstop mRNAs to ribonucleases, the nonstop mRNAs will be degraded rapidly [60]. Also, it has been reported that tmRNA can recruit the ribonuclease RNase R to degrade the nonstop mRNAs [93, 94]. In addition, the last three codons at the 3’ end of the tmRNA ORF are important for the decay of nonstop mRNAs. Mutations in


this region only slow down the nonstop mRNA degradation but don’t change the tmRNA tagging activity [30].

**IV e. Ribosome rescue**

There are 650-700 tmRNA molecules per cell, which is about 5% of the number of ribosomes. It has been reported that at least 13,000 translation events were stopped in each cell generation [35]. Ribosome stalling happens at such a high frequency that stalled ribosomes must be released somehow, or there will be no free ribosomes left after one generation. The tmRNA complex can enter the stalled ribosome and switch the template from nonstop mRNAs to tmRNA ORF, and release the ribosome when translation of the tmRNA ORF is completed. The release does not require the ribosome recycling factor RRF, which indicates that tmRNA has intrinsic ribosome release activity [48].

Rescue of the ribosome is important for bacterial physiology. Therefore it is not surprising that there are alternative pathways besides *trans*-translation to release stalled ribosomes, so that the cell will have a better chance to survive. For example in *E. coli*, if the tmRNA pathway is the only mechanism for ribosome rescue, then the *ssrA* deletion should be lethal. But the fact is that the *E. coli ssrA* deletion strain has only mild phenotypes, such as slower recovery from stationary phase [95], reduced motility on semisolid agar plate [12], and inability to support the lytic growth of bacteriophage λimmP22 [96]. The alternative pathway is still under investigation and it may be related to the peptidyl-tRNA hydrolase (Pth). Pth hydrolyzes the ester bond in peptidal-tRNAs during normal translations, and increase the peptidal-tRNA drop-off from stalled ribosomes, thereby release the stalled ribosomes [28, 97].
V. tmRNA associated protein factors

The tmRNA must be charged by AlaRS, and bound by SmpB and EF-Tu·GTP. Without these three proteins, the tmRNA will have no activity. There are also other proteins associated with tmRNA, such as RNase R, which are important for the stability and activity of tmRNA.

Va. SmpB

The smpB gene has been identified in all sequenced bacterial genomes, and is highly conserved. Deletion of the smpB gene has the same phenotype as deletion of ssrA [15, 83, 98, 99]. SmpB was first characterized in Salmonella typhimurium, in which mutations in smpB caused the same bacteriophage P22 growth defect as seen in the E. coli ssrA deletion [15, 24, 96, 99]. Although the smpB gene is immediately upstream of the ssrA gene in both species, the tmRNA level was not affected by polar effect on transcription. Therefore, SmpB must be related to the activity of tmRNA. Now we know that SmpB binds to the tRNA-like domain of tmRNA in a 1:1 stoichiometric ratio with high affinity [16, 100]. The binding of SmpB to tmRNA is important for tmRNA stability. For example, in C. crescentus, without the protection of SmpB, tmRNA will be degraded rapidly by RNase R [101]. The binding of SmpB is also important for the association of tmRNA and 70S ribosomes [102]. If the smpB gene is deleted, tmRNA cannot associate with 70S ribosomes [15]. Moreover, the C terminus of SmpB can mimic the anticodon in tRNAs and interact with the stalled ribosome near the decoding center [15, 40]. Deletion of this region won’t affect the association of tmRNA and 70S ribosomes, but the tmRNA will have no tagging activity [41].
V b. Ala-RS

As previously mentioned, tmRNA has the conserved characteristics of tRNA^{Ala}, such as the G:U wobble base pair of the acceptor arm, and A as the fourth base from the 3’ end. AlaRS can recognize these characteristics and charge tmRNA both \textit{in vitro} and \textit{in vivo} [12, 24]. As uncharged tmRNA has no activity [24], AlaRS is a required protein factor for the tmRNA pathway.

V c. EF-Tu

EF-Tu is the elongation factor that ensures the fidelity of protein synthesis. During normal translation, EF-Tu·GTP binds to the aminoacylated tRNA to protect the ester bond, and the complex will enter the A site of the ribosome. If the correct tRNA is brought to the ribosome, fast GTP hydrolysis is triggered by the codon-anticodon pairing and EF-Tu·GDP will dissociate, and transpeptidation occurs. If the tRNA is not the cognate one, then GTP hydrolysis is slow and the whole complex will dissociate before transpeptidation occurs. In the tmRNA pathway, EF-Tu·GTP binds to the T arm and acceptor arm of tmRNA, and protects the ester bond of charged tmRNA [103, 104]. The GTP hydrolysis of the tmRNA complex is triggered by the interaction of SmpB with the ribosome, which mimics the codon-anticodon pairing [40].

V d. Other protein factors

RNase R is a highly conserved exoribonuclease. This enzyme is responsible for the degradation of tmRNA in \textit{C. crescentus} when there is no SmpB [101]. It is also important for nonstop mRNA decay in \textit{E. coli} [93, 94]. RNase R is required for the expression of virulence genes in \textit{Shigella flexneri} [105, 106], which may be related to the tmRNA pathway.
Ribosomal protein S1 can aid translation initiation by interacting with mRNA templates. S1 can bind to tmRNA both *in vitro* and *in vivo*, and may help tmRNA association with the ribosome [107]. However, S1 mutations have little effect on tmRNA activity [108].

VI. Physiology of the tmRNA pathway

The tmRNA pathway is highly conserved throughout the bacterial kingdom. It is important not only for protein quality control, but also for gene regulation. The tmRNA pathway plays a central role in several processes including ribosome rescue, tagged protein degradation and the nonstop mRNA decay. These processes have different roles in bacterial physiology.

VI a. tmRNA pathway and bacteria pathogenesis

The tmRNA pathway is important for bacterial pathogenesis. In *Neisseria gonorrhoeae*, charged tmRNA is essential for viability [109]. In *Yersinia pseudotuberculosis*, if the *ssrA* or *smpB* gene is deleted, bacteria do not cause lethality in mice, and cannot proliferate in macrophages [110]. In *Salmonella typhimurium*, the *ssrA* and *smpB* genes are required for full virulence in mice [99], and mutations in *smpB* reduce bacterial survival in macrophages [111]. In *Shigella flexneri*, lack of tmRNA is lethal (Keiler lab, unpublished data). However, the molecular mechanism of the tmRNA pathway in the pathogenesis of bacteria is still not elucidated. One possibility is that the entry into host cells and the initiation of pathogenesis require large changes in gene expression profile, and the tmRNA pathway can facilitate these processes by its ribosome rescue, tagged protein degradation, and nonstop mRNA decay functions [59].
VI b. tmRNA pathway and bacteria stress response

In *E. coli*, the *ssrA* mutant has mild but multiple phenotypes. The cells have delayed recovery from stationary phase [95] and are more sensitive to amino acid starvation [72], heat shock [112] and antibiotic treatment [113]. Additionally, the *ssrA* deletion strain of *E. coli* has delayed induction of the *lac* operon in diauxic growth, because *lacI* autoregulation is disrupted [69]. The *ssrA* deletion *E. coli* also has lower mobility on a semisolid agar plate [12] and is defective in bacteriophage lytic growth [96]. While these phenotypes are varied, they all relate to stress responses and involve large changes in gene regulation profiles. In *B. subtilis*, the tmRNA pathway will affect stress responses to temperature changes, high concentrations of ethanol, or cadmium chloride [114, 115]. In *Y. pseudotuberculosis*, disruption of *ssrA* or *smpB* leads to increased sensitivity to oxidative and nitrosative stresses, and antibiotic treatments [110].

VI c. tmRNA pathway and bacterial development

In *C. crescentus*, the tmRNA pathway is required for correct timing of chromosomal DNA replication initiation [98]. When *ssrA* or *smpB* is deleted, the initiation of chromosomal DNA replication has a specific delay. Our studies showed that tmRNA can regulate the expression of the *dnaA* gene, which encodes the replication initiator protein DnaA. It also regulates the transcription from a strong promoter located within the origin. The disruption of *dnaA* regulation is responsible for the replication initiation delay. These results will be demonstrated and discussed in Chapter 2. Also, *C. crescentus* lacking tmRNA activity cannot maintain a set of plasmids derived from pBBR1 (The Ph.D. dissertation of Sue-Jean Hong), although the mechanism is still not known. Of interest is the fact that the cell-cycle delay phenotype cannot be complemented by the tmRNA variant tmRNA-DD [98], but the plasmid
maintenance defect can be rescued (The Ph.D. dissertation of Sue-Jean Hong). One possible reason is that both tagging and proteolysis are important for the cell-cycle regulation, but only tagging is important for plasmid maintenance. In *B. subtilis*, the tmRNA pathway will tag and degrade the *kinA* translation in undomesticated strain ACTT6051, and inhibit sporulation of ACTT6051 in the rich sporulation medium [116]. In *Bradyrhizobium japonicum*, the tmRNA pathway is important for symbiosis. *B. japonicum* lacking tmRNA can still form root nodules, but cannot differentiate to the nitrogen-fixing form [117, 118]. In *Thermotoga maritima*, tmRNA levels increase during biofilm formation [118].

In general, the tmRNA pathway has important roles in bacterial pathogenesis, stress response and development. Deletion or mutation in the *ssrA* or *smpB* gene results in various phenotypes in different species of bacteria. The various phenotypes may relate to the presence of different alternative pathways in different species. The alternative pathway might partially complement the defect of lack of tmRNA and ensure cell survival. This model would explain why deletion of tmRNA is lethal in some species but not others. Also, tmRNA activity is required more when there are large changes in gene expression profiles, such as switching of carbon source, stress response or starvation, entry to stationary phase, or treatment with antibiotics. It seems that the main role of the tmRNA pathway is to ensure that the bacterium responds to environmental stimuli promptly and maintains homeostasis effectively at the same time.

VII. **tmRNA pathway is a good target for novel antibiotic development**

Although tmRNA is conserved in all sequenced bacteria, there are no current commercial antibiotics targeting the tmRNA pathway. The tmRNA pathway is essential for some species of pathogenic bacteria, and also increases the sensitivity of some bacteria to antibiotic treatment. In
addition, there is no homolog of tmRNA in humans. Therefore, development of novel antibiotics targeting the tmRNA pathway is promising. The antibiotics could be broad-ranged, with low side effects, and could be combined with current drugs for treating infectious diseases. The antibiotics could either target the tagging step, or the proteolysis step. We have successfully developed strategies to look for inhibitors of ClpXP using genetic screens [89]. This strategy will be discussed in Chapter 3.

VIII. Future directions

Although many of the details of the tmRNA pathway have been elucidated in multiple species of bacteria, there are still important open questions. For example, how does the tmRNA pathway regulate the pathogenesis of bacteria? How is tmRNA integrated into the global gene regulation network? And why does the deletion of the ssrA gene lead to various phenotypes in different species of bacteria? These studies will contribute to not only further understanding of bacterial physiology, but also the development of novel antibiotics targeting the tmRNA pathway.
References


Chapter 2

tmRNA in the Genetic Network of Cell Cycle Regulation
Abstract

tmRNA is a unique RNA molecule transcribed from the ssrA gene. Its trans-translation function is important for both protein quality control and gene regulation in bacteria. In Caulobacter crescentus, deletion of ssrA causes a specific delay in the initiation of chromosomal DNA replication. To understand the molecular basis of the tmRNA phenotype, the impacts of tmRNA activity on dnaA regulation and transcription from the origin-proximal strong promoter (Ps) were investigated. The dnaA gene encodes the DnaA protein, which essential for replication initiation. Ps is located within the replication origin (Cori), and transcription from this promoter may help replication initiation. Our studies demonstrated that transcription of dnaA was disrupted by the ssrA deletion, and the synthesis and accumulation of DnaA protein were consequently disrupted. We excluded the possibilities that tmRNA regulates dnaA transcription through DnaA degradation or DnaA boxes. Furthermore, we identified a TCAA motif in the dnaA promoter that is important for the timing of dnaA transcription. We also demonstrated that transcription from Ps was disrupted by ssrA deletion. A possible model is that tmRNA can regulate an unknown transcriptional regulator, and the unknown transcriptional regulator could bind to the TCAA motif in both the PdnaA and Ps, thereby regulating the timing of transcription from both and ensuring the correct timing of DNA replication initiation. In addition, our data showed that the 5'-UTR of dnaA can strongly repress dnaA expression. This repression could be a conserved mechanism on dnaA regulation in other bacteria.
Introduction

tmRNA is a special RNA molecule having both tRNA and mRNA properties. It is found in all bacteria, some chloroplasts and mitochondria [1]. The 5’- and 3’- ends of tmRNA fold into a tRNA-like structure [2, 3] and can be charged with alanine by alanyl-tRNA synthetase [2]. The open reading frame (ORF) within tmRNA encodes a peptide tag containing degradation determinants for multiple proteases [4-8]. tmRNA is transcribed from the ssrA gene. The current model of tmRNA function is trans-translation [4]. When ribosomes stall during translation, the charged tmRNA forms a complex with SmpB and EF-Tu-GTP [9, 10], recognizes and enters the ribosomal A site, and switches the open reading frame from the nonstop mRNA to the ORF within tmRNA. Then translation of the tmRNA ORF adds a tag to the C terminus of the nascent peptide chain for rapid proteolysis. This process can avoid the accumulation of detrimental truncated proteins and rescue the stalled ribosomes [4, 11]. The tmRNA pathway is important for protein quality control in all bacteria. Lack of tmRNA activity causes various phenotypes in different species of bacteria. For example, deletion of ssrA causes a specific delay in DNA replication initiation in C. crescentus. Cell cycle related processes after initiation of DNA replication were all consequently delayed, and the cell cycle was longer than in the wild type [12]. This phenotype suggests that tmRNA is also important in gene regulation.

