2015

ELUCIDATION OF A NOVEL PATHWAY IN STAPHYLOCOCCUS AUREUS: THE ESSENTIAL SITE-SPECIFIC PROCESSING OF RIBOSOMAL PROTEIN L27

Erin A. Wall
Virginia Commonwealth University

ELUCIDATION OF A NOVEL PATHWAY IN STAPHYLOCOCCUS AUREUS: THE
ESSENTIAL SITE-SPECIFIC PROCESSING OF RIBOSOMAL PROTEIN L27

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at Virginia Commonwealth University

by

ERIN ASHLEY WALL
B.Sc., Texas A&M University, 2010

Director: Gail E. Christie, Ph.D.
Professor, Department of Microbiology and Immunology
Virginia Commonwealth University
Richmond, Virginia

Virginia Commonwealth University
Richmond, Virginia
April 2015

Dedication

To Mom - for being there for me no matter what. You're my rock. You look out for me first, above yourself. You taught me how to love and be loved, to read, to learn, to work hard, and to treat others with respect no matter who they might be. If I ever become a mother, I'll look to your example. I love you.

To Ciarán - you always encouraged me to work hard so that I could accomplish great things. You've shown me the world, and inspired in me a longing to experience all of it. Your infectious interest in a variety of subjects set an incredible example that I will continue to follow. I love you!

To Justin - for showing me just how fun life can be, and a new way for my heart to beat. I love you more than I can possibly say, and we both know how wordy I can be! You're an inspiration, a confidant, a partner in all the adventures to come and you've given me the happiest time of my life. I look forward to what life holds for us. No matter what it is, I'm with you.

To Gail - you've been all I could have wished for in a mentor and even more. You took me on and got me through this whole project and stayed supportive when I couldn't deal with my anxiety or even face it. Thank you for everything - your understanding, your expertise and for the truly massive amount time and consideration you've given me. As Neil DeGrasse Tyson said of Carl Sagan in the first episode of Cosmos 2014, you've shown me what type of scientist, and what type of person I want to be. I look forward to all the work that we will continue together.

To Darrell - there is no one like you and I would be lost without your expertise. We always have great talks and you've been much more than a resource and a mentor - you're one of the most valuable, wonderful friends I have. You've also shown me what type of scientist and what type of person I want to be.

To Aunt Sherry and Uncle Doug - for showing me a home away from home. Holidays and visits with you were completely necessary breaks that got me through graduate school.

To Candace, Alan and little Stella - your happy family is inspiring!

To John - though you're five years my junior I always seek and value your advice. Thanks for being like a brother to me.

To Pat and Majonka - thanks for your continual support! You two are a blast. Much love and gratitude.

To Paul, Kathy and Melissa - thank you for welcoming me as you have. I love you all so very much.

To Al - I don't know where I would be without our Sundays. Thanks for always being there, full of understanding, and willing to chat with me about everything from NMR physics to South Park. You're truly an inspiration to me! Much love.

To Mary and Frances, my little VCU family.

To Craig Shipp and Tim Diggs, my first science mentors after Ciarán.

To Mr. Dennis Phillips and Mr. Benjamin Griffith, high school teachers who taught me so much more than their subjects. Rest in peace.

To my extended family and friends without whom I would be lost.

Acknowledgments

The most outstanding feature of life's history is a constant domination by bacteria.

— Stephen Jay Gould

Nature uses only the longest threads to weave her patterns, so that each small piece of her fabric reveals the organization of the entire tapestry.

— Richard P. Feynman

I seldom end up where I wanted to go, but almost always end up where I need to be.

— Douglas Adams

Contents

List of Abbreviations

A	alanine
Å	Angstrom (10^{-10} meters)
a.a.	amino acid
β	beta
BHI	brain heart infusion media
bp	base pair
BSA	bovine serum albumin
C	cysteine
CDS	coding sequence or coding region
cfu	colony forming units
Chl	chloramphenicol
cm	centimeter
Da	Dalton
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
EDTA	ethylenediamine tetraacetic acid
E. coli	*Escherichia coli*
Eco	*Escherichia coli*
g	gram
G	glycine
gDNA	genomic DNA
gp	gene product
H	histidine
IPTG	isopropyl-D-1-thiogalactopyranoside
Kan	kanamycin
kbp	kilobase pairs
kDa	kiloDalton
L	liter
LB	Luria-Bertani media
M	molar
ml	milliliter
mmol	millimole
mM	millimolar
mol	mole
mRNA	messenger RNA
msec	millisecond
MW	molecular weight
MWCO	molecular weight cut-off
MQ	milli-Q purified water
n	nano
NCTC	National Collection of Type Cultures
nm	nanometer
ºC	degree Celsius
OD	optical density

orf	open reading frame
p	pico
PCR	polymerase chain reaction
PTC	peptidyl transferase center
RBS	ribosome binding site
RNA	ribonucleic acid
rRNA	ribosomal RNA
rpm	revolutions per minute
rRNA	ribosomal RNA
S. aureus	*Staphylococcus aureus*
Sau	*Staphylococcus aureus*
SDS	sodium dodecyl sulfate
Spec	spectinomycin
SUMO	Small Ubiquitin-like Moiety; has specific Ulp-1 protease cleavage site
TAE	Tris-acetate-EDTA
TE	Tris-EDTA
Tet	tetracycline
TSB	tryptic soy broth
tRNA	transfer RNA
wt/vol	weight per unit volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
µg	microgram
µl	microliter
µm	micrometer
µM	micromolar

List of Figures and Tables

Abstract

ELUCIDATION OF A NOVEL PATHWAY IN STAPHYLOCOCCUS AUREUS: THE ESSENTIAL SITE-SPECIFIC PROCESSING OF RIBOSOMAL PROTEIN L27

Erin A. Wall

Director: Gail E. Christie

Department of Microbiology and Immunology, Virginia Commonwealth University School of Medicine, PO Box 980678 Richmond, VA 23298

Ribosomal protein L27 is a component of the eubacterial large ribosomal subunit that has been shown to play a critical role in substrate stabilization during protein synthesis. This function is mediated by the L27 N-terminus, which protrudes into the peptidyl transferase center where it interacts with both A-site and P-site tRNAs as well as with 23S rRNA. We observed that L27 in *S. aureus* and other Firmicutes is encoded with a short N-terminal extension that is not present in most Gram-negative organisms, and is absent from mature ribosomes. The extension contains a conserved cleavage motif; nine N-terminal amino acids are post-translationally removed from L27 by a site-specific protease so that conserved residues important for tRNA stabilization at the peptidyl transferase center are exposed. We have identified a novel cysteine protease in *S. aureus* that performs this cleavage. This protease, which we have named Prp, is conserved in all bacteria containing the L27 N-terminal extension. L27 cleavage was shown to be essential in *S. aureus*; un-cleavable L27 did not complement an L27 deletion. Cleavage appears to play an essential regulatory role, as a variant of L27 lacking the cleavage motif could not complement. Ribosomal biology in eubacteria has largely been studied in *E. coli*; our findings indicate that there are aspects of the basic biology of the ribosome in *S. aureus* and other related bacteria that differ substantially from that of *E. coli*. This research lays the foundation for the development of new therapeutic approaches that target this novel, essential pathway.

Chapter 1

Introduction

Staphylococcus aureus:
Taxonomic classification:
Domain: Bacteria
Kingdom: Eubacteria
Phylum: Firmicutes
Class: Bacilli
Order: Bacillales
Family: Staphylococcaceae
Genus; Species: ***Staphylococcus aureus***

Staphylococcus aureus is a Gram positive coccus which is both a natural commensal of the human skin and nares and a dangerous pathogen. *S. aureus* is the most common cause of community acquired skin and soft tissue infections (Morgan, 2011) and the leading cause of hospital acquired infection (Klein, 2007). This versatile bacterium can cause osteomyelitis, septic arthritis, necrotizing pneumonia and toxic shock; severe infection is possible at nearly every location in the body. *S. aureus* is also the leading cause of infectious endocarditis in North America (Murdoch, 2009). In a study of one thousand individuals with infectious endocarditis, *S. aureus* was the most common pathogen associated with vegetations on pacemakers and native or prosthetic valves (Leone, 2012). In a separate study by the American Heart Association, *S. aureus* accounted for 60 to 80% of infections of cardiovascular implantable electronic devices.

Incidence of infection of these devices has risen faster than their increased rate of usage and ultimately requires costly and dangerous device removal (Baddour, 2010).

S. aureus has become profoundly resistant to the latest generation of antibiotics, resulting in widespread methicillin resistant *S. aureus* or MRSA. In the hospital setting, MRSA infections debilitate immunocompromised patients and those recovering from surgery or device implantation. Presence of MRSA is a predictor of arthroplasty failure; treatment of these infections demands costly resection of prosthetics for debridement or replacement, dramatically increasing hospital stay for the patient (Kurtz, 2008; Parvizi, 2009). MRSA strains are now becoming resistant to recently developed, clinically important PhLOPSA antibiotics (**Ph**enicols, **L**incosamides, **O**xazolidinones, **P**leuromutilins, and **S**treptogramin **A**). This resistance pattern was detected as early as 2006 and requires a single gene (Long, 2006). Current practices of chlorhexidine bathing of patients who are colonized with harmful skin flora has selected for chlorhexidine-resistant strains of MRSA (Batra, 2010).

The prevalence of vancomycin-intermediate or resistant *S. aureus* (VISA/VRSA) has also increased. Multiple cases of VRSA have already been identified in the United States (Lubin, 2011; CDC, 2002; Tenover, 2004). The first documented case of VRSA bacteremia occurred in August 2012 in Brazil. The vancomycin resistance genes in the causative strain were determined to reside on a novel mobile genetic element, and the resistance cassette was demonstrated to spread readily between strains of *S. aureus*. The VRSA isolate in question was also resistant to methicillin, teicoplanin, erythromycin,

clindamycin, ciprofloxacin, gentamicin and trimethoprim-sulfamethoxazole (Rossi, 2014). The treatment arsenal for MRSA and VRSA grows smaller as linezolid resistance continues to be reported among MRSA strains (Bonilla, 2010; Mendes, 2008; Sanchez-Garcia, 2010).

Linezolid, one of the last lines of defense against MRSA and VRSA, is an extremely clinically important antibiotic that seems to target the active site of the ribosome, the peptidyl transferase center. It is usually highly effective against *Staphylococcus epidermidis,* yet between 2008 and 2010 certain strains of *S. epidermidis* isolated from bloodstream infections were not only highly linezolid-resistant (unimpaired growth at 128μg/ul, where therapeutic dosing is 2μg/ul), they grew _faster_ when linezolid was present at high concentration. These strains exhibited increased peptidyl transfer activity and what appeared to be reliance on linezolid for proper ribosomal subunit assembly. Other such resistance-to-dependence phenotypes have been observed with vancomycin resistant/dependent enterococci (VRE/VDE) (Kokkori, 2014).

From the above data it has become abundantly clear that continuous antibiotic development is of primary importance to protect patients from emerging resistant bacterial pathogens. Current antibiotic therapies are already of limited utility, yet the novel drugs that are currently being developed are often "me too" drugs that work via the same mechanism as existing antimicrobials. To create unique drugs, new therapeutic targets and new data on the function of existing targets are needed.

The eubacterial ribosome has historically been a prime target for numerous antibiotics, including macrolides (e.g. erythromycin), lincosamides, oxazolidinones and tetracylines, which work by blocking aminoacyl-tRNA binding, preventing peptidyl transfer or causing premature peptidyl-tRNA release (Auerbach, 2002; Yonath, 2005; Colca, 2003; Tejedor, 1986). Though these drugs are beginning to fail, protein synthesis and ribosome formation remain extremely important yet underdeveloped targets in those bacteria that are only distantly related to the bacterial model organism *Escherichia coli*. This dissertation presents new findings regarding fundamental *S. aureus* ribosome function that are important to future strategies for antibiotic development against *Staphylococci* and related pathogens.

Bacterial ribosomes

Ribosomes, the molecular machines that carry out protein biosynthesis, are ribozymes composed of ribosomal RNA (rRNA) and proteins. Ribosomes consist of two ribonucleoprotein complex subunits, one small and one large. The subunits and assembled ribosome are designated by their sedimentation rate - the small 30S and large 50S in bacteria; a complete ribosome containing both subunits is 70S. The large and small subunit each contain part of the A, P and E sites through which transfer RNAs (tRNAs) charged with amino acids enter to add their residue to the nascent peptide chain, are deacylated, and exit, respectively (Figure 1). A "ratcheting" motion occurs due to the independent movement of the small and large subunits such that a tRNA can, for example, occupy the A site on the small subunit and the P site on the large subunit. The

hybrid states, in order, are A/A, A/P, P/P, P/E, E/E, with the site occupied on the 30S subunit noted first. In this way the movement of the two subunits, fueled by GTP, aids the passage of the tRNAs (Frank, 2000). The nascent peptide chain is continually polymerized as the C-terminal end is transferred to the P/P site tRNA after each peptide bond formation at the A/P hybrid site. The 16S rRNA forms the 30S subunit and contains the decoding center where mRNA codons are paired with the appropriate anticodons on incoming charged tRNAs. The 23S rRNA forms the major nucleic acid component of the large subunit and is complexed with the 5S rRNA that forms the central protuberance of the ribosome. The 50S subunit contains the Peptidyl Transferase Center (PTC) on Domain V of the 23S rRNA and the exit tunnel through which the nascent peptide must pass. The PTC is the location of the active site nucleotides that catalyze peptide bond formation and tRNA de-acylation. Here we will explore in detail the bacterial PTC components - rRNA and ribosomal proteins.
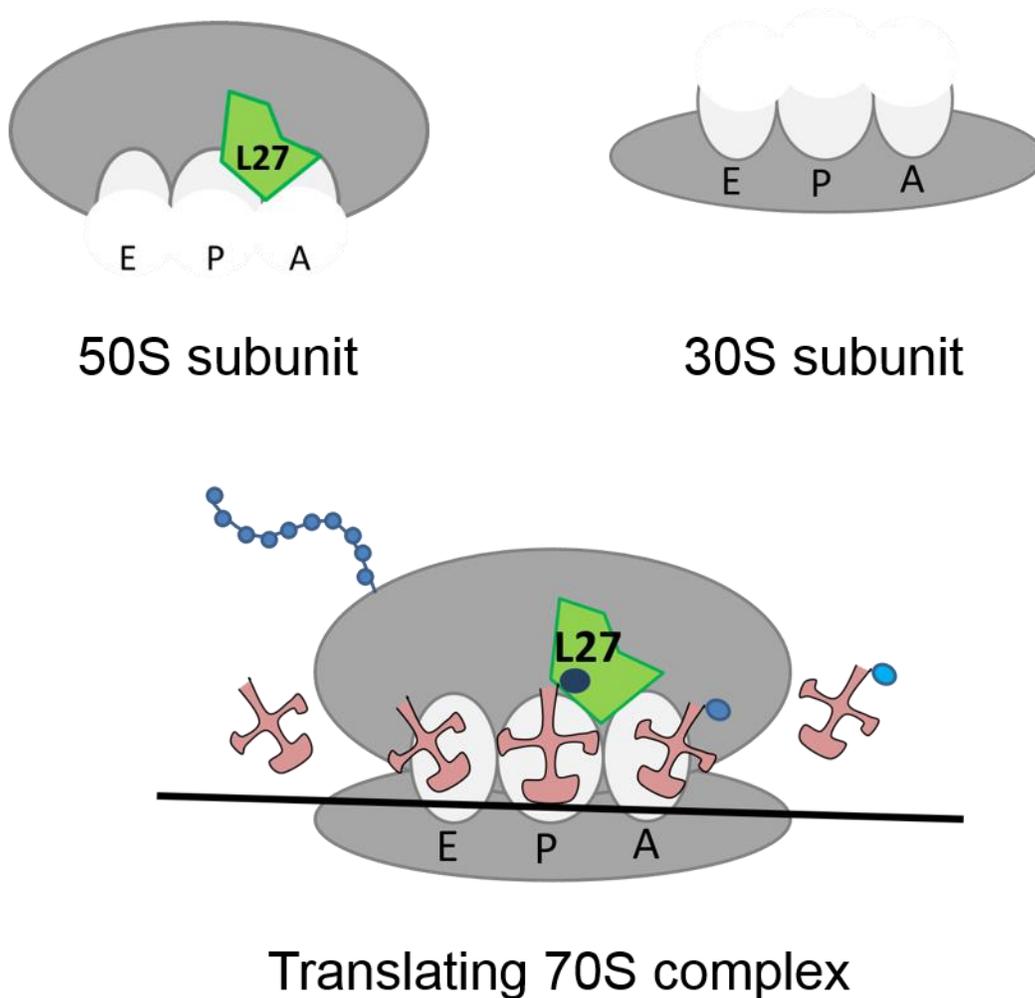
Figure 1. The translating bacterial ribosome uses ribosomal protein L27 to increase tRNA binding efficiency

Normal 50S and 30S subunits assemble independently and translate mRNA (black line) using tRNAs (pink) charged with amino acids (blue) to form a nascent peptide (blue chain). Charged tRNAs enter at the A site, undergo peptidyl transfer to accept the nascent peptide, move to the P site, release the nascent peptide to the next A site tRNA and then exit from the E site. L27 (green) is located between the A and P sites, where it stabilizes incoming tRNAs and peptidyl transfer. The nascent peptide passes through a tunnel in the 50S subunit and remains attached to a tRNA until the chain is terminated via release factor hydrolysis. In the absence of L27, the 50S subunit assembles incorrectly and translation efficiency is decreased due to impaired A site and P site tRNA binding.

The protein components of the bacterial ribosome play a large role in efficient ribosomal assembly and function. Ribosomal proteins (r-proteins) are among the most universally conserved cellular components - about half of them occur in all three domains of life, while some remain domain-specific. The study of r-proteins is complicated by nomenclature, and recent efforts have been undertaken to change the naming system to clarify homologous proteins in Archaea, Eukaryota and Bacteria (Ban, 2014). In general, if an r-protein is incorporated into the 50S/60S or "large" subunit, it will be designated Lx where x is an integer. The integer only represents the order in which the r-protein was found, it has nothing to do with size or function. 30S/40S or "small" subunit proteins are labeled Sx in the same way. A major source of confusion comes from non-homologous ribosomal proteins that have been given identical names in each domain of life, e.g. the protein known as L1 in bacteria and yeast has been called L10A in humans and bears no homology to human L1 (Ban, 2014). It is important to note this fact when exploring the distribution of ribosomal proteins in bacteria. Despite the novel nomenclature proposed by Ban *et al*, recently it has been the practice to append the letter "a," "e," or "p" to the end of the ribosomal protein designation to indicate whether the mentioned protein occurs in Archaea, Eukaryota or Prokaryotes, respectively. Sometimes instead of "p," "b" will be used instead to indicate Bacteria.