To further understand the role of tmRNA in the regulation of gene expression, we took advantage of the ssrA deletion phenotype in C. crescentus, and studied how tmRNA is integrated into the global cell cycle regulatory network. C. crescentus is a model system for bacterial cell cycle regulation studies. As Fig. 2-1A shows, C. crescentus is a dimorphic aquatic a-proteobacterium with swarmer and stalked cell states [13]. Swarmer cells have a single polar flagellum and are motile [14-17], but they cannot initiate DNA replication until they differentiate into sessile stalked cells [18-21]. Stalked cells initiate chromosomal DNA replication once and
only once per cell cycle [22]. They also shed the flagellum and chemotaxis apparatus, and assemble a stalk containing a holdfast at the same pole [23-26]. Each stalked cell then differentiates into a predivisional cell, and generates a progeny swarmer cell and a progeny stalked cell after an asymmetric cell division [13, 16, 27]. The progeny stalked cell will initiate DNA replication immediately, while the progeny swarmer cell needs to differentiate into a stalked cell before DNA replication initiation [18]. Therefore, the morphological transition from swarmer to stalked cell is coincident with the G1 to S phase transition [18, 19]. In addition, a population of G1 phase swarmer cells can be isolated easily by density gradient centrifugation [28]. Once the swarmer cell population is released to fresh medium, it will go through the cell cycle synchronously [28]. All these characteristics of *C. crescentus* make cell cycle studies practical and convenient.

The current model of *C. crescentus* cell cycle regulation is driven by the temporal and spatial expression of three master regulators: DnaA, GcrA, and CtrA. DnaA is the replication initiator found in all bacteria. In *E. coli*, DnaA binds to the DnaA boxes (the DnaA binding site) within the origin of replication (*oriC*), unwinds the AT-rich region, and recruits other proteins such as primase and helicase to form the replisome to initiate DNA replication [29-34]. In *C. crescentus*, DnaA is essential [35]. It is believed to function with the same mechanism as *E. coli* DnaA. The peak amount of DnaA protein is observed just before DNA replication initiation [36], and decreases afterward. The degradation of DnaA is dependent on ClpP [37], but the details are still unknown. DnaA is also a transcriptional activator that controls the expression of more than 40 genes including *gerA* [38, 39]. GcrA is a global transcriptional regulator that regulates 125 genes, and 49 of them are cell cycle regulated [40]. GcrA is transcribed at a low level in swarmer cells, and reaches a maximum in stalked cells [40]. GcrA represses *dnaA* expression, and activates the transcription of *ctrA* [40]. CtrA protein is a response regulator that can be phosphorylated and regulates the transcription of 55 operons in *C. crescentus* [41]. Both its
phosphorylation and its proteolysis are important for cell cycle control [42-46]. CtrA represses gcrA transcription [39, 40]. It is also responsible for the regulation of multiple physiological processes such as flagellum synthesis, DNA methylation, and cell division. In addition, CtrA functions as a DNA replication inhibitor in swarmer cells by binding to five CtrA binding sites within C. crescentus origin (Cori) [21]. The proteolysis of CtrA is necessary but not sufficient for replication initiation [12, 21, 35]. CtrA accumulates in the swarmer cell compartment before cell division, is eliminated before the G1-S transition, and is synthesized again after DNA replication initiation in stalked cells [41]. This dnaA-gcrA-ctrA genetic circuit could control the cell cycle in C. crescentus [47-49].

Several events must occur for DNA replication to initiate in C. crescentus: 1) degradation of the replication repressor CtrA protein [21], 2) synthesis of the replication initiator DnaA protein [35], and 3) transcription from a strong promoter (P,) within Cori [50]. Any of these processes could be regulated by tmRNA. Our former study showed that in C. crescentus lacking tmRNA, the proteolysis of CtrA is not affected [12]. Thus, to understand the molecular basis of the tmRNA phenotype, the impacts of tmRNA activity on dnaA regulation and P, transcription were investigated.

Multiple factors have been reported relating to the expression of dnaA in C. crescentus. The dnaA promoter has the canonical -35 and -10 sequence for σ73, the major sigma factor in C. crescentus [50, 51]. The binding of σ73 and RNA polymerase at this site could be responsible for the basal level of dnaA transcription throughout the cell cycle. There are also two CcrM methylation sites in the dnaA promoter [49]. CcrM is the only methyltransferase in C. crescentus [52]. Its transcription is activated by CtrA [22, 41, 52]. CcrM methylates the adenine in the GANTC sequence [52, 53]. The C-GT mutation (the transcription start site is +1) in methylation site 2 (Met2) can affect dnaA gene expression [52, 54]. In addition, there are two DnaA boxes in dnaA, one is between the transcription start site and the Shine-Dalgarno sequence, the other one is
in the coding sequence from 800 to 809 bp [36]. The DnaA protein-DnaA box binding could have a role in the autoregulation of dnaA.

In addition to DnaA, transcription from P_s is important for DNA replication initiation. P_s is a promoter inside Cori and upstream of the hemE gene [50]. Transcription from the P_s does not lead to synthesis of protein because of the secondary structure in the RNA leader sequence, but it may help unwind the negatively supercoiled DNA [50]. P_s transcription has an auxiliary role in chromosomal DNA replication initiation, and is essential for Cori plasmid replication [55]. CtrA is a repressor of P_s and proteolysis of CtrA is necessary for P_s transcription [50]. Depletion of DnaA does not affect the transcription from P_s [35].

We report here that tmRNA regulates the transcription of dnaA and P_s, and through its activity ensures the correct timing of DNA replication initiation.
Results

The cell cycle regulation of dnaA transcription is disrupted in the ΔssrA strain

*C. crescentus* strains lacking tmRNA activity due to a mutation in either *ssrA* or *smpB* grow slowly and do not initiate DNA replication at the correct time [12]. To identify likely mediators of this phenotype, microarray assays were used to examine the effects of tmRNA activity on the timing of cell-cycle regulated gene expression. Swarmer cells were isolated from wild-type and ΔssrA cells and allowed to pass synchronously through the cell cycle. Cells were harvested from the synchronized cultures at various times and gene expression was compared to an unsynchronized culture of the same strain. The microarray data for *dnaA*, *gcrA*, and *ctrA* genes are plotted in Fig. 2-1B. Data from the wild-type cells followed the expected pattern, but in the *ssrA* deletion strain, expression of *dnaA* was delayed, resulting in *dnaA* RNA peaking at the same time as *gcrA* RNA. Since the delay disrupts the *dnaA*-gcrA cascade, this result raises the possibility that changes in *dnaA* expression may be responsible for the cell cycle phenotype of the ΔssrA cells.

To confirm the microarray results, quantitative RT-PCR was performed. Synchronized cells of both wild type and the *ssrA* deletion strains were collected at various times throughout the cell cycle. The mRNA level of *dnaA* was determined using Taqman probe specific for *dnaA*. 16S rRNA was used as endogenous control. In wild-type cells, the amount of *dnaA* mRNA peaked at 15 min, then decreased to the basal level, and increased again at the beginning of the next cell cycle (Fig. 2-1C, in blue). In ΔssrA cells, the maximum amount of *dnaA* mRNA was delayed by about 30 min (Fig. 2-1C, in purple), which is consistent with the microarray data.
Figure 2-1: The cell cycle regulated dnaA mRNA level is disrupted in ΔssrA cells.

A) *C. crescentus* starts the cell cycle as a swarmer cell, then differentiates into a stalked cell and develops to a predivisional cell. After asymmetric cell division, the predivisional cell generates a progeny swarmer cell and a progeny stalked cell. The progeny stalked cell can initiate DNA replication right away, while the progeny swarmer cell needs to differentiate into a stalked cell first. The swarmer to stalked transition is coincident with DNA replication initiation (G1-S transition).

B) In *C. crescentus*, the expression of dnaA activates gcrA, and GcrA activates ctrA. This cascade could ensure the correct timing of cell cycle regulated gene expression. Microarray data showed that the cascade was normal in wild-type cells. While in ΔssrA cells, the peak level of dnaA mRNA was as late as the gcrA mRNA peak. (Keiler lab,
C) The quantitative RT-PCR result confirmed the microarray data. In wild-type cells, the dnaA mRNA level increased at the beginning of cell cycle, then peaked just before DNA replication initiation, and decreased after initiation of DNA replication. In ΔssrA cells, the peak of dnaA mRNA was delayed by 30 min, which is consistent with the delay of DNA replication initiation.

The changes in mRNA level could be caused by alterations of transcriptional regulation, and/or the differences in mRNA stability. We did not test the mRNA stability of dnaA, because the mRNA level of dnaA in wild-type cells is relatively low, and is difficult to detect by Northern blot. Our strategy was to use a pulse-labeling experiment to determine the transcription levels from the dnaA promoter throughout the cell cycle. A P_{dnaA-lacZ} transcriptional fusion reporter (plasmid pLC32) was transformed into both wild type and the ssrA deletion strains. The cells were synchronized and collected at various times. The pulse-labeling ensured that only newly synthesized proteins were labeled by $[^{35}\text{S}]-\text{methionine}$ at each time point. Then total β-galactosidase (β-gal) was immunoprecipitated and separated on SDS-PAGE, and newly synthesized β-gal can be detected by phosphor screen. In wild-type C. crescentus, the synthesis of β-gal is low at the beginning of the cell cycle, and reaches a maximum just before the initiation of DNA replication, then decreases after initiation and remains at a low level until the start of the next cell cycle (Fig. 2-2, in blue). When ssrA was deleted, the curve of β-gal synthesis retained a similar shape to that of wild-type cells. However, the maximum synthesis occurred 30 min later (Fig. 2-2, in purple). Since DNA replication initiation was also delayed by 30 min, the maximum was still just before replication initiation. Assuming that translation of P_{dnaA-lacZ} is the same in both wild-type and ΔssrA cells, the amount of newly synthesized β-gal at each time point reflects the transcription activity of P_{dnaA}. Synthesis of β-gal throughout cell cycle could reflect the timing of dnaA transcription. Therefore, delay of the maximum synthesis of β-gal means that the timing
of *dnaA* transcription was delayed by the *ssrA* deletion. These data are consistent with both the microarray and the quantitative RT-PCR results.

Figure 2-2: The transcription of *dnaA* gene is disrupted by the *ssrA* deletion.

The transcriptional fusion reporter on pLC32 (pRK-P~*dnaA*-369+1-lacZ) was used for pulse-labeling experiments in both wild-type and Δ*ssrA* strains. Cells were synchronized and swarmer cells were released to fresh M2G medium. Newly synthesized proteins were labeled by [35S]-methionine at each time point. Then β-gal was immunoprecipitated by β-gal antibody, and separated by SDS-PAGE. Gels were dried and exposed to phosphor screen. The relative band intensity was plotted. The transcription from the *dnaA* promoter had the same cell cycle regulation pattern as the *dnaA* mRNA level from quantitative RT-PCR. The peak of *dnaA* transcription was delayed by about 30 min in Δ*ssrA* cells.
Both the synthesis and the accumulation of DnaA are disrupted by the ssrA deletion

The transcription from \( P_{dnaA} \) was disrupted in the \( \Delta ssrA \) strain, but it is still possible that \( dnaA \) expression is regulated by tmRNA at other levels as well, such as translation or post translation. To determine the effect of tmRNA on temporal DnaA synthesis during the cell cycle, both wild type and the \( ssrA \) deletion cells were synchronized. Newly synthesized proteins were labeled by \([^{35}S]\)-methionine at each time point, and then DnaA protein was immunoprecipitated using anti-DnaA antibody. In wild-type \( C. crescentus \), the amount of newly synthesized DnaA increased and reached its peak at 15 min, then decreased to the basal level, and increased again at the beginning of the next cell cycle (Fig. 2-3A, in blue). This is the same cell-cycle-regulated pattern as \( dnaA \) mRNA. In the \( ssrA \) deletion strain, the shape of the DnaA synthesis curve was very similar to the curve in the wild type, but the synthesis peak was delayed by about 30 min (Fig. 2-3A, in purple). This difference is consistent with the changes of \( dnaA \) mRNA level in \( \Delta ssrA \) cells. There appears to be no additional defect in protein synthesis.