In bacteria, the peptidyl transferase center (PTC) is composed largely of rRNA, with only very few r-protein components in close proximity. Universally, the PTC is positioned within a pseudo two-fold symmetrical region which contains conserved nucleotides that catalyze protein formation in two ways. First, there are nucleotides that seem to assist in

catalysis of peptide bond formation through a proposed "proton wire," composed of coordinated water molecules, that stabilizes the nucleophilic attack by the incoming residue on the ester bond between the P-site tRNA and the nascent peptide chain (Polikanov, 2014). Second, there are nucleotides that contribute entropically through induced-fit conformational changes that stabilize and facilitate the A-site tRNA rotatory movement required for substrate-mediated acceleration (reviewed in Bashan and Yonath, 2008). The ribosome has been thought to be a ribozyme for at least the past fifteen years (Cech, 2000). This conclusion was initially drawn based on the crystal structure of the 50S subunit from Archaeon *Haloarcula marismortui* in which the PTC is composed of rRNA exclusively (Nissen, 2000). However, later high resolution structures of the bacterial ribosome (as opposed to the archaeal ribosome) demonstrated the presence of the N-terminal region of a ribosomal protein only eight angstroms from the center of the PTC that participates in the proton wire - ribosomal protein L27 (Voorhees, 2009; Polikanov 2014).

Bacterial ribosomal protein L27, product of the *rpmA* gene, is a component of the large ribosomal subunit found only in eubacteria and in the ribosomes of mitochondria and chloroplasts. This homology follows the endosymbiont theory - mitochondria and chloroplasts (plastids) are thought to have evolved from small rickettsial bacteria and cyanobacteria, respectively, which became engulfed by larger cells (Sagan, 1967). L27 is highly conserved, and although deletions can be tolerated in some bacterial species (albeit with severe growth defects), *rpmA* is generally considered an essential gene. L27 consists of a C-terminal β-sandwich domain and a long N-terminal arm that extends into

the PTC. Its rRNA contacts include domain II, helix 34 and domain V helix 81 and 86 on the 23S rRNA. It is also in contact with the 5S rRNA, positioned at the central protuberance in the midst of the ribosome (Fox, 2010; helix numbers from Yusupov, 2001). All bacteria seem to contain only a single copy of L27, despite the fact that r-protein paralogs are common, especially in the low GC phylum Firmicutes to which *S. aureus* belongs (Yutin, 2012).

The role of r-protein L27 in substrate stabilization at the PTC and its entropic contributions to catalysis have been increasingly appreciated in the last decade (Wower, 1998; Maguire, 2005; Voorhees, 2009; Polikanov, 2014). The new work on the proton wire model implicates the N-terminus of r-protein L27 (after loss of the initial N-formyl methionine) as an extremely important participant in the coordination of a water molecule that supports the proton wire, while also serving as an agent of P-site tRNA stabilization (Polikanov, 2014; Wang, 2012). It has been chemically demonstrated to play a critical role in tRNA substrate stabilization during the peptidyl transfer reaction (Wang, 2004a; Maguire, 2005). The 3' ends of both A- and P-site tRNAs in the PTC can be crosslinked to L27 in *E. coli* (Wower, 1998). In a genetic complementation assay, deletion of the first three N-terminal amino acids (A2 H3 K4; the universal fMet is removed) from *E. coli* L27 resulted in a drastically decreased growth rate, loss of tRNA crosslinking and a defect in peptidyl transferase activity (Maguire, 2005). FRET analysis of peptidyl-tRNA dynamics implicated L27 residue K4 in stabilization of the P-site tRNA (Wang, 2012; Xiao, 2012). Deletion of the entire *E. coli rpmA* gene led not only to a severe growth defect but also incomplete assembly of the 50S ribosomal subunit, indicating a role for L27 in ribosome

assembly (ribosomes lacking L27 also lost r-proteins L16, L20 and L21) as well as catalysis (Wower, 1998). The role of the L27 N-terminus in translation and thus bacterial biology is large and well-studied - any variations in the conserved region would be of considerable importance and highly worth investigating.

In this study we report a novel N-terminal extension of L27 in Firmicutes and related bacteria. For *S. aureus* to survive, this extension must be cleaved away by the newly discovered Prp protease. This is a completely unprecedented essential pathway in *S. aureus* that represents a new target for antibiotic design.

Chapter 2

Discovery of the L27 N-terminal extension in *S. aureus*

Our work on ribosomal protein L27 began in a seemingly unrelated area of research- the structural biology of bacteriophage capsid formation. In the course of studying capsid assembly in staphylococcal phage 80α, a cleavage event that occurred in the phage scaffold and major capsid proteins, gp46 and gp47, was examined. It is very common for phage to utilize a "prohead" protease to cleave their scaffold away after procapsid assembly, in order to form a mature capsid. However, 80α does not encode a protein with any predicted homology to known proteases. It was instead found that these proteins were cleaved correctly at their conserved motif upon expression in *S. aureus* in the absence of any other phage proteins, implicating a host protease in this process (see Fig. 2). Over-expression of the processed versions of both capsid and scaffold proteins in *S. aureus* led to formation of polyheads, or polymer sheets and tubes of capsid protein. This seemed to indicate that the cleavage of both capsid and scaffold proteins might have regulatory implications during capsid assembly (Spilman, 2012).
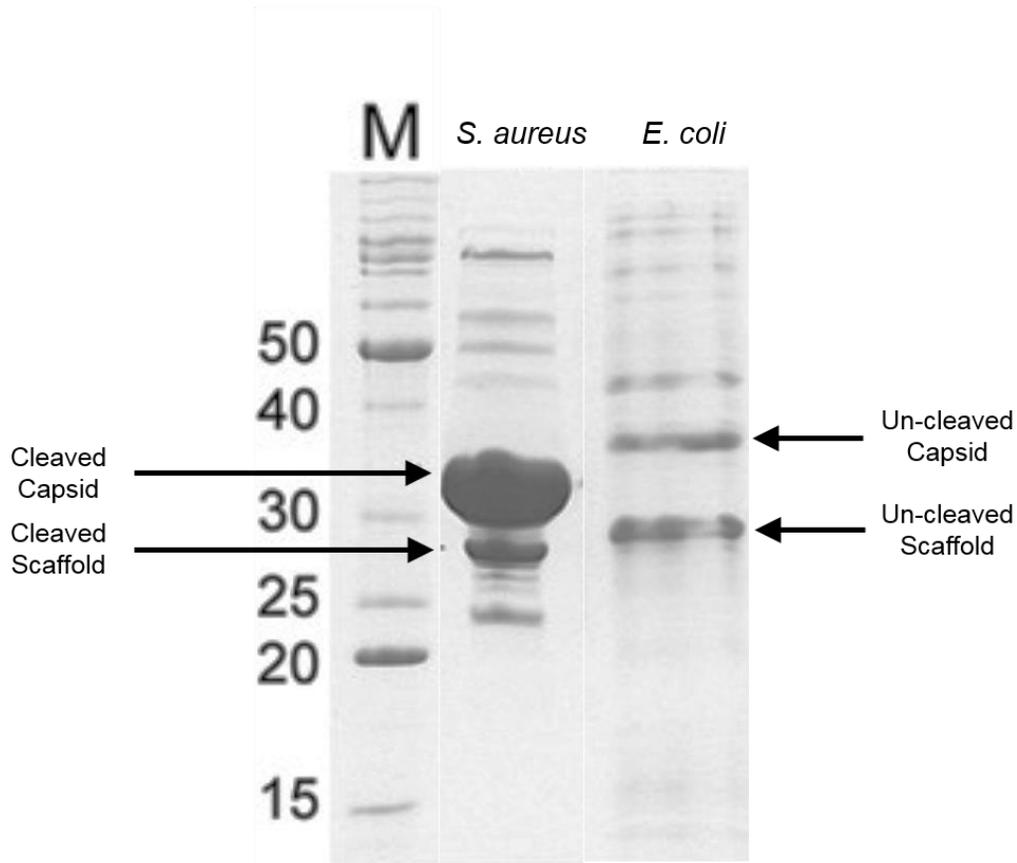
Figure 2. Staphylococcal phage 80α proteins are processed by a host protease not present in E. coli.

This is a composite gel of 80α scaffold and capsid protein expression adapted from Spilman *et al*, 2012. The lane on the left contains molecular weight markers in kDa, the *S. aureus* lane shows the protein bands that were produced when 80α scaffold and capsid were expressed in *S. aureus* in the absence of any other phage proteins. The right lane shows the results of 80α scaffold and capsid expression in *E. coli*. The bands in the right lane represent un-processed proteins, where no cleavage has occurred. The downward-shifted bands in the *S. aureus* lane have been N-terminally cleaved at a conserved motif, confirmed by mass spectrometry.

Spilman, M.S., Damle, P.K., Dearborn, A.D., Rodenburg, C.M., Chang, J.R., Wall, E.A., *et al*. (2012) Assembly of bacteriophage 80alpha capsids in a *Staphylococcus aureus* expression system. *Virology* **434:** 242–250.

Images reproduced with permission

The cleavage did not appear to be autoproteolytic, and no equivalent cleavage of either protein ever occurred upon expression in *E. coli*. This was the first demonstration of the involvement of a bacterial host protease in bacteriophage assembly. We did not believe that *S. aureus* encodes a protease for the purpose of phage maturation. It seemed likely that the phage was making use of something the bacterium already needed for cellular purposes. The function of such a protease would have to be so important that the cell would never lose the enzyme during the course of its evolution, and there would be no pressure for the phage to retain such an enzyme in its own space-limited genome.

A search for a host protein with a potential homologous cleavage site to that of gp46 and gp47 identified a similar sequence at the N-terminus of *S. aureus* ribosomal protein L27 (Spilman, 2012). This motif is conserved in all known Firmicutes, including *Bacillus, Listeria, Clostridium* and *Streptococcus.* It represents an N-terminal extension that occludes well-studied residues that have been found in the PTC. This motif is not found in bacteria such as *E. coli*. Importantly, sequence specific cleavage in the Firmicute L27 was predicted to generate an N-terminus highly similar to that found in *E. coli* L27 (see Figure 3).

Figure 3. The 80α scaffold and capsid proteins share a conserved cleavage motif with L27 in _Staphylococcus aureus_ and _Bacillus subtilis_ but not L27 in _E. coli_

ClustalW alignment of the N-termini of L27 from _E. coli_, _B. subtilis_ and _S. aureus_ with those of bacteriophage 80α scaffold and capsid proteins. Identical residues are highlighted in black; similar ones in gray. A dotted line separates the N-terminus encoded in the genome from that found in the mature protein.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in Staphylococcus aureus and related bacteria. _Molecular Microbiology, 95_(2), 258-269. doi:10.1111/mmi.12862 [doi]

Image reproduced with permission

Several lines of evidence were consistent with this postulated cleavage. Based on the structure of L27 and its position in known ribosome structures (Voorhees, 2009; Polikanov, 2014), the N-terminal extension would reach into the PTC and interfere with the peptidyl transfer reaction via massive steric hindrance. Further, the N-terminal four residues of *E. coli* L27 constitute a highly conserved A(S/H)KK motif that is also found in *S. aureus* L27, immediately following the 9-residue N-terminal extension. Antibiotics that inhibit the peptidyltransferase reaction by obstructing the PTC, such as oxazolidinone, crosslink to *S. aureus* L27 residues in the A(S/H)KK motif (Colca, 2003; Leach, 2007) and not to any residues N-terminal to those. These observations suggested that the N-terminal extension is not present in the mature, functioning ribosome.

Consistent with this expectation, no tryptic peptide corresponding to the N-terminal extension was found in ribosomes isolated from *S. aureus* (Colca, 2003). The *Bacillus subtilis rpmA* gene also encodes an L27 protein with a similar N-terminal extension, and these extra amino acids are lacking in the L27 protein isolated from *B. subtilis* ribosomes; this discrepancy had been noted previously, but attributed to misannotation of the gene (Lauber, 2009). Since L27 is positioned at the heart of the ribosomal active site, where it plays a key role in translation (Maguire, 2005), we sought to elucidate further the role of this N-terminal extension in *S. aureus* biology.

This work sets out to describe the function and distribution of L27 proteins with this N-terminal extension motif, which immediately precedes and occludes the highly conserved PTC-participating region on the N-terminus of L27 in *E. coli.* This report

demonstrates that this extension is cleaved post-translationally in *S. aureus*, prior to or concurrent with ribosome assembly. We have identified a universally conserved cysteine protease in bacteria containing the L27 N-terminal extension that performs this cleavage. Both L27 and the protease, which we have termed Prp, are essential in *S. aureus.* This indicates that the cleavage event is essential as well. The recent emergence of virulent, antibiotic resistant strains of *S. aureus* that spread among otherwise healthy individuals is a considerable public health concern that necessitates new antibiotics. Prp provides an attractive target for the development of antibiotics specific to *S. aureus* and other Firmicute or Tenericute pathogens in which this specific L27 processing occurs.

Chapter 3

Methods

Detailed step-by-step instructions for many of the techniques described here have been included in a section called Protocols.

Bioinformatics. The National Center for Biotechnology Information (NCBI) website was the source for most bioinformatics data presented in this work. The Conserved Domain Database (CDD) was where it was first noticed that *ysxB* was part of a family known only as Domain of Unknown Function (DUF) 464. The CDD and the Wellcome Trust protein family site, Pfam, list species that contain members of any specified protein domain family, and this allowed the determination of which bacteria phyla contained a DUF464 homolog. Members of these phyla were manually inspected to determine whether they contained N-terminally extended L27 using the NCBI Graphics View of each fully annotated genome.

The L27 protein alignment (Figure 3) was generated through the Mobyle Portal of the Pasteur institute using BOXSHADE v3.31 C (beta, 970507 http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::boxshade) (Neron, 2009). A MUSCLE alignment of the five proteins was generated using the CLUSTALW algorithm and default

preset values in the bioinformatics software package Geneious (Kearse, 2012). This alignment file was submitted to the Mobyle Pasteur BOXSHADE server. All other alignments were generated in Geneious using the MUSCLE algorithm.

The L27 phylogenetic tree (Figure 6) was generated in Geneious using the BLOSUM80 cost matrix. The L27 tree is derived from representative L27 protein sequences from each bacterial phylum. These sequence data were subjected to global alignment with free end gaps (gap open penalty set to 12, gap extension penalty set to 3). The genetic distance model utilized was Jukes Cantor and the tree build method was neighbor-joining with no outgroup.

Molecular Modeling. The existing crystal structure of *S. aureus* Prp (PDB ID: 2P92) lacks the loop that includes the active site residues His 22 and Cys 34. The PDB file of the incomplete *Staphylococcus aureus* Prp/YsxB homolog dimer was read into SYBYL-X version 2.1. Modeller version 9.12 was used to model the loop that was not present in the crystal structure. One loop model was created using the loop_model.py script written by Dr. Hardik Parikh, and the loop was refined using the loop_refine.py script written by Dr Parikh. Models were created and read into SYBYL and one was chosen that contained a favorable His/Cys pro-catalytic conformation. The four residue loop (residues 60-63), had a closed conformation that partially obscured the active site of the enzyme. For this reason, the loop_refine.py script was again utilized to allow the four residue loop to adopt a conformation that did not obscure the active site. The model with greatest solvent exposure of the catalytic site was chosen. This model was read into SYBYL, along with

the original crystal 2P92 which formed the template for modeling the dimer. This dimer was read into Gold 5.2, and docking was achieved using the blocked peptide Ac (Acetyl)-Asn-Leu-Gln-Phe-Phe-Ala-Ser-Am (Amide).

Molecular dynamics (MD) simulations were carried out with the NAMD 2.8 package developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign (Phillips, 2005). CHARMM (Charmm-27) was used as the force field (MacKerell, 1998). Prior to simulation, the ionization states of His residues were checked using PROPKA 3.1 (Olsson, 2011). The analysis of the MD trajectory was done in VMD (Humphrey, 1996).

Bacterial Culture. All *S. aureus* strains used in this work are derivatives of the phage-cured, restriction-defective strain RN4220 (Kreiswirth, 1983). The *S. aureus* expression strain SA178RI carries a T7 RNA polymerase expression cassette under *lac* operator control (D'Elia, 2006). The plasmid used for expression in SA178RI was pG164, an *E. coli-S. aureus* shuttle vector carrying the T7 late promoter into which a *lac* operator, a multiple cloning site, an optimized gram-positive ribosome binding site, and a constitutively expressed copy of the *lac* repressor gene were introduced (D'Elia, 2006). *S. aureus* strains were grown in Trypticase Soy Broth (TSB) (Remel, Lenexa, KS) or Brain Heart Infusion (BHI) (Remel, Lenexa, KS) at 32ºC.

Cell growth was routinely measured with a Klett-Summerson colorimeter. *E. coli* strains were grown in Luria Bertani (LB) (Difco, Franklin Lakes, NJ) medium with appropriate antibiotics at 37ºC at 200 rpm to Klett of 90, then shifted to 32ºC at 200 rpm for protein induction. *E. coli* strains containing a plasmid with the P_{BAD} promoter were induced using 0.2% arabinose. All strains containing isopropyl β-D-1−thiogalactopyranoside （IPTG)-inducible promoters were induced with 1mM IPTG (Goldbio; St. Louis, MO).

DNA manipulations. Restriction endonucleases, T4 DNA ligase, Antarctic phosphatase, Polynucleotide kinase, specific buffers and BSA used for DNA manipulation were purchased from New England Biolabs (NEB; Ipswich, MA) and used as recommended by the manufacturer. Polymerases for PCR were exclusively high-fidelity - either Pfu Ultra II Fusion (Agilent, Santa Clara), Advantage HD Polymerase (Clontech-Takara; Mountainview, CA) or Phusion (NEB; Ipswich, MA). DNA was extracted from agarose gels using the QIAquick Gel Extraction Kit (Qiagen; Valencia, CA) or the Nucleospin Gel and PCR Cleanup Kit (Macherey-Nagel Inc; Bethlehem, PA) as described by the manufacturer. PCR products were purified using the QIAquick PCR purification kit (Qiagen; Valencia, CA) or the Nucleospin Gel and PCR Cleanup Kit (Macherey-Nagel Inc; Bethlehem, PA) as recommended by the manufacturer. Plasmid DNA was isolated from *E. coli* transformants that were grown overnight in 4 ml of LB broth with appropriate antibiotic, using a QIAprep Spin Miniprep kit (Qiagen; Valencia, CA) as described by the manufacturer. Gibson assembly was performed using the In-Fusion Kit (Clontech-Takara; Mountainview, CA).

Polymerase Chain Reaction (PCR). Polymerase chain reactions (PCRs) were performed in a T-Gradient thermocycler (Whatman Biometra, Goettingen, Germany). Oligonucleotides used in this study are listed in Table 2. These oligonucleotides were designed from the *Staphylococcus aureus* NCTC8325 genome (GenBank NC_007795.1) and were purchased from Integrated DNA Technologies (Coralville, IA). Oligonucleotides were reconstituted in HPLC grade water (Mallinckrodt Baker; Phillipsburg, NJ) to 1mM stock solution. Working solutions of oligonucleotides were prepared by diluting these stock solutions to 10 µM concentration.