The accumulation of DnaA was demonstrated by immunoblots with anti-DnaA antibody. Synchronized cultures of both wild type and the \( \Delta ssrA \) strains were collected at various times during cell cycle. As Fig. 2-3B shows, the accumulation of DnaA protein has the same cell-cycle-regulated pattern as \( dnaA \) mRNA in both strains. The peak level of DnaA protein accumulation occurred just before the DNA replication initiation, and was delayed by about 45 min in cells lacking tmRNA activity.

Therefore, the primary defect is at the level of transcription, and any additional defects in synthesis or steady-state levels are relatively small.
A) The synthesis of DnaA protein was tested by pulse-labeling experiments. Cultures of KCK2 (wild type) and KCK116 (ΔssrA) were synchronized and swarmer cells were released to fresh M2G medium. Newly synthesized proteins were labeled by [35S]-methionine at each time point. Then DnaA protein was immunoprecipitated by anti-DnaA antibody. The relative band intensity was plotted. The synthesis of DnaA protein has the same cell cycle regulation pattern as the transcription of dnaA gene. And the peak of synthesis was delayed by about 30 min in the ΔssrA cells as well.

B) The accumulation of DnaA protein was checked by immunoblots. Cultures of KCK2 (wild type) and KCK116 (ΔssrA) were synchronized and swarmer cells were released to fresh M2G medium. Cells were then collected at various times and lysed. The DnaA protein was separated by SDS-PAGE and detected by anti-DnaA antibody. The relative band intensity was plotted. The accumulation of DnaA protein has the same cell cycle regulation pattern as the synthesis. The peak amount of DnaA showed up about 45 min later in the ΔssrA cells. The heat shock protein GroEL can interact with anti-DnaA antibody and appears on the immunoblots as well [37].

Figure 2-3: Both the synthesis and the accumulation of DnaA are disrupted by ΔssrA.
The above results indicate that the temporal transcription of dnaA is regulated by tmRNA. Because tmRNA acts primarily at the level of translation and protein stability, it is more likely that tmRNA affects certain dnaA transcriptional regulators by trans-translation, rather than acting as a transcriptional regulator itself. The tmRNA-dependent transcriptional regulator could be DnaA, or some unknown proteins. There are two DnaA boxes in C. crescentus dnaA, one is in the 5′-UTR, and the other one is within the coding region. It has been reported that in E. coli, binding of DnaA to the DnaA box in the 5′-UTR can repress transcription from the dnaA promoter [56]. In C. crescentus, it is not known whether the binding of DnaA has any autoregulatory effects. If it does, then a possible model is: tmRNA can tag DnaA which would lead to rapid proteolysis of DnaA, thereby relieving negative autoregulation. However, when the ssrA gene is deleted, DnaA cannot be tagged and degraded. Thus, residual DnaA binds to the DnaA box in the 5′-UTR of dnaA and represses its transcription. However, we have excluded this model by testing whether the tmRNA can tag DnaA, whether tmRNA can change the degradation rate of DnaA, and whether the DnaA boxes are important for tagging or repression.

1) tmRNA tags DnaA

To test whether DnaA can be tagged by tmRNA, a plasmid carrying his6-dnaA under the control of a xylose inducible promoter [57], and a plasmid carrying a proteolysis resistant variant of tmRNA (tmRNA-DD) under control of the ssrA promoter were transformed into wild-type C. crescentus. Cells were induced by adding xylose into the medium and the whole cell lysate was used for immunoblots. Two bands were detected by anti-DnaA, anti-His6, and anti-tmRNA-DD antibodies. One is nearly the same size as the full length DnaA, and the other band is about 30
KDa. This result indicates that DnaA can be tagged by tmRNA at two sites, one is near the C terminus and the other is in the middle of the gene (Fig. 2-4).

Figure 2-4: DnaA is tagged by tmRNA.

The left lane is KCK2 without plasmid, which was used as wild-type control. The right lane is KCK2 with pKK842 (tmRNA-DD) and pKK940 (his6-dnaA). The whole cell lysate was separated by SDS-PAGE, and DnaA-DD protein was detected by anti-DnaA antibody. The blot was then incubated with anti-His6 or anti-tmRNA-DD antibody (data not shown). Two His6-DnaA-DD bands were observed with all antibodies. One has near the full length DnaA, another one has about half length of DnaA.

To look for the precise tmRNA tagging sites in DnaA, His6-DnaA was produced and purified, and separated on the SDS-PAGE gel. The 30 KDa band could not be detected on the Coomassie blue stained gel, so only the full length DnaA band was excised and in-gel Asp(N) digestion was performed. The peptides were then detected by MALDI-TOF. The sequence of the tmRNA-DD tag is AANDNFAEEFAVDD. Therefore, the peptide that had been tagged ended with AAN at its C terminus after digestion. Peptides that covered more than 50% of the full length of DnaA were detected by mass spectrometry. One of the peptides DLETLTRKLAAN has both DnaA sequence DLETLTRK and tag sequence AAN. Its observed mass is 1344.68 m/z, which is near to the expected mass 1344.74 m/z. The C. crescentus DnaA has 490 amino acids in
total. This peptide is from the C terminus of DnaA, and the Leu before tag sequence AAN is the 488th amino acid. Therefore, Leu488 is one of the tmRNA tagging sites.

There is a DnaA box from 800 to 808 bp in the dnaA coding sequence. The sequence of this DnaA box is TGTTCCACA, which has 8 out of 9 bases matched to the E. coli DnaA box consensus TTA/TTNCACA [36]. It is possible that C. crescentus DnaA protein can bind to this DnaA box and somehow provides signals for tmRNA tagging. Synonymous mutations (silent mutations) in the DnaA box were made and the sequence was changed to TCTTTTCATA. The tagging of DnaA was determined by immunoblots with anti-DnaA, anti-His6, and anti-tmRNA-DD antibodies. No change of DnaA tagging was detected (data not shown), indicating that the DnaA box is not the motif that leads to tmRNA tagging.

2) The degradation of DnaA is not affected by tmRNA

To check whether the tagging affects the degradation rate of DnaA, a pulse chase experiment was carried out in cells with or without tmRNA activity. In wild-type C. crescentus, the DnaA half life was 29 ± 4 min. In the ssrA deletion strain, it was 34 ± 4 min (Fig. 2-5). Therefore, there is not a large difference between these two strains in the rate of DnaA degradation.

Moreover, dnaA mutations with either the tagging site changed to a stop codon (L488Z) or with the tmRNA-AA or DD tag encoded in the gene were made. If the truncated or tagged DnaA is responsible for the temporal transcription of dnaA, then expression of mutant DnaA in ΔssrA cells should be able to rescue the ssrA deletion phenotype. However, overexpression of these mutations did not rescue the delay of replication initiation in ΔssrA cells (data not shown). Therefore, the tagging does not change the degradation of DnaA, and does not affect the transcription of dnaA.
Figure 2-5: The degradation of DnaA is not affected by the ssrA deletion.

The half life of DnaA in both KCK2 (wild type) and KCK116 (ΔssrA) were determined by pulse-labeling experiments. Cells were grown in M2G until mid log phase, and pulsed with [35S]-methionine, then chased with unlabeled methionine, and collected at various times. DnaA was then immunoprecipitated, separated by SDS-PAGE, and detected by phosphor screen. The relative band intensity was plotted. The half life of DnaA protein in wild-type cells was 29 ± 4 min, and 34 ± 4 min in ΔssrA cells. There was not a large difference between these two strains on the degradation rate.
3) Mutations of DnaA box in the 5'-UTR does not change the temporal transcription of dnaA in the ΔssrA strain

The DnaA box in the 5'-UTR of dnaA gene is 18 bp downstream of the transcription start site, and upstream of the dnaA Shine-Dalgarno sequence. In E. coli, it has been reported that binding of DnaA to the DnaA box in the 5'-UTR of dnaA will lead to transcription termination [56, 58]. If C. crescentus uses the same strategy, mutations that disrupt the DnaA box will change transcription of dnaA. We mutated the DnaA box from TTATCCAAG to TGAGCGCAG, and tested whether the dnaA transcription changed. Both the dnaA promoter and the 5'-UTR of dnaA were inserted upstream of a lacZ reporter with its own Shine-Dalgarno sequence to make transcriptional fusions, and the β-gal activity was measured and compared. As Fig. 2-6A shows, the reporter with the mutant DnaA box (pLC17) had roughly 2-fold higher β-gal activity than the reporter with the wild type DnaA box (pLC20), and the increase is the same whether the cells have tmRNA or not. Hence, the DnaA box represses dnaA expression, but the repression is not related to tmRNA activity. On the other hand, no matter which reporter was used, the β-gal activity was always lower in ΔssrA cells, which means loss of tmRNA represses dnaA expression, but this effect is not related to the DnaA box. Furthermore, the β-gal synthesis of the reporter with the mutant DnaA box (pLC17) was tested in ΔssrA cells. The cell-cycle-regulated pattern of β-gal synthesis was not affected (Fig. 2-6B), which indicates that the temporal transcription from P_dnaA is not affected by mutation of the DnaA box in the 5'-UTR of dnaA.
Figure 2-6: The DnaA box in the 5'-UTR of dnaA is not responsible for the ΔssrA phenotype.

A) There is a DnaA box on the antisense strand in the 5'-UTR of dnaA gene. The relative β-gal activity of pLC20 (wild-type DnaA box) and pLC17 (mutant DnaA box) was plotted. Mutations in DnaA box were shown in red. The mutant DnaA box led to a 1.58-fold increase of β-gal activity in wild-type cells and a 2-fold increase in ΔssrA cells.

B) The pLC17 was used for a pulse-labeling experiment in the ΔssrA strain. Compared to pLC32 (P\textsubscript{dnaA-lacZ}, no 5'-UTR) in ΔssrA cells, the timing of transcription from P\textsubscript{dnaA} was not affected by the DnaA box mutation. Here pLC32 was used for the assay instead of pLC20, because reporter on pLC20 has low expression and was difficult to detect.
A motif in dnaA promoter is responsible for the ssrA deletion phenotype

To look for the tmRNA-dependent transcriptional regulators of dnaA expression, we started with studies of important cis-acting elements in the dnaA promoter. Once the cis-acting elements are elucidated, they will help us to identify the trans-acting factors.

There are two CcrM recognition sites (GANTC) within the dnaA promoter. Methylation site 1 (Met1) is from -160 to -156 bp, and Met2 is from -75 to -71 bp (relative to the transcriptional start site). Mutation in Met2 decreased the activity of the dnaA promoter [54]. To test whether Met2 is responsible for the ΔssrA phenotype, we constructed a set of point mutations in Met2 G-75A-74G-73T-72C-71. The relative β-gal activities of these point mutations are shown in Fig. 2-7A. The C-71T mutation decreased the P_{dnaA} activity, consistent with published result [54]. The P_{dnaA} activity was decreased about 6-fold by C-71T mutation, and the T-72G mutation had a similar effect. The G-73T mutation decreased the promoter activity about 2-fold. Surprisingly, the A-74T mutation, which is the site of CcrM-mediated methylation [52, 53], had little if any effect on promoter activity. Therefore, the methylation state of Met2 is not important for P_{dnaA} activity, but the T-72C-71 sequence is important. To further characterize the cis-acting element, we made point mutations both upstream and downstream of Met2. The relative β-gal activities of these mutations are shown in Fig. 2-7A as well. When the upstream sequences G-77 or C-76 was mutated, the transcription activity of P_{dnaA} was not affected. However when the downstream bases A-70 or A-69 was mutated, the P_{dnaA} transcription activity decreased about 5-fold. This result indicates that the 5’-TCAA-3’ motif is important for dnaA promoter activity. More point mutations of downstream sequence will be tested to identify the full length motif.

To test whether the TCAA motif is responsible for the delay of dnaA transcription in ΔssrA cells, lacZ transcriptional fusions with P_{dnaA} (pLC32) or the C-71T mutant P_{dnaA*} (pLC33) were used for pulse-labeling experiments. As formerly mentioned, the transcription from P_{dnaA}
was delayed by 30 min in ΔssrA cells (Fig. 2-2, Fig. 2-7B). When the C.71T mutant promoter
PdnaA* (pLC33) was used in ΔssrA cells, the cell-cycle-regulated pattern of β-gal synthesis
occurred 15 min earlier (Fig. 2-7B). Pulse labeling experiments of these two reporters were also
performed in cells with tmRNA, and the C.71T mutation did not affect dnaA transcription (data
not shown). Therefore, the C.71 is important for the timing of dnaA transcription in ΔssrA cells.
We have not tested other mutations for the temporal transcription of dnaA. Assuming that all
bases in this motif have the same effect on PdnaA transcription timing as on PdnaA transcription
intensity, then this TCAA motif is important for the temporal transcription of dnaA, and the C.71T
mutation can partially restore dnaA transcription in ΔssrA cells.