Standard PCR mixtures were set up on ice using MQ water or HPLC grade water and prepared as follows: 1X PFU Reaction Buffer, DNA template (2-50 pg plasmid or 50-500 ng genomic template), oligonucleotides at 0.5 µM final concentration, 100 µM dNTPs each (Invitrogen, Carlsbad, CA) and 1.25 units of PFU DNA Polymerase (Agilent, Santa Clara CA) in a total reaction volume of 50 µl. Initial denaturation was performed at 95ºC for 2 minutes followed by 34 cycles of the following operations: denaturation at 95ºC for 10 seconds, primer annealing at hybridization temperature (TM -5) for 15 seconds and primer extension at 72ºC for 15 seconds per kb of extension. A final extension at 72ºC for 5 minutes was performed after the final cycle and the reactions were chilled to 4ºC.

Agarose gels. Agarose gels were prepared by dissolving agarose in 1X TAE buffer at 100ºC (Fisher Scientific, Pittsburgh, PA). Ethidium bromide was added to 0.1 µg/ml prior to pouring the gel. Generally, PCR products were resolved on 1% agarose gels,

plasmids and genomic DNA (gDNA) on 0.7% agarose gels. 5X DNA loading dye (50% glycerol and 0.25% bromophenol blue in TAE) was added to the DNA solution in the ratio of 1:4 to reach 1X dye concentration. DNA was loaded on agarose gels and subjected to electrophoresis at about 5 volts/cm until the bromophenol blue had migrated roughly 3/4 of the gel length and then visualized under UV light. DNA size was measured against the Hyperladder (BioLine, Taunton MA) series of DNA MW markers.

Plasmid screening. Plasmids used in this study are listed in Table 1 and were constructed by Gibson assembly (Gibson, 2009) using the In-Fusion kit from Clontech-Takara (Mountainview, CA). Genes were amplified by PCR using Pfu Ultra II Fusion polymerase (Agilent) with the primers listed in Table 2. PCR products and linearized vectors were gel purified (Nucleospin Kit, Macherey Nagel; Bethlehem, PA) and assembled using an In-Fusion Kit (Clontech). Assembled plasmids were introduced initially into *E. coli* Stellar® Competent cells (Clontech) and verified by sequence analysis (MWG Biotech). Protein expression was carried out in *E. coli* BL21 (DE3) RIL (Invitrogen) or *S. aureus* SA178RI.

Preparation of electrocompetent cells. Electrocompetent *S. aureus* cells were prepared from cultures grown in BHI at 37°C with shaking at 200 rpm, to ~2 x 10^8 cells/ml. The cells were then chilled in an ice water bath for 15 min to arrest growth and harvested by centrifugation at 12,000 x g for 15 min at 4°C. The supernatant was carefully removed and the cell pellet was suspended in 50 ml of sterile, ice-cold water and centrifuged again at 12,000 x g for 15 min at 4°C. The cells were washed two more times in 50 ml of sterile,

ice-cold water. The cell pellet was suspended in 25 ml of sterile, ice-cold 10% glycerol, then the cells were pelleted at 4,000 x g for 15 min at 4°C and resuspended in 2 ml final volume of 10% glycerol to a final concentration of about 1 x 10^{10} cells/ml. Cells were distributed in 65 µl aliquots into sterile 1.5 ml microfuge tubes and stored at -70°C.

Transformation of *E. coli* Gibson assembly reactions (2.5 µl), or ~30 ng of purified plasmid DNA, were added to 50 µl of Stellar *E. coli* chemically competent cells (Clontech) previously thawed on ice. This mixture was transferred to a cold 15ml conical tube. This tube was immersed in a 42⁰C water bath for 45 seconds. 1 ml of SOC medium (2% (wt/vol) tryptone, 0.5% (wt/vol) yeast extract, 85.5 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 20 mM glucose) warmed to 37⁰C was immediately added to the tube and cells were gently but quickly resuspended by flicking the tube. The cells were incubated for one hour at 37°C with shaking. Following this recovery period the cells were plated on LB plates with appropriate antibiotic selection.

Transformation of *S. aureus* Purified plasmid DNA (5 µl) was added to 65 µl aliquots of electrocompetent *S. aureus* and incubated on ice for 20 minutes. The mixture was then transferred to an electroporation cuvette with a gap length of 0.1 cm and pulsed one time using the MicroPulser (Bio-Rad, Hercules, CA) pre-set *S. aureus* setting Sta (1.8 kV, 2.5 msec, 25 µF). Following electroporation, one ml of Brain Heart Infusion (BHI) broth (Remel, Lenexa, KS) was added immediately to the cuvette and this mixture was then transferred to a 15ml conical tube. The cells were incubated for 1.5–2 hours with shaking at 30°C. Aliquots (5 µl, 100 µl and the rest of the cells pelleted to a volume not greater

than 100 µl) were then spread on TSA plates supplemented with appropriate antibiotics and incubated at 30°C for 48 hours.

Mutant Creation and Allelic Exchange. The strategy used for creating mutants can be divided into the following steps: (1) The mutant allele was created using Gibson assembly of overlapping pieces of DNA that contained the desired nucleotide changes in the manipulable 15 base pair (bp) overlap regions. Gibson assembly requires that at least 15 bp of any piece of DNA to be assembled to overlap with (be identical to) the neighboring piece to be joined (see Fig. 4). These 15 bp can be included on the primer, and thus can be of any sequence desired. For example, if a point mutation was desired, the codon change could be incorporated in this 15 bp region at the 5' end of the primer with no concern for how it would affect primer annealing. If a deletion or insertion was required, pieces of noncontiugous sequence could be knitted together in this manner with no intervening or unwanted sequence in between. This provides a mutant allele that can be exchanged using the shuttle plasmid pMAD. (2) Isolation of plasmid from *E. coli* for sequencing and then transforming the appropriate *S. aureus* strain and finally, if appropriate, (3) allelic exchange.
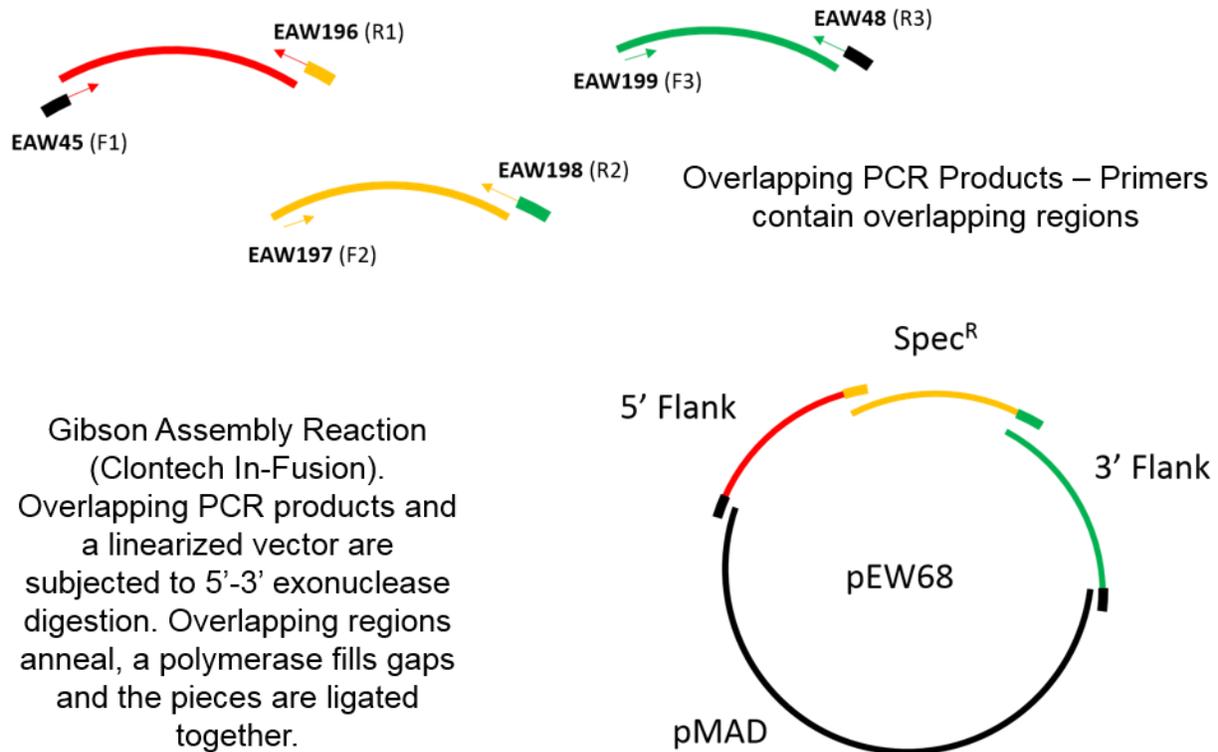
Figure 4. Gibson assembly of a plasmid that replaces native L27 in S. aureus with a spectinomycin resistance cassette – pEW68

Using the In-Fusion cloning technique, it is possible to produce a complete allelic exchange vector (or any multi-piece construction) from PCR products made with overlapping sequence. Primer EAW45, for example, has 15 base pairs of 5' overlap identical to BamHI cut pMAD (black), and EAW196 has 15 base pairs of 5' overlap to the next primer, EAW197, which will link each segment of DNA when the reaction is complete. Plasmid pEW68 is a typical allelic exchange vector that contains 5' and 3' flanking homology to the gene that will be replaced, in this case L27.

Allelic exchange was carried out using the methods described by Arnaud (Arnaud, 2004) with some modifications. See Allelic Exchange in Protocols. A shuttle vector (a vector that can replicate in *E. coli* and in *S. aureus*, thereby simplifying the process of plasmid isolation and cloning) called pMAD contains a temperature-sensitive *S. aureus* origin of replication that can only replicate independently in *S. aureus* at temperatures of around 30-32ºC. The pMAD vector also encodes an erythromycin resistance cassette for selection and beta galactosidase to allow blue-white screening. This allows us to select for bacteria containing the plasmid using an erythromycin resistance cassette, but then raise the temperature of growth, selecting against extrachromosomal plasmid maintenance. Cells cannot lose the plasmid and continue to live, so this selects for cells in which the plasmid has integrated into the bacterial chromosome. This cointegrate, when plated in the presence of 200 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (Xgal) will produce a blue colony. If designed correctly, the homologous sequence on either side of the mutation that we desire will allow this process to occur with efficiency. We then choose blue colonies, culture without selection for plasmid maintenance (in this case, erythromycin) and grow the culture at a permissive temperature for the plasmid. This should allow the plasmid to excise from the chromosome via homologous recombination, hopefully leaving behind the desired mutation. Then we select against plasmid maintenance by raising the temperature. This should allow curing of the plasmid after successive rounds of growth with no erythromycin. To ensure that the plasmid has been cured, white candidate colonies are patched onto erythromycin and must show sensitivity to be considered a possible allelic exchange candidate.

Colony PCR was performed on candidates that had lost the plasmid. For colony PCR, candidates are struck out to single colonies, and a medium-sized colony can serve as the template for PCR. The colony was rubbed onto the bottom interior of a PCR tube. The tube was microwaved on high for one minute, to rupture and dry out the cells, releasing DNA. After microwaving, 10 μl of MQ or HPLC water was added to the cells and they were vortexed, and then centrifuged to remove large chunks of cellular debris which can inhibit PCR. One μl of the supernatant served as template for one 50 μl PCR.

A note on the PCR and sequencing of allelic exchange candidates: care was always taken to amplify the region of interest with primers far outside the boundaries of the regions where homologous recombination with the pMAD derivative could have occurred, and then to sequence with primers only slightly interior to the amplifying ones (still outside the boundaries of possible recombination). This check confirms that there was no accidental incorporation of vector sequence into the chromosome due to a faulty recombination event.

Around the Horn mutagenesis. In lieu of the In-Fusion reaction, point mutations were sometimes introduced to a plasmid using divergent primers containing the desired change. The primers were first phosphorylated using Polynucleotide Kinase (PNK) so that they can later undergo ligation. The entire plasmid was amplified by PCR using these primers, then the template vector was digested by addition of 20 units Dpn1 to the reaction. DNA was purified by gel purification or PCR purification, depending on whether non-specific amplification products were present. The linearized product was re-

circularized via traditional ligation and introduced into competent *E. coli* cells by electroporation. Plasmids pEW21, 22 and 23 were created in this manner during the hunt for an un-cleavable point mutant of *S. aureus* L27.

Expression of proteins in *E. coli*. Expression of proteins in *E. coli* strain BL21 DE3 RIL (Invitrogen) was accomplished by growth until OD 0.6 or Klett 90 at 37°C and then induction with 1mM IPTG at 30-32°C. The cells were harvested 4 hours post induction by centrifugation at 4000 rpm for 10 minutes and frozen at -80°C until needed for protein purification. The method of lysis depended on the volume of cells used. For protein purification from a 3 liter culture, cells were resuspended in nickel-NTA column buffer (300 mM NaCl, 25 mM Tris pH 8, 10 mM imidazole) and were then lysed by passage through an EmulsiFlex C3 high-pressure homogenizer (Avestin, Inc., Ottawa). If the cell pellet came from a 400 ml culture or smaller, the cells were pelleted, resuspended in Ni-NTA column buffer, often with EDTA-free protease inhibitor tablets (Roche, Nutley NJ) and then lysed via mini bead-beater. Samples were centrifuged at high speed (greater than or equal to 12000 rpm) to separate lysate from pellet. The lysate was collected and used for purification, though a pellet sample would be saved for analysis.

Expression of proteins in *S. aureus*. For protein purification, overnight cultures of respective *S. aureus* strains containing pG164 derivatives were added to fresh media to a density of around Klett 10. The cells were allowed to grow till mid-late exponential phase (Klett 90). Inducer IPTG was added to a final concentration of 1 mM and the cells were shifted to 30 °C for four hours post induction. Cells were pelleted by centrifuging at 4000

rpm for 10 minutes and resuspended in PBS buffer with EDTA-free protease inhibitor (Roche, Nutley NJ). Cells were then mechanically disrupted using a mini bead beater and cell lysate was treated as above.

For growth curve analysis of mutant over-expression strains, overnight cultures were inoculated directly into media containing inducer and appropriate antibiotics (TSB with 2 µg/ml tetracycline, 30 µg/ml chloramphenicol and 1 mM IPTG) and monitored at OD_{600} for 16 hours in a 96-well plate format Biotek Microplate reader.

Growth curves for the complementation strains were performed differently, using a Klett spectrophotometer instead of a microplate reader. These strains contain wild type L27 under an IPTG-inducible T7 promoter. Overnight cultures grown in BHI tetracycline (Tet) 2 µg/ml, chloramphenicol (Chl) 15 µg/ml, spectinomycin (Spec) 250 µg/ml and 1 mM IPTG were inoculated to Klett=3 in fresh media and allowed to grow until the beginning of exponential phase, Klett=30. Five hundred microliter aliquots of those cultures were subsequently harvested at 6000 rpm for 2 minutes in a microcentrifuge. These cells were washed twice with 1ml BHI (free of IPTG) in order to deplete them of their initial inducer. In the case of the L27 depletion experiment, cells were then inoculated into BHI Tet 2 µg/ml, Chl 15 µg/ml, Spec 250 µg/ml and allowed to continue growing, with Klett measurements taken every hour for 10 hours. In the case where expression was switched to an arsenite-inducible copy of L27, these cells were inoculated into BHI Tet 2 µg/ml, Chl 15 µg/ml, Spec 250 µg/ml, $NaAsO_2$ 5 µM.

Polyacrylamide Gel Electrophoresis. Protein samples were analyzed using the Criterion XT protein system (Biorad, Hercules, CA) and were separated on either pre-cast Tris -Tricine 16.5% gels or 4-20% gradient gels (Biorad Hercules, CA). 2X Tricine loading buffer (for Tricine gels) or 4X XT loading buffer (for glycine gels) was added to samples to achieve a 1X concentration and these samples were then heated for 10 minutes at 75⁰C prior to gel loading. The proteins were separated under constant voltage at 200V in 1X Tris-Tricine running buffer (Bio-Rad, Hercules, CA). A Precision Plus Dual Xtra Protein Standard dual color marker (2 kDa-150 kDa) (Bio-Rad, Hercules, CA) was run with samples to estimate the molecular weight of the protein. After electrophoresis, the gels were rinsed in deionized water three times for 5 minutes each and then stained with Bio-Safe™ Coomassie (Bio-Rad, Hercules, CA) followed by destaining in 40% methanol, 10% acetic acid until the image was clear. The bands were visualized under visible light and images were taken using a Fotodyne imaging system (Hartland, WI).

Western blotting. Protein samples were loaded onto 16.5% Tris-Tricine precast gels (Biorad) for SDS PAGE and run until the 10kDa ladder band of the Precision Plus Protein Prestained Standards (Biorad) approached the bottom centimeter of the gel. For detection of L27, gels were electroblotted to a 0.45 micron PVDF membrane for 3 hours at 300 mA in pH 8.3 Tris-Glycine transfer buffer containing 0.05% SDS, to ensure movement of the small, positively charged L27 protein. For detection of the small, acidic Prp protein, gels were electroblotted in Towbin buffer (Tris-Glycine pH 8.3, 20% methanol) for 1 hour 45 minutes at 200 mA. The membranes were blocked with phosphate-buffered saline/ 0.1% tween-20 (PBST) containing 5% skim milk for one hour

at room temperature and then rocked overnight at 4ºC with a 1:5000 dilution of primary antibody. L27Myc-His$_6$ was detected with monoclonal mouse anti-Myc antibody (gift from Dr. William Barton) while untagged L27 and Prp were detected using rabbit sera with custom polyclonal antibodies raised to each C-terminally His$_6$-tagged protein in its denatured form (New England Peptide; Gardner, MA; Thermo Fisher; Waltham, MA respectively). All antibody dilutions were made in 5% skim milk PBST. The membrane was then washed thoroughly with PBST and probed with a 1:10000 dilution of Protein A conjugated to Horseradish Peroxidase (Invitrogen) in 5% skim milk PBST and protected from light for one hour at room temperature. The membrane was protected from light, washed thoroughly with PBST and then Pierce ECL Western Blotting substrate was added for 5 minutes per manufacturer instructions. The blots were then exposed to x-ray film (Bioexpress; Kaysville, UT) and developed.

Polysome analysis. Overnight cultures of strains for polysome analysis were inoculated into a flask of media containing inducer and appropriate antibiotics (TSB with 2 µg/ml Tet, 30 µg/ml Chl and 1 mM IPTG) beginning at a 1:100 dilution and grown until they reached a Klett reading of 90 or for 4 hours. As noted above, *S. aureus* expressing Prp C34A will not grow above Klett 50 for over 8 hours, so this strain was grown at a double volume compared to the others so that an equivalent number of cells could be harvested. Cell pellets from at least 100 ml of culture were resuspended in buffer PA (20 mM Tris pH 7.8, 100 mM NH$_4$Cl, 10 mM MgCl$_2$, 6 mM beta mercaptoethanol), broken by bead beating and then clarified by centrifugation for 10 minutes at 13000 rpm at 4ºC. The clarified cell lysates were applied to the top of 10-40% sucrose gradients in buffer PA.

The gradients were centrifuged for 17 hours at 19000 rpm in an SW-28 rotor at 4ºC. Gradients were analyzed for absorbance at 254 nm using a Biocomp Piston Gradient Fractionator with a BIORAD Econo UV Monitor with a Full Scale of 1.0. Data were recorded using DataQ DI-158-UP data acquisition software, as has been previously reported (O'Farrell, 2012). Areas under each peak were calculated using the trapezoid method.