If the delay of dnaA transcription is the only cause for the ΔssrA phenotype, then C.71T
mutation should be able to partially rescue the mutant phenotype, because it can partially restore
dnaA transcription. We tested this possibility using growth curve experiments (Fig. 2-8). When
tmRNA is active, cells bearing plasmid pLC51 (C.71T mutant PdnaA*-dnaA) grew at the same rate
as the control which has pRKlacZ empty vector. The doubling time was 136 min. When there is
no tmRNA, cells with pRKlacZ had a doubling time of 165 min. However, ΔssrA cells with
plasmid pLC51 grew faster. The doubling time is 146 min. Cells with pLC33 (C.71T mutant PdnaA*
-lacZ) were used as control as well, because they have high yield of β-gal, which may cause
problem when lacking tmRNA for protein quality control. The result showed that the expression
of extra β-gal did not affect the growth rate of ΔssrA cells. Therefore, expressing DnaA under the
control of dnaA promoter with the C.71T mutation can partially rescue the growth defect in ΔssrA
cells. The ΔssrA cells grow slower because the DNA replication initiation is delayed, so the
replication initiation time of ΔssrA cells bearing pLC51 will be tested with synchronies to
confirm the rescue.
Figure 2-7: The TCAA motif is important for P\textsubscript{dnaA} activity and timing of dnaA transcription.

A) Point mutations in P\textsubscript{dnaA} were used for $\beta$-gal activity assays. pLC20 has lacZ under the control of wild type P\textsubscript{dnaA}. The TCAA motif (in red circle) affected the transcription activity of P\textsubscript{dnaA}. Met2 site is in black box.

B) The ssrA deletion cells with pLC33 plasmid were used for pulse-labeling experiments. pLC33 has P\textsubscript{dnaA-369+1} C\textsubscript{71} T lacZ. Compared to pLC32 which has P\textsubscript{dnaA-369+1} -lacZ, C\textsubscript{71} T mutation led to earlier transcription in ΔssrA cells.
Figure 2-8: C\textsubscript{71}T mutation in P\textsubscript{dnaA} can partially rescue the \textit{ΔssrA} phenotype.

Growth curve experiments were performed and the doubling time of each plasmid in both wild type and the \textit{ssrA} deletion strains were plotted. pRKlacZ has a promoterless \textit{lacZ}, pLC51 has P\textsubscript{dnaA-369+1} \textit{C-71 T}\textsuperscript{-}\textit{dnaA}, and pLC32 has P\textsubscript{dnaA-369+1} \textit{C-71 T}\textsuperscript{-}\textit{lacZ}. All constructs have doubling time about 136 min in cells with tmRNA activity. Without tmRNA activity, the doubling time will be longer because the DNA replication initiation is delayed by 30 min. But cells having pLC51 grew faster than cells with the other two plasmids. The doubling time was shortened from 165 min to 146 min.

The 5'-UTR of the \textit{dnaA} gene represses \textit{dnaA} expression

During the study of \textit{dnaA} transcription, we noticed that \textit{dnaA} has a 155 bp long 5'-UTR. To test whether the leader sequence of \textit{dnaA} can regulate \textit{dnaA} expression, we constructed \textit{lacZ} transcriptional fusion reporters with P\textsubscript{dnaA} and truncations of the leader sequence (Fig. 2-9A). The leader sequence was truncated from its 3'-end. The relative β-gal activity of these constructs are shown in Fig. 2-9A. Interestingly, the β-gal activity was not affected when the truncated leader sequence was 87 bp or longer. But a leader sequence of 67 bp or less resulted in dramatically increased activity, about 10- to 20-fold depending on the length of truncated leader sequence. Therefore, the leader sequence exhibits strong repression of \textit{dnaA} expression, and the 20 bp sequence between 87 bp and 67 bp downstream from the transcription start site is required for this repression.
The activity of all constructs was tested in the *ssrA* deletion strain as well. Without repression from the leader sequence, all reporters have a slightly higher activity in Δ*ssrA* cells than in the wild type. This difference could be due to the relatively large sample errors. In addition, when the timing of β-gal synthesis in Δ*ssrA* cells was tested by pulse-labeling experiment, there is no difference between reporters with or without the leader sequence (data not shown). Therefore, the leader sequence does not affect the timing of *dnaA* expression in Δ*ssrA* cells.

To test whether the repression by the leader sequence is specific for *dnaA*, we added the 140 bp leader sequence to *lacZ* transcriptional fusion under the xylose inducible promoter P<sub>xyl</sub> (pLC41), and compared its activity to P<sub>xyl-lacZ</sub> with no leader sequence (pLC42). As Fig. 2-9B shows, insertion of the leader sequence decreased the reporter activity about 50-fold. Hence, the *dnaA* leader sequence has ubiquitous repression in *C. crescentus*. It represses not only the gene expression of *dnaA*, but also the expression of unrelated genes such as P<sub>xyl-lacZ</sub>.

DnaA is highly conserved in all bacteria. It is possible that other bacteria use similar mechanism to regulate *dnaA* expression. *E. coli* *dnaA* gene has a 151 bp long 5'-UTR. When this leader sequence is integrated into the *E. coli* P<sub>dnaA-lacZ</sub> reporter (pLC43), the relative β-gal activity was 5-fold lower than the activity of *E. coli* P<sub>dnaA-lacZ</sub> without the leader sequence (pLC44). Both reporters were examined in Δ*lacX74* MG1655 with or without tmRNA activity (Fig. 2-9C), and there was no obvious change caused by *ssrA* deletion. Therefore, the leader sequence in *E.coli dnaA* can also repress *dnaA* gene expression.
Figure 2-9: The leader sequence in dnaA can repress the expression of dnaA.

A) The 5’-UTR of dnaA gene has a leader sequence between transcription start site (+1) and Shine-Dalgarno sequence (+140). Serial deletions of the leader sequence were inserted into P_{dnaA}-lacZ transcriptional fusion. The length of the leader sequence in each plasmid is shown with a bar. Relative β-gal activity of each plasmid in cells with or without ssrA was plotted. The leader sequence showed strong repression when it is longer than 87 bp. The 20 bp between +67 and +87 is required for repression.

B) The 140 bp leader sequence from C. crescentus dnaA was integrated into the P_{xyr}-lacZ reporter (pLC41), and its β-gal activity was compared to the reporter without the leader sequence (pLC42). The activity of pLC42 was about 50-fold higher than pLC41.

C) The 151 bp leader sequence from E.coli dnaA gene was integrated into E.coli P_{dnaA}-lacZ reporter (pLC43), and its β-gal activity in ΔlacX74 MG1655 with or without ssrA was compared to the reporter without leader sequence (pLC44). The activity of pLC44 was about 5-fold higher than pLC43.
**Pₘ transcription is delayed by the ssrₐ deletion**

Since tmRNA can regulate *dnaA* expression at the transcriptional level, it is possible that it regulates other genes important for DNA replication initiation. Transcription from the strong promoter (Pₘ) does not lead to any protein product, but it may help to unwind the negative supercoiling near *Cori* and therefore has an auxiliary role in replication initiation [50]. To explore the effect of *ssrₐ* deletion on Pₘ transcription, *lacZ* under the control of Pₘ was used as a reporter. The synchronized culture of both wild type and the *ssrₐ* deletion strains were examined. Newly synthesized proteins were labeled by [³⁵S]-methionine at each time point, and β-gal was immunoprecipitated. The timing of Pₘ transcription was coupled with CtrA degradation and peaked before DNA replication initiation in the wild type. In the *ssrₐ* deletion strain, CtrA was degraded normally [12], but Pₘ transcription did not reach its peak until 30 min later. This change in timing of Pₘ transcription was the same as the change in timing of *dnaA* transcription, and the delay of replication initiation (Fig. 2-10). This result indicates that tmRNA can regulate unknown transcription regulators and thereby ensures the correct timing of Pₘ transcription.

The *dnaA* gene is essential for chromosomal DNA replication initiation, and Pₘ transcription can help replication initiation. Transcriptions from P₃ₐₜₜ and Pₘ have the same cell cycle regulated pattern, and both can be delayed for 30 min by the *ssrₐ* deletion. This similarity leads to the assumption that they may share the same tmRNA-dependent transcriptional regulator. We noticed that there is one TCAA motif 2 bp upstream of the -35 region of Pₘ. Whether this motif is important for Pₘ transcription is still under investigation. tmRNA could regulate an unknown transcriptional regulator by *trans*-translation, and the transcriptional regulator could bind to the TCAA motif in both P₃ₐₜₜ and Pₘ and regulate the transcription from both, thereby ensuring the correct timing of DNA replication initiation.
Figure 2-10: The transcription from Pₜ was disrupted by the ssrA deletion.

Transcription from the Pₜ-lacZ reporter (pGM976) was tested by pulse-labeling experiments in both wild type and ΔssrA cells. Cells were synchronized and swarmer cells were released to fresh M2G medium. Newly synthesized proteins were labeled by [³⁵S]-methionine at each time point, and β-gal was immunoprecipitated. The relative band intensity was plotted. The peak transcription from Pₜ was delayed by about 30 min in ΔssrA cells.
**Discussion**

Here we report that tmRNA can regulate transcription of the *dnaA* gene and the strong promoter P<sub>s</sub>, and therefore ensure the correct timing of DNA replication initiation in *C. crescentus*. Although the tmRNA-dependent transcriptional regulator has not been isolated, we identified the *cis*-acting element in P<sub>dnaA</sub> that is partially responsible for the delay of DNA replication initiation in ΔssrA cells. The regulatory role of tmRNA in DNA replication initiation would have some advantages. In wild type cells under normal conditions, tmRNA regulates the unknown transcriptional regulator of DnaA, and ensures the correct timing of DNA replication initiation. When cells are under stresses such as starvation, mRNA cleavage occurs and results in nonstop mRNAs, stalled ribosomes, and truncated translation products [59]. tmRNA can sense these problems, use its *trans*-translation function to release stalled ribosomes, tag truncated proteins for degradation, and lead to the decay of nonstop mRNAs. When the tmRNA machinery is occupied, the unknown transcriptional regulator of DnaA cannot be regulated on time, thereby delaying the initiation of DNA replication. Only after all the problems are cleaned up, the tmRNA pathway could be available to regulate the unknown transcriptional regulator of DnaA, and DNA replication could be initiated. Therefore, tmRNA may work as a checkpoint, and integrating tmRNA into the global regulation network is important for the bacterial physiology.

The TCAA motif we identified in P<sub>dnaA</sub> is a *cis*-acting element, and it is thought to interact with a *trans*-acting transcriptional regulator. Mutations in this motif decreased the transcription activity of P<sub>dnaA</sub>, suggesting that the mutations disrupted the binding site for an activator. We also found that the C<sub>-71</sub>T mutation in the TCAA motif affects both the intensity and the timing of *dnaA* transcription. Assuming other bases in this motif are also important for correct timing, the effect of the activator on the timing of *dnaA* transcription could be explained as follows. Without binding of the activator, *dnaA* expression is at a basal level and maintains
correct timing. When the activator binds to the motif, it adds another layer of regulation to *dnaA* transcription. The sum of the basal level expression and the activator’s effect will result in the final pattern and level of *dnaA* transcription. In cells with tmRNA, the activation effect overlaps with the basal level expression, such that the timing of *dnaA* transcription is not affected. In ΔssrA cells, the binding of the activator is somehow delayed. This delay leads to a higher but delayed *dnaA* expression compared to the basal level. The total effect on *dnaA* expression is the peak of transcription will be delayed. However, when the cis-acting motif is disrupted by mutations in ΔssrA cells, the activator does not bind to the P<sub>dnaA</sub>, which results in a basal level of *dnaA* transcription with normal timing. This hypothesis can be tested once the activator is identified. Another possibility is that the motif actually contains binding sites for multiple transcription regulators. For example, one regulator might be important for the intensity of P<sub>dnaA</sub> transcription, while another one could be important for the timing of P<sub>dnaA</sub> transcription. The C-71T mutation affects both the intensity and the timing of *dnaA* transcription because cis-acting elements overlap. Identifying important elements in the motif will lead to a better working model for the transcriptional regulation of the *dnaA* gene, and will help us to identify the trans-acting factors.

The TCAA motif in P<sub>dnaA</sub> is important for the correct timing of *dnaA* expression, although the rescue by C-71T mutant P<sub>dnax</sub>-*dnaA* is only partial. There are several possible explanations for this observation. The first explanation is that there is also a chromosomal copy of *dnaA* under the control of wild type P<sub>dnaA</sub>. The transcription of the chromosomal copy is delayed in ΔssrA cells. The delayed transcription from chromosomal copy could interfere with the normal timing of transcription from plasmid copy. The second possibility is that the TCAA motif may have multiple bases important for the binding of a tmRNA-dependent transcriptional regulator. The C-71T mutation alters only one base, and may not completely disrupt the binding site. The third possibility is that the motif we identified is only one of the multiple cis-acting elements
responsible for the correct timing of \textit{dnaA} expression. We will study more mutations in \( P_{dnaA} \) to identify all of the important bases, then make a construct with mutations that disrupt the entire motif, and express the mutant \( P_{dnaA^{-}}-dnaA \) in a strain without interference from wild-type \( P_{dnaA} \). The results from that experiment should tell us whether the motif is the only \textit{cis}-acting element important for the timing of \textit{dnaA} transcription.