Edman Degradation. Ni-NTA purified L27 Myc-His$_6$ protein samples were loaded onto a 16.5% Tris-tricine precast gel (Biorad) and run until the 10 kDa band of the Precision Plus Protein Prestained Standards (Biorad) approached the bottom of the gel. These proteins were electroblotted to a 0.45 µm PVDF membrane for 4 hours at 300 mA in a Criterion Tank Blotter (Biorad). The membrane was then stained with Coomassie and the appropriate bands were marked and coded and sent to the Iowa State University Protein Facility for Edman degradation. There the protein samples on the membrane were washed six times with deionized water and loaded onto a 494 Procise Protein Sequencer/140C Analyzer (Applied Biosystems, Inc).

Mass Spectrometry. Ni-NTA purified L27 Myc-His$_6$ protein fractions were separated on a 16.5% Tris-Tricine precast gel (Biorad) and stained with Coomassie Blue. The co-purifying protein was excised from the gel and the gel band was subjected to trypsin digestion and LC-MS analysis of tryptic peptides using a Thermo Electron hybrid LTQ-Orbitrap mass spectrometer.

Pooled sucrose gradient fractions containing the 50S ribosomal assembly intermediates were concentrated and de-salted prior to trypsin digestion and LC-MS/MS. The solutions were diluted with 50 µL 100 mM ammonium bicarbonate. The samples were reduced with 5 µL of 10 mM dithiothreitol in 0.1 M ammonium bicarbonate at room temperature for 0.5 h. Then they were alkylated with 5 µL 50 mM iodoacetamide in 0.1 M ammonium bicarbonate at room temperature for 0.5 h. The samples were digested with 1 µg trypsin overnight and then quenched with 5% (v:v) glacial acetic acid. 3 µL of the final solutions were injected for analysis.

The LC-MS system consisted of a Thermo Electron hybrid LTQ-Orbitrap mass spectrometer system with a nanospray ion source interfaced to a Waters nanoAcquity UPLC system equipped with a Waters NanoAcquity C18 column. 5 µL of the extract was injected and the peptides eluted from the column by an acetonitrile/0.1 M acetic acid gradient at a flow rate of 0.4 µL/min over 60 minutes. The nanospray ion source was operated at 3.5 kV. The digest was analyzed using the double play capability of the instrument acquiring full scan mass spectra to determine peptide molecular weights and product ion spectra to determine amino acid sequence in sequential scans. This mode of analysis produces approximately 10000 CAD spectra of ions ranging in abundance over several orders of magnitude. Not all CAD spectra are derived from peptides.

The data were analyzed by database searching using the Mascot search algorithm against NCBI's non-redundant database. The relative amounts of selected ribosomal proteins in each sample were determined from the MS1 spectra by measuring the areas

under the peaks of individual peptides from each protein and normalizing within each sample to the peptides from the known early assembly proteins L3 and L4.

Development of Prp Assay In order to study Prp enzymology in vitro, an assay was designed using a quenched fluorescent peptide substrate derived from the cleavage motif on L27. The peptide KLNLQFFASKK was chosen for its content of hydrophilic residues and fluorophore 2-amino benzoic acid (2-abz) was conjugated to the N terminus, with a dinitrophenol (Dnp) quencher conjugated to the C-terminus. The resulting peptide substrate (2abz-KLNLQFFASKK-Dnp) was ordered from United Peptide in Herndon, VA (now United Biosystems). The substrate was initially resuspended to 2 mM with a buffer containing 25 mM Tris pH8 and 300 mM NaCl. Fluorescence was measured using an excitation wavelength of 325 nm and an emission wavelength of 425 nm using a Photon Technology International QuantaMaster steady-state fluorescent spectrometer. Prp was initially purified and analyzed as a His_6-SUMO N-terminal fusion protein from plasmid pEW34. His_6-SUMO fusions were also made of PrpC34A (pEW39) and PrpC34S (pEW40).

Chapter 4

Staphylococcus aureus L27 is cleaved by novel protease Prp

We previously noted that L27 proteins in *S. aureus* and other Firmicute bacteria have an N-terminal extension that is not present in *E. coli* L27 (Spilman, 2012). Cleavage of the extension was predicted to generate an N-terminus highly similar to that found in *E. coli* L27, which is known to aid in peptidyl transfer. Our bioinformatic analyses revealed that all sequenced Firmicutes, Fusobacteria, and Synergistetes, as well as some Thermatogae and Tenericutes, encode an L27 containing an N-terminal extension with the conserved cleavage motif. As noted in the introduction, everything thus far known about L27 suggested that this N-terminal extension would be highly toxic to the ribosome because it would occlude the conserved residues that stabilize tRNAs at the PTC. Only the post-cleavage form of L27 was found in mature *B. subtilis* ribosomes, supporting the validity of this theory (Lauber, 2009).

Taken together, these data support the existence of a group of evolutionarily related bacteria that exhibit a fundamental difference in the basic biology of the ribosome involving a previously undescribed ribosomal protein processing event. Since L27 is positioned at the heart of the ribosomal active site, where it plays a key role in translation

(Maguire, 2005), we sought to elucidate further the role of this N-terminal extension in *S. aureus* biology.

L27 cleavage

We tested the predicted N-terminal cleavage of *S. aureus* ribosomal protein L27 by comparing the products of the cloned full length gene in *S. aureus* and *E. coli,* as we had previously done for the phage 80α capsid and scaffold proteins (Spilman, 2012)*.* To identify and purify *S. aureus* L27 in *E. coli*, we introduced C-terminal Myc and His$_6$ tags using vector pBADMycHis A (Invitrogen). The same tagged L27 cassette was introduced into the T7 expression plasmid pG164 (D'Elia, 2006) for expression in *S. aureus.* The His-tagged proteins in clarified lysates from each overexpression strain were batch adsorbed to Nickel NTA resin (Clontech) and purified proteins were then examined on Western blots probed with anti-Myc antibodies. L27 isolated from *S. aureus* is approximately 1 kDa smaller than the same protein isolated from *E. coli* (Fig. 5), consistent with the predicted N-terminal cleavage by a *S. aureus* protease that is absent from *E. coli*. The L27 protein expressed in *S. aureus* was extracted from the gel and subjected to N-terminal protein sequencing by Edman degradation. The first four residues were ASKK, confirming the predicted processing of L27 in *S. aureus* at the conserved cleavage motif*.* Apart from the common removal of the N-terminal formyl methionine, this is the first example of a specific N-terminal cleavage of a bacterial ribosomal protein precursor.
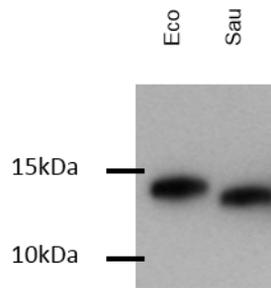
Figure 5. *S. aureus* **L27 is cleaved at the conserved motif when it is expressed in its native host but not when expressed in** *E. coli*

Western blot analysis of L27 N-terminal cleavage. Plasmid-encoded *S. aureus* L27 carried C-terminal Myc and His$_6$ tags; the blot was probed with anti-Myc antibodies. In the lane labeled Eco, L27 was produced from pEW5 in *E. coli* BL21. In the lane labeled Sau: L27 was produced from pEW7 in *S. aureus* SA178RI. L27 in the Sau lane has undergone specific cleavage at the conserved motif to yield the N-terminus ASKK, confirmed by Edman degradation.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

Image reproduced with permission

Prp discovery

A protease or peptidase is an enzyme that can catalyze cleavage of a protein or peptide substrate. There are seven known classes of proteases, the most well-studied of which are the hydrolases. Cysteine, serine and threonine proteases use their respective side-chain nucleophiles to attack the carbonyl carbon of the substrate peptide bond with cleavage of its link to the amide nitrogen and formation of a covalent enzyme/substrate acyl intermediate. This intermediate is hydrolyzed and the product is released. In contrast, aspartate, glutamate and metalloproteases perform acid/base catalysis using an activated water molecule. The seventh and most recently characterized group of proteases are not hydrolases: asparagine amidine lyases self-cleave using a coordinated water molecule to form a succinimide intermediate that cleaves the Asn-Asn peptide bond. Intein auto-cleavage domains utilize this mechanism (Rawlings, 2011).

From sequence comparison of *S. aureus* L27 with phage 80α capsid and scaffold proteins as well as expression of those proteins in *E. coli*, it seemed extremely unlikely that the specific N-terminal cleavage event was autocatalytic. That left one of six possible hydrolases, all of which have some identifiable set of catalytic residues that would have to be completely conserved across mechanistic homologs. There have been many classified proteases and the MEROPS database has an extensive collection of annotated and experimentally confirmed protease families.

We expected the protease responsible for *S. aureus* L27 cleavage to be essential and conserved among bacteria with the L27 N-terminal cleavage motif. Our attention was drawn to an open reading frame of unknown function, first designated *ysxB* in *B. subtilis* (locus tag SAOUHSC_01756 in *S. aureus* NCTC8325), located between the genes encoding L21 (*rplU*) and L27 (*rpmA*). The occurrence of this intervening reading frame was first noticed in a sequencing project in 1985 which showed both that by sequence, L27 in *B. subtilis* was N-terminally extended compared to that in *E. coli*, and that there was a gene upstream of L27 that bore no homology to L21 (Ferrari, 1985). It was also noted in a bioinformatic study of horizontal gene transfer between Gram positive bacteria that the intervening reading frame belonged to an established conserved domain family named COG2868 and was maintained between L21 and L27 among *Bacilli*, *Spirochaetales* and *Theromotogales*. (Garcia-Vallve, 2002).