In our study, we demonstrated that tmRNA can regulate both the transcription from \( P_{dnaA} \) and the transcription from \( P_{s} \). If \( P_{s} \) shares the same \textit{trans}-acting factor with \( P_{dnaA} \), it should have the same \textit{cis}-acting motif. \( P_{s} \) is located within the \textit{C. crescentus} origin of replication [50], which is a sequence between the \textit{CC0001} and \textit{hemE} genes [60]. We searched the full length \textit{Cori} sequence and found three TCAA motifs upstream of \( P_{s} \). One is only 2 bp upstream from the -35 region of \( P_{s} \), one is 90 bp, and the other one is 291 bp upstream. Mutations in these TCAA motifs have not been tested. If the timing of \( P_{s} \) transcription in \( \Delta ssrA \) cells is changed by mutations in these TCAA motifs, then the \( P_{s} \) does share the same tmRNA-dependent transcriptional regulator as \textit{dnaA}.

We also demonstrated that the 5’-UTR of \textit{dnaA} can strongly repress \textit{dnaA} expression. Within the 5’-UTR, there is a 20 bp sequence between +67 and +87 with respect to the start of \textit{dnaA} transcription that is required for repression. When this 20 bp sequence was deleted, expression of \textit{dnaA} increased more than 10-fold. Whether the upstream 67 bp is necessary for the repression is still unknown. Mutations that delete the 5’-UTR from the left side of the 20 bp will be tested. The result will narrow down the \textit{cis}-acting element responsible for repression. We examined the full length leader sequence of \textit{dnaA} for possible transcription termination signals. We did not find a hairpin structure followed by U rich regions, which is usual for intrinsic transcription terminators [61]. The leader sequence is not C rich either, which is common for \( \rho \)-dependent transcription terminators [62]. The leader sequence is likely to interact with either proteins or small RNAs, and inhibit the transcription or translation of the \textit{dnaA} gene. For
example, binding of proteins to the mRNA leader sequence could switch it to an alternative secondary structure, and inhibit the readthrough of RNA polymerase or sequester the Shine-Dalgarno sequence. The only known protein that binds to adjacent sequence is DnaA. Its binding site, the DnaA box, is 50 bp upstream of the 20 bp sequence. Mutations of the DnaA box only decreased the \textit{dnaA} expression by 2-fold. Therefore, DnaA is not responsible for the 10-fold repression. It is also possible that a small RNA with complementary sequence can bind to the leader sequence at the DNA or RNA level, thereby block the transcription or translation, and/or lead to the degradation of \textit{dnaA} mRNA. We used the RNATarget program [63] to search for potential small RNAs with complementarity to the \textit{dnaA} mRNA leader sequence, and got several hypothetical small RNAs. The small RNA with highest complementarity is between gene \textit{CC0009} encoding the DNA alkylation damage repair protein AlkB, and gene \textit{CC0010} encoding the heat shock protein DnaK. The small RNA was predicted by bioinformatics methods, and its function is unknown. Although it has a sequence that is complementary to the \textit{dnaA} leader sequence, the biological relationship between these two RNAs has not been tested.

The amount of DnaA protein is important for the initiation of DNA replication, which is the basis for correct chromosome segregation and cell division. Overexpression of DnaA can result in additional replication initiation events that are often not complete in \textit{E. coli} [64, 65], and lead to excess DNA content per cell in \textit{C. crescentus} [36]. We examined the \textit{E. coli dnaA} 5′-UTR and found that it repressed \textit{E. coli dnaA} expression as well. The mechanism of repression caused by the \textit{dnaA} 5′-UTR in \textit{C. crescentus} and \textit{E. coli} has not been elucidated, but it could be a conserved mechanism. DnaA is highly conserved in all bacteria, and this mechanism could ensure the amount of DnaA in cells does not exceed a critical level.
Materials and methods

Strains and plasmids

All bacterial strains and plasmids used in this study are listed in table 2-1. Quikchange mutagenesis kit from Stratagene was used and the manufacture protocol was followed. The tmRNA tag was added to the 5’-end of the target gene by PCR amplification with the primers containing –AA or –DD tag codons. *C. crescentus* strains were grown in PYE broth [66], or M2 minimal media [67] with 0.2% glucose (M2G) or 0.3% xylose (M2X) at 30ºC. Spectrometry was used to monitor the increase of culture optical density at 660 nm. Antibiotics were added at the following final concentration where indicated: 5 mg/ml kanamycin, 1 mg/ml chloramphenicol, 1 mg/ml oxytetracycline.

Table 2-1: Plasmids and strains used in this study

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGM976</td>
<td>pRK290-(P_{x-y}^-)lacZ [50]</td>
</tr>
<tr>
<td>pKK842</td>
<td>pRK290- (P_{x-y}^-)ssrA-DD [68]</td>
</tr>
<tr>
<td>pRKlacZ</td>
<td>For making lacZ transcriptional fusion [69]</td>
</tr>
<tr>
<td>pKK940</td>
<td>pML81-his6-dnaA- under the control of (P_{x-y}) This study</td>
</tr>
<tr>
<td>pLC1</td>
<td>pML81-his6-dnaA-L to Z, Leu488 was changed to stop codon This study</td>
</tr>
<tr>
<td>pLC2</td>
<td>pML81-his6-dnaA-L + tmRNA-AA tag This study</td>
</tr>
<tr>
<td>pLC3</td>
<td>pML81-his6-dnaA-L + tmRNA-DD tag This study</td>
</tr>
<tr>
<td>pLC7</td>
<td>pML81-his6-dnaA*, the DnaA box in coding sequence was changed from TGTTCCACA to TCTTTCATA This study</td>
</tr>
<tr>
<td>pLC11</td>
<td>pRK290-(P_{x-y}^-)his6-dnaA-L488Z, Leu488 was changed to stop codon This study</td>
</tr>
<tr>
<td>pLC12</td>
<td>pRK290-(P_{x-y}^-)his6-dnaA-L + tmRNA-AA tag This study</td>
</tr>
<tr>
<td>pLC13</td>
<td>pRK290-(P_{x-y}^-)his6-dnaA-L + tmRNA-DD tag This study</td>
</tr>
<tr>
<td>pLC17</td>
<td>pRK-(P_{dnaA}^-)lacZ, has (P_{dnaA}) from -369 to +164, the DnaA box in 5’- UTR was changed from TTATCCACA to TGAGCGCAG This study</td>
</tr>
<tr>
<td>pLC20</td>
<td>pRK-(P_{dnaA}^-)(369 to +164)-lacZ This study</td>
</tr>
<tr>
<td>pLC21</td>
<td>pRK-(P_{dnaA}^-)(369+164 G-75T)(\gamma)lacZ This study</td>
</tr>
<tr>
<td>pLC22</td>
<td>pRK-(P_{dnaA}^-)(369+164 A-74T)(\gamma)lacZ This study</td>
</tr>
<tr>
<td>pLC23</td>
<td>pRK-(P_{dnaA}^-)(369+164 T-72G)(\gamma)lacZ This study</td>
</tr>
<tr>
<td>pLC24</td>
<td>pRK-(P_{dnaA}^-)(369+164 C-71T)(\gamma)lacZ This study</td>
</tr>
<tr>
<td>pLC25</td>
<td>pRK-(P_{dnaA}^-)(369+164 G-77A)(\gamma)lacZ This study</td>
</tr>
<tr>
<td>pLC26</td>
<td>pRK-(P_{dnaA}^-)(369+164 C-76T)(\gamma)lacZ This study</td>
</tr>
</tbody>
</table>
Synchrony and flow cytometry

*C. crescentus* strains were growing at 30°C in M2G medium with antibiotics to mid-log phase. The warmer cells were then isolated by Ludox (Ludox LS, Du Pont, Wilmington, DE) density gradient centrifugation as described previously [28] and released to fresh medium. The release point was counted as 0 min. Culture was then collected every 15 min throughout the entire cell cycle. For wild-type *C. crescentus*, one cell cycle was about 150 min, while in *ssrA* deletion *C. crescentus*, it was around 180 min.

Collected cultures were then treated as described in the following sections. Part of the culture from each time point was incubated with 150 μg/ml rifampicin at 30°C for 3 hours, and then the cells were fixed in 70% ethanol for flow cytometry assays. The flow cytometry result was used to evaluate the synchronization and determine the time required for 50% of the cells to initiate DNA replication. Before analysis by flow cytometry, the fixed cells were spun down,
resuspend in TMS buffer (Tris·HCl, MgCl₂, NaCl) [14], and stained by 2 µM Invitrogen SYTO13 green fluorescent nucleic acid stain. The fluorescence was monitored on a Coulter XL-MCL machine at the Penn State University Cytometry Facility. The addition of rifampicin prevents transcription in those cells that had not already initiated replication, resulting in 1X DNA mass. Those cells that had initiated replication precious to rifampicin addition have 2X DNA mass. The 50% initiated time is the time required for half of the cells in the synchrony to possess 2X DNA mass.

Quantitative real time RT-PCR

Synchronized C. crescentus culture was collected at each time point throughout the entire cell cycle. RNA was extracted using Qiagen RNeasy mini kit and treated with RNase free DNase I (Qiagen) twice on the column. RNA samples were diluted to 200 ng/µl and tested by PCR to exclude genomic DNA contamination before use in RT-PCR. The ABI high capacity RT kit was used for reverse transcription. cDNA was added to the PCR reaction containing a Taqman 2X universal mix and amplified with the following protocol in triplicate: 50ºC for 2 min, 95ºC for 10 min, 40 repeats of 95ºC for 15 sec and 60ºC for 1 min.

All primers and Taqman probes were designed by the software provided by ABI. The primers for dnaA gene were: forward primer 5’-GAGTTCGCGACGCTGTAG-3’ and reverse primer 5’-CGTACGGGCCGTGGAA-3’. The Taqman probe for dnaA gene was 5’-CGGACGTCACCTCAATCCTGTGCT-3’. The 16S rRNA gene was used as endogenous control. It was amplified by forward primer 5’-GGGTTAAGTCCCACGAACGA-3’ and reverse primer 5’-ATGATTAGAGTGCCCAGCCAAA-3’. The Taqman probe for 16S rRNA was 5’-CGCAACCTCGTGATTTGCCATC-3’. Both Taqman probes were synthesized by Applied Biosystems, with 6-carboxyfluorescein (6-FAM) labeled at the 5’-end as the reporter,
and 6-carboxy-tetramethyl-rhodamine (TAMRA) labeled at the 3’ end as the quencher. ABI 7300 Sequence Detection System was used to record data.

The data were analyzed by the comparative $C_T$ method for relative quantification. The threshold was set up when the fluorescent intensity of reaction was above background. The cycle threshold ($C_T$) was defined as the fraction of the cycle when the fluorescent intensity reached the threshold. For each time point, the average of $dnaA$ $C_T$ was compared to the average $C_T$ of 16S rRNA to get $\Delta C_T$, and $\Delta C_T$ of this time point was then compared to the $\Delta C_T$ of time 0 min to get $\Delta\Delta C_T$. The relative amount of $dnaA$ mRNA was calculated by formula $2^{-\Delta\Delta C_T}$.

**Pulse labeling, pulse chase and immunoprecipitation**

For pulse labeling experiments, *C. crescentus* strains were grown in M2G minimal medium with antibiotics to an OD$_{660}$ of 0.3-0.4. Swarmer cells were isolated by synchronization, released to fresh media and monitored throughout the whole cell cycle. At each time point, 1 ml of culture was sampled and mixed with 1 µl [$^{35}$S]-methionine (10-15 µCi), and grown in a 30°C shaker for 5 min. The reaction was stopped by adding 950 µl labeled culture to an Eppendorf tubes containing 50 µl TCA and vortexing. The tubes were then stored at -80°C.

For pulse chase experiments, *C. crescentus* strains were grown in M2G minimal medium with antibiotics to an OD$_{660}$ of 0.2. Protein was labeled with 1 µl [$^{35}$S]-methionine per ml culture at 30°C for 5 min. 2% cold methionine was added to the culture for the chase at time 0 min. The culture was grown at 30°C for 1-2 hrs. Every 10 min during the growth, 950 µl culture was sampled and mixed with 50 µl TCA, and stored at -80°C.

For immunoprecipitation, the TCA-protein sample was first pelleted by centrifugation. After the liquid containing TCA was thoroughly removed, 50 µl IP-SDS buffer (10 mM
Tris·HCl, pH=8, 1% SDS, 1 mM EDTA) was added to each tube. The protein pellet was resuspended by alternating vortexing and boiling at 95°C. Then 750 μl RIPA buffer was added, and the number of counts in each sample was determined by scintillation counting. The same number of counts for each sample was aliquoted to a new tube, and the volume was brought up to 500 μl by adding RIPA buffer when needed. Samples were then rocked at 4°C overnight with 1 μl of a specific antibody and 15 μl protein A conjugated with sephorase bead.

Samples were washed twice with 900 μl RIPA buffer the next day, resuspended in 2X Laemmlli loading buffer, and separated on 6% SDS-PAGE gel. The gel was dried at 80°C for 2 hrs, and then exposed to the phosphor screen overnight. The image was scanned by Typhoon and analyzed by Imagequant 5.2 software (Molecular Dynamics).

**β-galactosidase activity assay**

Log phase culture of *C. crescentus* strains harboring a variant *dnaA* promoter and *lacZ* reporter was sampled every 30 min for at least 4 time points, and β-galactosidase activity assay was performed using a modified Miller assay. Briefly, at each time point, 1 ml culture was taken and OD_{660} was measured. Then 50 μl culture was lysed by 50 μl chloroform, and mixed with 750 μl Z buffer and 200 μl 4 mg/ml ONPG. The reaction was incubated at 30°C until the color changed to yellow. Then 500 μl Na_{2}CO_{3} was added to stop the reaction, and the reaction time was recorded in minute. After centrifuging for 3 min at maximum speed, the supernatant was taken and the OD_{420} was measured. The unit was calculated by plotting OD_{420}/reaction time (Y-axis) vs. OD_{660} (X-axis). The β-galactosidase assay in *E. coli* differed in that cells were grown at 37°C, and then 100 μl culture was lysed with 25μl 0.01% SDS and 50 μl chloroform, and 500 μl Z buffer was used for assay. All strains were then compared to the control for normalization.
**Immunoblotting**

The cells from 500 μl culture were pelleted by centrifugation, and resuspended in 2X Laemmli loading buffer. For synchronized culture, the volume of loading buffer was adjusted according to the OD$_{660}$ of culture, so that samples of different time points had the same concentration of total protein. Then the same volume of each sample was loaded onto SDS-PAGE gel. After electrophoresis, the proteins were transferred to Amersham Hybond-P PVDF transfer membrane, and incubated with 1 μl anti-DnaA antibody in 10 ml TBST buffer with 5% milk at 4ºC overnight. The membrane was then washed twice with TBST buffer, and incubated with 1:1000 alkaline phosphatase labeled anti-rabbit IgG (Vector Laboratories) in TBST buffer for 1 hr at room temperature. Amersham ECF substrate was used to detect the protein, and the image was scanned by Typhoon and analyzed by Imagequant software. The membrane was then stained with ponceau S for the loading control.

**Protein purification and mass spectrometry**

The His6-tagged protein was purified by Ni-NTA affinity chromatography under denaturing condition. *C. crescentus* cells were grown in PYE broth at 30ºC till OD$_{660}$=0.2. Xylose was added to a final concentration of 0.3%. The culture was grown for additional 3 hrs, and the cells were collected by centrifugation. Cell pellets were resuspended with 5 ml buffer B (100 mM NaH$_2$PO$_4$, 10 mM Tris-HCl, 8 M urea, pH=8) per gram wet weight, and sonicated 6 passes with 30 sec on and 1.5 min off. The lysate was spun at 15000 rpm for 10 min and the supernatant was transferred to a new tube. A volume of 50% Ni-NTA slurry (Qiagen) equals to 1/50 the volume of the supernatant was added and the mixture was rocked at room temperature for 1hr. The mixture was then loaded onto Biorad Econo column, washed by buffer C (100 mM NaH$_2$PO$_4$, 10 mM
Tris-HCl, 8 M urea, pH=6.3) until OD$_{280}$ < 0.1, and eluted with buffer D (100 mM NaH$_2$PO$_4$, 10 mM Tris-HCl, 8 M urea, pH=5.9) followed by buffer E (100 mM NaH$_2$PO$_4$, 10 mM Tris-HCl, 8 M urea, pH=4.5). All fractions were checked by SDS-PAGE. Fractions containing protein were combined and concentrated by Millipore centricon centrifugal filter unit.

The concentrated protein was loaded onto SDS-PAGE gel and stained by Coomassie blue. The gel piece containing the protein was excised and destained by 25 mM NH$_4$HCO$_3$ in 50% acetonitrile, and dried in speed vacuum. To reduce disulfide bonds, 10 mM dithiothreitol (DTT) in 100 mM NH$_4$HCO$_3$ was added in addition to iodoacetamide for alkylation. The gel piece was then digested by trypsin or Asp (N) in NH$_4$HCO$_3$ at 37°C overnight. The digested peptides were prepared by Ziptip (Millipore), and analyzed by MALDI-TOF at the Penn State Proteomics and Mass Spectrometry Core Facility, or by LC-MS in Stanford Protein and Nucleic Acid Facility. Peptide fingerprint searches were performed using the MS-Digest tool from UCSF Mass Spectrometry Facility.
References


67. Ely, B. and R.C. Johnson, Generalized Transduction in CAULOBACTER CRESSENTUS.


Chapter 3

Discovery of Antibacterial Cyclic Peptides that Inhibit the ClpXP Protease

Publication:

ABSTRACT

A method to rapidly screen libraries of cyclic peptides in vivo for molecules with biological activity has been developed and used to isolate cyclic peptide inhibitors of the ClpXP protease. Fluorescence Activated Cell Sorting was used in conjunction with a fluorescent reporter to isolate cyclic peptides that inhibit the proteolysis of tmRNA-tagged proteins in E. coli. Inhibitors shared little sequence similarity and interfered with unexpected steps in the ClpXP mechanism in vitro. One cyclic peptide, IXP1, inhibited the degradation of unrelated ClpXP substrates and has bactericidal activity when added to growing cultures of Caulobacter crescentus, a model organism that requires ClpXP activity for viability. The screen used here could be adapted to identify cyclic peptide inhibitors of any enzyme that can be expressed in E. coli in conjunction with a fluorescent reporter.
INTRODUCTION

The discovery of antibiotics and pharmacological reagents requires identification of small molecules that can act in vivo. The use of in vivo screening methods to identify inhibitors of specific enzymes ensures that inhibitors will have biological as well as biochemical activity. To establish methods for rapidly screening a pool of possible inhibitors that will be effective in bacteria, inhibitors were sought from a library of cyclic peptides produced in bacteria using Split Intein Circular Ligation of Proteins and Peptides (SICLOPPS) technology [1]. SICLOPPS employs the chemistry of inteins, naturally occurring protein-splicing sequences, to ligate the amino and carboxy termini of a peptide (Fig. 3-1A). The intein-catalyzed reaction is spontaneous in vivo and does not require extensive sequence conservation in the cyclic peptide, so a wide variety of cyclic peptides and proteins can be efficiently produced. Because the cyclic peptides produced by SICLOPPS are genetically encoded, complex libraries can be produced by randomizing the DNA sequence encoding the peptide. Such libraries have been successfully screened for inhibitors of protein-protein interactions [2].

The tmRNA protein tagging and degradation pathway was chosen as a target for inhibition because this pathway is found in all bacteria, and is essential for virulence in several pathogenic species, including species of Neisseria, Salmonella, and Yersinia [3-5]. tmRNA is a specialized RNA that can enter a ribosome and add a peptide tag to the C terminus of the nascent protein [6]. The tmRNA-encoded peptide tag contains epitopes for several intracellular proteases and most tagged proteins are rapidly degraded. In E. coli, most cytoplasmic proteins tagged by tmRNA are recognized by a proteolytic specificity factor, SspB, which facilitates degradation by the ClpXP protease [7, 8]. ClpXP is a multisubunit protease that degrades a variety of substrates in addition to tmRNA-tagged proteins [9]. Although there are no known inhibitors that are specific for ClpXP, activators of ClpP activity have antibacterial activity against Gram-positive
bacteria [10]. To determine if inhibitors of ClpXP would have antibacterial activity against species that require this protease, a bacterial strain in which ClpXP is not essential was used to screen for cyclic peptides that block degradation of tmRNA-tagged proteins. Synthetic versions of these inhibitors were then tested for bactericidal activity against *Caulobacter crescentus*, a Gram-negative bacterium in which *clpX* and *clpP* are essential.
RESULTS

Screen for cyclic peptide inhibitors of ClpXP

To identify inhibitors of proteolysis of tmRNA-tagged proteins, a reporter was engineered by encoding the tmRNA peptide tag at the 3’ end of the egfp gene, such that expression of this gene produces a variant of GFP containing the tmRNA peptide tag (GFP-tag) (Fig. 3-1B). When GFP-tag was produced in wild-type E. coli, the cells showed little fluorescence (Fig. 1C), presumably because the protein was recognized as a tmRNA-tagged protein and rapidly degraded by ClpXP. Degradation required both the tmRNA peptide tag and ClpX. When a variant of GFP with no tag or with a tag lacking the ClpXP recognition sequence at the C terminus (GFP-tagDD) was produced, the E. coli were highly fluorescent (not shown). Likewise, in an E. coli strain deleted for clpX, production of GFP-tag resulted in cells that were highly fluorescent (Fig. 3-1C). These results suggested that cells producing GFP-tag would be fluorescent if an inhibitor of ClpXP was present.

A library of SICLOPPS plasmids was constructed that encodes the sequence SGW followed by five NN(G/C) codons. This SGWX5 library theoretically produces 3.2 X 10^6 different cyclic peptides. The SGW sequence allows efficient circular ligation, and the redundant codons can encode any of the 20 amino acids [11]. The use of NN(G/C) instead of fully redundant codons reduces the probability of a stop codon and results in a more even distribution of encoded amino acids.

To isolate cyclic peptides that inhibit proteolysis of tmRNA-tagged proteins, the SGWX5 SICLOPPS library was expressed in E. coli containing GFP-tag, and fluorescent cells were selected from a population of ~10^6 using FACS. Most cells producing a cyclic peptide had little fluorescence, indicating that most cyclic peptides do not inhibit ClpXP. Approximately 0.014%
of the population had fluorescence over the background level, and 96 of these cells were isolated for clonal growth and characterization. To eliminate any clones that resulted from sorting errors or spurious accumulation of GFP, cells from each colony were cultured and examined by epifluorescence microscopy. All selected clones produced some fluorescent cells (cells with fluorescence intensity at least 0.5-fold the level observed in $\Delta clpX$ cells producing GFP-tag), and two clones, containing the peptides IXP1 and IXP2, produced cells with fluorescence indistinguishable from the $\Delta clpX$ strain (Fig. 3-1C and Table 3-1).

To determine if other libraries of cyclic peptides also contained inhibitors of GFP-tag degradation, a SICLOPPS library of 9-mer peptides with the sequence SGX5PL was engineered and screened in the same manner as the SGWX5 library. Three clones (IXP3, IXP4, and IXP5) producing GFP fluorescence of similar intensity to the $\Delta clpX$ strain were isolated (Table 3-1).

Cultures producing IXP1, IXP3, or IXP4 contained over 70% fluorescent cells, indicating efficient inhibition of GFP-tag degradation (Table 3-1). In addition, the $\Delta clpX$ strain has a partially penetrant filamentous phenotype, and cells producing IXP1, IXP3, or IXP4 had a similar morphology (Fig. 3-1C), suggesting that the presence of these peptides mimics a genetic deletion of $clpX$. Fewer than 40% of the cells producing IXP2 or IXP5 were fluorescent, suggesting low intracellular concentrations of the cyclic peptide or inefficient inhibition of GFP-tag degradation in these clones. Although there is some sequence similarity between pairs of inhibitory cyclic peptides, there is no sequence conservation in the randomized region found in all of the peptides, indicating that they may inhibit the degradation of GFP-tag through different interactions.
Figure 3-1: Selection for cyclic peptide inhibitors of ClpXP.

(A) Schematic diagram of SICLOPPS library construction. Three fixed and five randomized codons were cloned in the coding region between the intein IC and IN domains. When this gene is expressed, the IC and IN domains promote circular ligation of the intervening peptide, resulting in 8-mer cyclic peptides with 5 random amino acids.

(B) Schematic diagram of the GFP-tag reporter. The tmRNA peptide tag sequence was encoded at the 3’ end of the gfp gene under control of an IPTG-inducible promoter. All GFP produced from this gene will have the tmRNA peptide tag at the C terminus.

(C) Result of expression of the GFP-tag reporter and inhibitory cyclic peptides in E. coli. Production of GFP-tag was induced in wild-type E. coli (wt), a strain lacking the clpX gene (ΔclpX), and a strain that was also producing the IXP1 cyclic peptide, and the cells were imaged by immunofluorescence (fluor.) to see fluorescent cells and differential interference contrast microscopy (DIC) to see all cells. In wild type, SspB binds to the tmRNA peptide at the C terminus of GFP-tag and tethers the protein to the ClpXP protease, resulting in rapid degradation and no fluorescent cells. In the ΔclpX strain, the absence of active ClpXP protease results in stabilization of GFP-tag and highly fluorescent cells. In wild-type cells producing IXP1, the cyclic peptide inhibits degradation of GFP-tag, resulting in fluorescent cells. DIC images show that the ΔclpX cells and wild-type cells producing IXP1 are also slightly filamentous.
Table 3-1: Cyclic peptides identified from in vivo screen.

<table>
<thead>
<tr>
<th>library</th>
<th>name</th>
<th>sequence</th>
<th>% fluorescent cells$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGWX₅ library</td>
<td>IXP1</td>
<td>SGWYGRRH</td>
<td>&gt; 80</td>
</tr>
<tr>
<td></td>
<td>IXP2</td>
<td>SGWHRGGM</td>
<td>30-40</td>
</tr>
<tr>
<td></td>
<td>Con62</td>
<td>SGWPKW</td>
<td>0</td>
</tr>
<tr>
<td>SGX₅PL library</td>
<td>IXP3</td>
<td>SGSKGVLPL</td>
<td>70-80</td>
</tr>
<tr>
<td></td>
<td>IXP4</td>
<td>SGWRVQGPL</td>
<td>70-80</td>
</tr>
<tr>
<td></td>
<td>IXP5</td>
<td>SGGRGRPL</td>
<td>10-20</td>
</tr>
</tbody>
</table>

$^a$Percentage of cells producing the indicated peptide from a SICLOPPS plasmid in conjunction with GFP-tag that have fluorescence intensity at least 0.5X that observed in ΔclpX cells producing GFP-tag, as determined by epifluorescence microscopy.

Inhibition of ClpXP in vitro

To ensure that the selected cyclic peptides are inhibitors of ClpXP and do not cause accumulation of GFP-tag in vivo by some other mechanism, cyclic peptides were synthesized and purified to examine their effects on ClpXP activity in vitro. Proteolysis of GFP-tag in the presence of ClpXP and SspB was monitored by loss of GFP fluorescence in a continuous fluorometric assay. In the absence of cyclic peptide, GFP-tag was degraded with kinetic parameters $k_{cat} = 1.79 \pm 0.08 \text{ min}^{-1}$, $K_M = 0.74 \pm 0.04 \mu\text{M}$, similar to previously published values [8]. No degradation was observed for GFP without a tmRNA tag or for GFP-tagDD when incubated with ClpXP and SspB (not shown). Likewise, no degradation was observed when ClpX or ClpP were omitted from the reaction (not shown). These results confirm that proteolysis of GFP-tag in vitro requires ClpXP recognition of the tmRNA peptide tag.
Inclusion of purified IXP1 reduced the rate of GFP-tag proteolysis, demonstrating that this cyclic peptide is a *bona fide* inhibitor of ClpXP (Fig. 3-2). Increasing the concentration of IXP1 decreased both the apparent $K_M$ and the apparent $k_{cat}$ of the reaction, suggesting uncompetitive inhibition. Fitting the data to an uncompetitive model gave a $K_I$ value of $136 \pm 35 \mu M$ (Fig. 3-2).

**Figure 3-2**: Cyclic IXP1 inhibits ClpXP *in vitro*.

GFP-tag was incubated with ClpXP and proteolysis was monitored using a continuous fluorometric assay. Representative assays without inhibitor and with IXP1 are shown. The assays were repeated using different concentrations of substrate to determine the apparent kinetic parameters. Eadie-Hofstee plots (inset) for proteolysis with no inhibitor (solid line), 50 $\mu M$ IXP1 (long dashes), and 100 $\mu M$ IXP1 (short dashes), which are consistent with an uncompetitive inhibition model.

To exclude the possibility that IXP1 is a substrate for ClpXP, IXP1 was incubated with ClpXP in the absence of GFP-tag and the amount of cyclic peptide was quantified by reverse-phase HPLC. The amount of cyclic peptide did not change over at least one hour, and no linear
peptide or smaller peptide products could be detected (Fig. 3-3). These results indicate that IXP1 is not degraded by ClpXP.

Figure 3-3: Interaction of IXP1 with ClpX and ClpP in vitro.

(A) IXP1 was incubated with ClpXP for 60 min and samples before (0 min) and after (60 min) incubation were analyzed by reverse-phase HPLC. Plots of the absorbance at 280 nm versus time after injection are shown with arrows indicating the retention time for cyclic IXP1 and linear IXP1 as determined from control assays without ClpXP. The area under the cyclic peptide peaks was unchanged after 60 min.

(B) The effects of IXP1 on the ATPase activity of ClpX with and without GFP-tag, and on the peptidase activity of ClpP, were measured. Each assay was normalized to the activity in the absence of IXP1. Error bars indicate the standard deviation at each IXP1 concentration.
The proteolytic mechanism of ClpXP involves both ATP-dependent unfolding of the substrate by ClpX and hydrolysis of peptide bonds by ClpP [12], so the effects of IXP1 on these individual reactions were investigated. The rate of ClpX ATP hydrolysis was not inhibited by IXP1 at concentrations up to 200 μM (Fig. 3-3). When GFP-tag was included in the ATPase assay the rate of hydrolysis increased by 1.6-fold, similar to previous reports [13], but IXP1 still had no effect on ClpX activity (Fig. 3-3). The peptidase activity of ClpP was assayed using the fluorogenic substrate Suc-Leu-Tyr-AMC [14], and addition of IXP1 decreased ClpP activity by less than 2% (Fig. 3-3). These results indicate that the individual activities of ClpXP are not affected by IXP1, and are consistent with an uncompetitive mechanism of inhibition.

Uncompetitive inhibition is characteristic of molecules that bind the enzyme-substrate complex, but not free enzyme. In the *in vitro* and *in vivo* proteolysis reactions above, the proteolytic adaptor SspB binds GFP-tag and tethers it to ClpXP [8]. In principle, IXP1 could act on the GFP-tag•ClpXP interaction or the SspB•ClpXP interaction. Because ClpXP can degrade GFP-tag in the absence of SspB, albeit at a slower rate [8], the proteolysis assays were repeated without addition of SspB. The degradation of GFP-tag by ClpXP was still inhibited by IXP1 in the absence of SspB (not shown), suggesting that IXP1 binds the substrate-ClpXP complex.

Because ClpXP recognizes substrates by at least 5 different motifs, it has been proposed that there are several substrate binding sites on the protease [9]. To determine if IXP1 inhibits proteolysis of ClpXP substrates recognized by an epitope distinct from the tmRNA tag, λ O protein was used as an assay substrate. Sequences at the N terminus of λ O are recognized by ClpXP, and there is no interaction between λ O and SspB [8, 15]. Degradation of λ O was assayed in the presence and absence of IXP1 by following the loss of intact λ O protein on SDS-polyacrylamide gels (Fig. 3-4). With no inhibitor, λ O was degraded with a half-life of 35 ± 2 min. Addition of 100 μM IXP1 increased the half-life to 73 ± 8 min, close to the value expected if
the $K_i$ with $\lambda$ O was the same as for GFP-tag. Therefore, IXP1 is a general inhibitor of ClpXP and affects degradation of substrates in addition to those tagged by tmRNA.

Figure 3-4: Cyclic IXP1 inhibits degradation of $\lambda$ O by ClpXP.

$\lambda$ O protein was incubated with ClpXP in the presence or absence of IXP1 and the loss of intact substrate was monitored by SDS-PAGE. Representative SDS-polyacrylamide gels stained with Coomassie blue showing the amount of $\lambda$ O protein at various times after addition of ClpXP are shown. The amount of $\lambda$ O protein remaining was plotted versus time and fit with a single exponential function to determine the substrate half-life. The average half-life for degradation of $\lambda$ O was $35 \pm 2$ min in the absence of IXP1, and $73 \pm 8$ min in the presence of 100 mM IXP1.

Purified IXP3 and IXP4 also inhibited ClpXP in vitro, but appeared to be competitive inhibitors of GFP-tag degradation (not shown). IXP2 and IXP5 did not inhibit the reaction at the concentrations tested, consistent with the observation that fewer cells producing these peptides
have high GFP-tag levels in vivo. The linear versions of IXP1, IXP2, IXP3, and IXP4 showed little inhibition of ClpXP in vitro at concentrations up to 1 mM, so the cyclic architecture of the peptides is important for inhibition (Table 3-2). An arbitrary cyclic peptide from the SGWX₅ library, Con62, also showed no detectable inhibition of ClpXP in vivo or in vitro at concentrations up to 1 mM (Tables 3-1 & 3-2). Therefore, both the sequence of the cyclic peptide and the cyclic architecture are important for inhibitory activity.

A linear peptide, XB, containing the 11 C-terminal residues of E. coli SspB, has been shown to bind to ClpX and inhibit the degradation of GFP-tag in the presence of SspB [16]. IXP5 has some sequence similarity to the C terminus of SspB, suggesting that cyclic versions of the XB peptide might also be potent inhibitors. To test this hypothesis, linear and cyclic versions of XB were synthesized and assayed in vitro (Table 3-2). As previously reported [16], the linear XB peptide was a competitive inhibitor of ClpXP degradation of GFP-tag in the presence of SspB, and as expected, the linear XB peptide had no effect on GFP-tag degradation in the absence of SspB and did not inhibit the degradation of λ O by ClpXP (not shown). Cyclic XB inhibited ClpXP proteolysis of GFP-tag in the presence of SspB with a \( K_i = 8 \pm 1 \) μM, 7-fold lower than the linear XB peptide. Like linear XB, cyclic XB did not inhibit the degradation of λ O (not shown). Therefore, the circular ligation of the XB peptide increases the efficiency of inhibition, perhaps by decreasing the entropy of the free peptide, thereby increasing the energy of binding to ClpX.

**Bactericidal activity of ClpXP inhibitors**

ClpXP is essential in C. crescentus [17], so the effect of adding purified inhibitory peptides to growing cultures was examined (Table 3-2). IXP1 killed C. crescentus with a minimum bactericidal concentration (MBC) of 279 ± 23 μM and a minimum inhibitory
concentration (MIC) of 219 ± 42 μM, suggesting that IXP1 can both enter *C. crescentus* cells and inhibit *C. crescentus* ClpXP. The linear XB peptide had a MBC of 146 ± 11 μM and a MIC of 139 ± 44 μM, and the cyclic XB peptide was more effective, with a MBC 40 ± 6 μM and a MIC 29 ± 2 μM. It is important to note that although *C. crescentus* has an SspB protein that performs the same functions as *E. coli* SspB, the sequence that interacts with ClpX is highly diverged [18]. Nonetheless, all residues of *E. coli* ClpX that make hydrophobic or hydrogen-bonding contacts with the XB peptide [19] are conserved in *C. crescentus* ClpX, so the XB peptide might bind *C. crescentus* ClpX in the same manner as for *E. coli* ClpX. Cyclic peptides are not generally toxic to *C. crescentus*, because Con62 had no effect on bacterial growth. Thus, despite anticipated problems with transporting a peptide across the membrane of Gram-negative bacteria, peptides isolated from the screen had bactericidal activity.

Table **3-2**: *In vivo* and *in vitro* properties of cyclic peptides.

<table>
<thead>
<tr>
<th>peptide</th>
<th>sequence</th>
<th>$K_I$ (μM)$^a$</th>
<th>MIC (μM)$^b$</th>
<th>MBC (μM)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IXP1</td>
<td>SGWYGRRH</td>
<td>136 ± 35</td>
<td>219 ± 42</td>
<td>279 ± 23</td>
</tr>
<tr>
<td>linear IXP1</td>
<td>SGWYGRRH</td>
<td>&gt; 1000</td>
<td>385 ± 28</td>
<td>481 ± 38</td>
</tr>
<tr>
<td>IXP2</td>
<td>SGWHRGGM</td>
<td>&gt; 400</td>
<td>304 ± 27</td>
<td>376 ± 37</td>
</tr>
<tr>
<td>linear IXP2</td>
<td>SGWHRGGM</td>
<td>&gt; 1000</td>
<td>430 ± 40</td>
<td>&gt; 600</td>
</tr>
<tr>
<td>Con62</td>
<td>SGWPYKWM</td>
<td>&gt; 1000</td>
<td>&gt; 600</td>
<td>&gt; 600</td>
</tr>
<tr>
<td>linear XB</td>
<td>CRGGRPALRVVK</td>
<td>56 ± 9</td>
<td>139 ± 44</td>
<td>146 ± 11</td>
</tr>
<tr>
<td>cyclic XB</td>
<td>RGGRPALRVVK</td>
<td>8 ± 1</td>
<td>29 ± 2</td>
<td>40 ± 6</td>
</tr>
</tbody>
</table>

$^a$Determined from kinetic assays with purified peptides and *E. coli* proteins.

$^b$Values for addition of purified peptides to *C. crescentus* cells.
DISCUSSION

Using a high throughput screen and SICLOPPS technology, efficient inhibitors of the degradation of tmRNA-tagged proteins were isolated. Kinetic assays and several lines of evidence suggest that IXP1 is an uncompetitive inhibitor of ClpXP. IXP1 is not a substrate for ClpXP, but inhibits proteolysis of at least two ClpXP substrates that are recognized by different epitopes, and inhibition is independent of SspB, suggesting that IXP1 does not compete for binding to ClpX. IXP1 does not inhibit the ATPase activity of ClpX or the peptidase activity of ClpP, consistent with a mechanism in which IXP1 binds to the ClpXP-substrate complex. One step in the proteolytic mechanism where IXP1 could act uncompetitively is the translocation of the substrate through the central pore of ClpXP. Further structural and biochemical experiments will be required to understand exactly how IXP1 inhibits ClpXP, but it clearly does not use the same mechanism as the rationally designed peptide XB. Therefore, the screen described here can identify multiple inhibitors with diverse sequences and unexpected mechanisms of action. In principle, this screening technique could be employed to identify cyclic peptide inhibitors for any pathway with a fluorescent reporter that can be expressed in *E. coli*. The method provides a set of lead compounds for reagent design or antibiotic development that includes diverse activities, and does not require any knowledge of molecular structures or co-factor requirements of the targeted pathway.

The cyclic architecture of the selected peptides was important for the inhibitory and bactericidal activities. Because libraries of linear peptides have not been screened, it is possible that there are linear peptides that would inhibit ClpXP, but each of the selected cyclic peptide sequences was less effective in a linear form. Even the XB peptide, which inhibits ClpXP by binding to the same site as the C-terminal tail of SspB, is more active as a cyclic peptide. The higher activity of cyclic peptides compared to linear versions could be the result of specific
structural features, or of tighter binding of cyclic peptides due to decreased loss of entropy. In either case, cyclic peptides are likely to be more stable in vivo than linear peptides, and are therefore more attractive for pharmacological and antibacterial agents.

Although the selected cyclic peptides are bactericidal, optimization of the sequence and length of the cyclic peptides might improve their bioactivity. In principle, further improvements could be made through modification or derivatization of the peptide, or the use of non-standard amino acids. Finally, because the selected peptides appear to inhibit ClpXP through different interactions, using them in combination could have synergistic effects on the efficiency of inhibition. Even without improvements in efficiency, biologically active inhibitors such as IXP1 provide the ability to study the role of specific pathways in vivo without the drawbacks associated with the genetic deletion or depletion of essential activities.
MATERIALS AND METHODS

Plasmids and bacterial strains

A GFP-based reporter for the proteolysis of tmRNA-tagged proteins was constructed by amplifying the egfp gene from pEGFP-N2 (BD Biosciences Clontech) using PCR with primers that add the codons for the tmRNA tag (AANDENYALAA) at the 3’ end of the gene before the stop codon, and cloning the product into pTrc99a. A similar strategy was employed for the control reporters containing egfp with no tag and egfp with the DD tag (AANDENYALDD). For fluorescence assays, plasmids bearing a GFP-based reporter were mobilized into E. coli strain BW7786 [20]. The ΔclpX strain was constructed from BW7786 using the Wanner method [21].

For over-production of GFP-tag, GFP, and GFP-tagDD, the genes were excised from pTrc99a, ligated into pQE8 (Qiagen) to produce an N-terminal His6-fusion under control of the T7 promoter, and mobilized into E. coli BL21(DE3) (Novagen). E.coli clpP was cloned into pQE70 (Qiagen) resulting in a C-terminally His6-tagged protein. E.coli clpX, sspB and the gene encoding λ O were cloned into pET28a (Novagen) to produce N-terminally His6-tagged proteins. All constructs were transformed into E.coli BL21(DE3). Unless otherwise noted, E. coli strains were grown at 37 °C in LB broth, with the addition of 100 μg/ml ampicillin, 30 μg/ml chloramphenicol, or 30 μg/ml kanamycin where appropriate. C. crescentus strain CB15N [22] was grown in PYE medium [23].

SICLOPPS libraries were constructed as previously described [11]. For the SGX5 library the initial PCR reaction combined degenerate oligonucleotide SGW+5 (5’-ggaattgccaatggggcagtcgccacaattcggcgtggnnsnnstgtaagtttggc-3’) and CBDRev (ggaattcaagtttcatgagctgccacaagg). For the second PCR reaction CBDRev was combined with a forward primer named zipper (ggaattgccaatggggcagtcgcc). Production of cyclic peptides from the SGX5PL library in E. coli was confirmed by butanol extraction and reversed-phase
chromatography followed by mass spectrometric analysis of the purified cyclic peptides as described [24].

Proteins and peptides

Histidine-tagged versions of ClpP, ClpX, SspB, GFP, GFP-tag, GFP-tagDD and λ O protein were purified from over-producing strains by metal-chelate chromatography followed by ion exchange chromatography, gel filtration, or both. In all cases, cells were grown at 30 °C in LB broth with the appropriate antibiotics to OD_{600} = 0.6, 1 mM IPTG was added to induce protein production for 3.5 h, and cells were harvested by centrifugation. Cell pellets were resuspended in Wash buffer (50 mM NaH2PO4 (pH 8.0), 300 mM NaCl, 20 mM imidazole), lysed by sonication, and cleared by centrifugation at 26,000 x g for 15 min. The cleared lysate was added to 0.1% v/v Ni-NTA resin (Qiagen) for 1 h, loaded into a column, and washed with 100 bed volumes of Wash buffer. Bound protein was eluted with 5 bed volumes of Wash buffer containing 500 mM imidazole, and fractions containing purified protein were identified by SDS-PAGE.

Fractions containing purified ClpX were combined, applied to a Superose 6 (GE Healthcare) gel filtration column equilibrated in buffer containing 50 mM Tris-HCl (pH 7.5), 300 mM NaCl, 2 mM DTT, and 10% glycerol, and fractions containing purified ClpX protein were identified by SDS-PAGE.

For purification of ClpP and λ O protein, fractions from metal-chelate chromatography were combined and dialyzed against buffer A1 (50 mM Tris-HCl (pH 8), 10 mM MgCl₂, 5 mM DTT, 10 mM KCl) and applied to a MonoQ HR5/5 column (GE Healthcare). The column was washed in buffer A1 and bound protein was eluted with a linear gradient from 10 to 1000 mM KCl. Fractions containing purified protein were combined and dialyzed against buffer A2 (50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 2 mM DTT, and 10% glycerol).
GFP, GFP-tag, and GFP-tagDD were purified as described for ClpP, except that fractions from the MonoQ column containing purified protein were applied to a Superdex 75 (GE Healthcare) gel filtration column equilibrated in buffer A2, and fractions containing the GFP variant were combined.

SspB was purified as described for ClpP, except that buffer A3 (50 mM MES (pH 6), 10 mM MgCl₂, 1 mM DTT, 100 mM KCl) was used in place of buffer A1. For all proteins, concentrations were determined by UV absorbance at 280 nm.

Linear peptides were synthesized by the PSU Huck Institutes of the Life Sciences Macromolecular Core Facility (Hershey, PA). Linear peptides were cyclized by incubating peptide with excess 1-ethyl-3-(3’-dimethylaminopropyl)carbodiimeide (EDC) and 1-hydroxy-7-azabenzotriazole (HOAt) in 50mL DMF. After 24 hours of incubation, an aliquot of each reaction was analyzed by RP-HPLC to confirm cyclization. Successful reactions were assumed based on increased retention time of peptide relative to the retention time of linear starting products. Reactions were evaporated and peptides partially purified by precipitation with diethyl ether. Final purification of cyclic peptides was accomplished by RP-HPLC. Mass was confirmed by use of electrospray-ionization mass spectrometry.

Screen for inhibitors of the tmRNA pathway

_E. coli_ BW7786 cells containing the GFP-tag reporter and the SICLOPPS library were grown in LB broth with 30 μg/ml chloramphenicol, 100 μg/ml ampicillin, and 0.0002% arabinose at 37°C to OD₆₀₀ = 0.3. IPTG was added to a final concentration of 1mM, and the culture was grown for 3 hours. Cells were sorted by fluorescence activated cell sorting (FACS) using a Beckman Coulter Elite cell sorter with Autoclone to isolate cells with GFP fluorescence, and selected cells were deposited on agar plates for clonal growth. Cells from each colony were
grown in liquid culture as described above and examined by epifluorescence microscopy. The fluorescence intensity and the number of cells with fluorescence above background were scored using ImagePro software (MediaCybernetics). SICLOPPS plasmid DNA was prepared from selected clones and the region encoding the cyclic peptide was sequenced. Peptide sequences were obtained from conceptual translation of the DNA sequences.

**In vitro proteolysis assays**

Proteolysis of GFP-tag was performed using a continuous fluorescence assay essentially as previously described [8]. Briefly, loss of GFP fluorescence was monitored at 507 nm after excitation at 395 nm at 30 °C using a Hitachi F-2000 Fluorescence Spectrophotometer in buffer R (25mM HEPES-KOH (pH 7.6), 5 mM MgCl₂, 50 mM KCl, 0.032% NP-40, 10% glycerol, and an ATP regeneration system containing 4 mM ATP (pH 7), 16 mM creatine phosphate and 0.32 mg/ml creatine kinase). Typically, reactions contained 0.1 μM ClpX₆, 0.3 μM ClpP₁₄, and equimolar concentrations (0.2 μM – 2.0 μM) GFP-tag and SspB. Degradation of GFP-tag protein was confirmed by SDS-PAGE assays. To examine inhibition of GFP-tag proteolysis, peptides were incubated with ClpXP in reaction buffer for 5 min prior to the addition of GFP-tag. Plots of fluorescence versus time were fit with a single exponential function to determine the initial rate of proteolysis. Kinetic parameters were estimated using Eadie-Hofstee plots. Curve fitting for competitive inhibitors was performed using the Scientist program (MicroMath Scientific Software, Inc.).

Peptidase activity against IXP1 was assayed by incubating 200 μM IXP1 with 0.1 μM ClpX₆, 0.3 μM ClpP₁₄ and the ATP regeneration system in buffer R at 37 °C. Sample were taken after 0, 5, and 60 min, and separated by reverse-phase HPLC using a Varian Microsorb-MV C18 column developed in a gradient from 0.1% trifluoroacetic acid in water to 0.1% trifluoroacetic
acid in acetonitrile. UV absorbance at 280 nm was monitored and the peak corresponding to the IXP1 cyclic peptide was determined by comparison to reactions containing only cyclic IXP1 or linear IXP1. The loss of cyclic IXP1 was determined by integrating the area under the cyclic IXP1 peak.

Proteolysis of λ O protein was assayed by incubating 1 μM λ O with 0.1 μM ClpX₆, 0.3 μM ClpP₁₄ in buffer R at 30 °C. At various times the reaction was sampled, the reaction was terminated by boiling in SDS-PAGE loading buffer, and analyzed using SDS-polyacrylamide gels stained with Coomassie blue. The intensity of the band corresponding to intact λ O protein was measured using ImageQuant software (Molecular Dynamics), and plots of the intensity versus time were fit with single exponential functions to determine the half-life of λ O protein.

**ATPase and peptidase assays**

ClpX ATPase activity was measured by monitoring the increase in phosphate using a ternary heteropolyacid assay [25]. 1 μM ClpX was incubated with or without 6 μM GFP-tag and varying concentrations of IXP1 in buffer P (4 mM ATP, 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 100 mM KCl, 10 mM MgCl₂, 2 mM DTT, and 10% glycerol) at 37 °C. At each time point 10 μl was removed from the reaction, added to 265 μl 0.88 M nitric acid for 2 min, and 225 μl color developing solution (44.4 mM bismuth nitrate, 31.1 mM ammonium molybdate, 0.11% ascorbic acid) was added and the absorbance at 700 nm was determined. The rate of ATP hydrolysis was determined from plots of phosphate accumulation versus time.

ClpP peptidase activity was measured using the fluorogenic peptide Suc-Leu-Tyr-AMC. 0.1 μM ClpP was incubated with 0.5 – 1.0 mM Suc-Leu-Tyr-AMC and varying concentrations of
IXP1 in buffer P at 37 °C and monitoring the fluorescence of AMC (excitation at 353 nm, emission at 442 nm).

**Antibacterial activity assays**

To measure the effects of peptides on bacterial growth, cultures of *C. crescentus* were diluted to OD$_{660}$=0.001, peptide was added at various concentrations, the cultures were incubated at 30°C for 14 h, and the MIC was determined as the lowest concentration of peptide that prevented growth. To determine the MBC, each culture was diluted 1:100 in PYE and 10 μl of the diluted culture was spread onto PYE agar plates, incubated at 30°C overnight, and the number of colonies on each plate was counted. The MBC was assigned as the concentration of peptide that reduced the number of colonies by 99.9% compared to cultures with no inhibitor.

**ACKNOWLEDGMENTS**

This work was supported by NIH grant GM68720 to KK and DTRA JSTO CBD grant W911NF-06-1-0144 to KK and SB.
REFERENCES


VITA

Lin Cheng

EDUCATION

2002-2008    Ph.D.    Genetics
              The Pennsylvania State University
              University Park, PA

1999-2002    M.S.    Genetics
              Shandong University School of Medicine
              Jinan, China

1994-1999    B.M.    Clinical Medicine
              (MD equivalent)    Shandong Medical University
              Jinan, China

TEACHING EXPERIENCE

Fall 2003    BIO110 (The introduction of Biology), Penn State University

AWARDS AND FELLOWSHIPS

2007    Genetics Student Travel Award, Penn State University
2005    Student Travel Award for American Society of Microbiology conference
2002    Genetics Student Fellowship, Penn State University

PUBLICATIONS

        Inhibit the ClpXP Protease    Protein Science 16, 1-8.

        Yi Chuan Xue Za Zhi. 20: 89-93.


PROFESSIONAL SOCIETIES

2005-2008    The American Society of Microbiology