YsxB was an attractive candidate for the protease for several reasons: First, all bacteria that encode an L27 with the N-terminal extension also carry a *ysxB* gene (Fig. 8). Second, both *rpmA* and *ysxB* were classified as essential in *S. aureus* by saturation transposon mutagenesis (Chaudhuri, 2009) and in an earlier antisense RNA study (Ji, 2001). Third, the YsxB protein has features characteristic of a protease. The YsxB structure was solved in a structural survey of *S. aureus* proteins (PDB ID: 2P92) and was grouped with other similar structurally characterized proteins containing a common domain in NCBI's Conserved Domain Database (CDD) under the designation DUF464. The crystal structure of a DUF464 member from *Thermatoga maritima* (1S1L) was published in 2005, and the authors noted that the conserved residues could possibly

constitute a catalytic center (Shin, 2005). The structures of two additional family members from *Streptococcus mutans* (PDB ID: 2G0I) and *Streptococcus pneumoniae* (PDB ID: 2IDL) have also been solved. DUF464 proteins are distinguished by a pair of invariant histidine and cysteine residues with conserved spacing that form the classic catalytic dyad of a cysteine protease (Fig. 7A) and are clustered together in a cleft on observed DUF464 structures (Fig. 7B). In addition to the *ysxB* homologues present upstream of *rpmA*, there are a number of phages that encode a DUF464 family member. One member of this family has been shown to be a bona fide bacteriophage prohead protease, responsible for capsid protein cleavage in streptococcal phage Cp-1 (Martín, 1998).
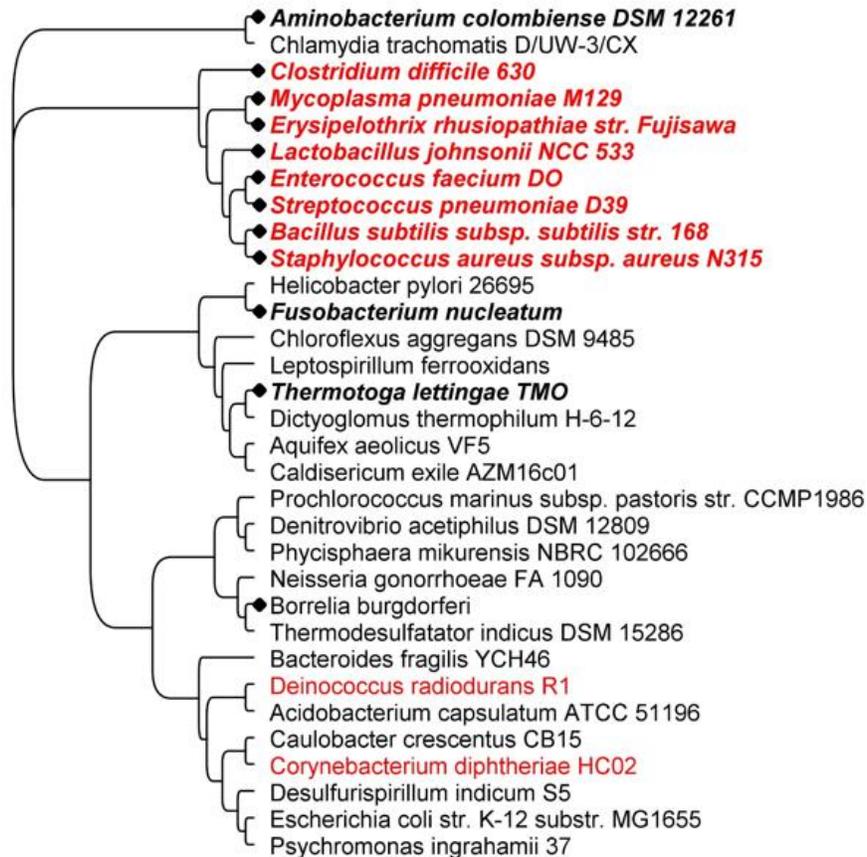
Figure 6. Phylogenetic tree of L27 shows that upstream gene *ysxB* (*prp*) is co-conserved with the N-terminal cleavage motif

L27 phylogeny across major bacterial phyla. This neighbor-joining phylogenetic tree (Jukes Cantor; BLOSUM80) consists of the sequences of ribosomal protein L27 from representative species of each major bacterial phylum. Gram positive species are designated in red, Gram negative in black. Species containing the conserved N-terminal extension of L27 are bolded and italicized. Species containing a ysxB/DUF464/Prp homolog are indicated by black diamonds. <u>All species containing the L27 N-terminal extension also contain a Prp homolog.</u>

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

Image reproduced with permission

Figure 7. YsxB contains conserved catalytic residues His and Cys in a pocket

(A). ClustalW alignment of YsxB sequences from *S. aureus* (Sau), *B. subtilis* (Bsu), *Streptococcus pneumoniae* (Spn), *Streptococcus mutans* (Smu), *Aminobacterium colombiense* (Aco), *Fusobacterium nucleatum* (Fnu), *Thermotoga lettingae* (Tle) and *Mycoplasma pneumoniae* (Mpn). Identical residues are highlighted in solid red, conserved residues in yellow. Asterisks designate the proposed active site residues. The secondary structure elements from the crystal structure of *S. aureus* Prp (PDB ID: 2P92) are shown above the alignment. Figure made with ESPript 2.2 (Gouet, 1999).

(B) Molecular surface of the complete *S. aureus* YsxB monomer model generated by I-TASSER (Roy *et al.*, 2010) using the partial crystal structure of the *S. aureus* protein (PDB ID: 2P92) and the complete structures of *S. pneumoniae* (PDB ID: 2IDL) and *S. mutans* (PDB ID: 2G0J) as templates. Residues are colored by conservation from white (non-conserved) to yellow (partially conserved) and then red (completely conserved) based on the alignment in (B). The active site Cys and His residues are indicated by arrowheads. Figure created in UCSF Chimera (Pettersen, 2004)

Figure adapted from Wall, 2015 Figure 2 with permission.

When *S. aureus* L27 Myc-His$_6$ and untagged YsxB were co-expressed in *E. coli*, the resulting tagged and purified protein was the same size as L27 produced in *S. aureus;* both are approximately 1 kDa smaller than L27 expressed in *E. coli* in the absence of YsxB (Figure 10). Edman degradation confirmed that the N-terminus of *S. aureus* L27Myc-His$_6$ co-expressed with YsxB in *E. coli* was also ASKK. These data demonstrate that YsxB can perform specific N-terminal cleavage of *S. aureus* L27 in *E. coli* and confirm that this conserved gene of previously unknown function encodes the L27 protease. We have named this gene *prp*, for p̲hage-related r̲ibosomal p̲rotease. Further evidence to support the classification of this protein as a cysteine protease was provided by mutation of the predicted catalytic cysteine residue, C34, to Ala or Ser, which resulted in the inactivation of Prp (Fig 8A, lane "Eco + PrpC34A"; Fig. 8B lane 2). Furthermore, when the un-cleaved L27Myc-His was purified from a strain co-expressing Prp C34A or C34S, the mutant Prp co-purified as an 11 kDa band on the gel accompanying the 13 kDa L27Myc-His$_6$ protein (Fig. 9). The identity of the 11 kDa co-purifying band was confirmed by mass spectrometry. These observations indicate that the mutant protease remains tightly bound to its substrate in the absence of catalysis. They also show that replacement of cysteine with serine in the catalytic site does not allow Prp to function as a serine protease, despite the presence of a conserved aspartic acid or asparagine in the DUF464 motif.
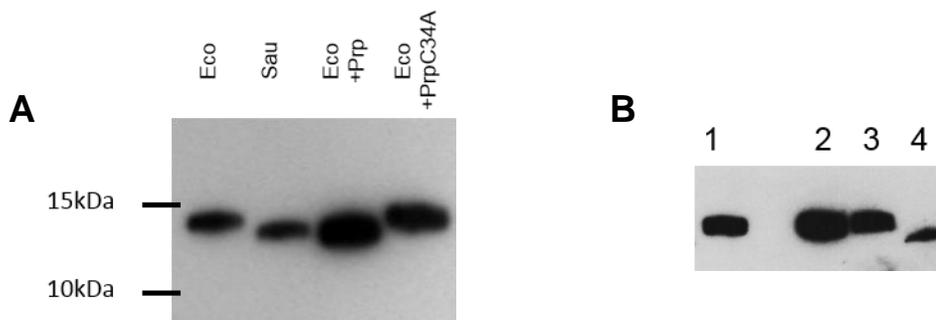
Figure 8. Prp is a cysteine protease that cleaves L27 at the conserved motif; Prp C34S mutant is unable to cleave L27 despite D/N residue at the catalytic core

(A) Western blot analysis of L27 N-terminal cleavage. Plasmid-encoded *S. aureus* L27 carried C-terminal Myc and His$_6$ tags; the blot was probed with anti-Myc antibodies. Lane "Eco": L27 produced from pEW5 in *E. coli* BL21; lane "Sau": L27 produced from pEW7 in *S. aureus* SA178RI; lane "Eco + Prp": L27 co-expressed with Prp in *E. coli* BL21 (pEW5 + pEW15); lane "Eco + PrpC34A": L27 co-expressed with inactive Prp mutant C34A in *E. coli* BL21 (pEW5 + pEW28). L27 in lanes Sau and Eco + Prp has undergone specific cleavage at the conserved motif to yield the N-terminus ASKK, confirmed by Edman degradation.

(B) Western blot (anti-Myc) of *S. aureus* L27MycHis$_6$ produced alone in *E. coli* BL21 (lane 1), co-expressed with Prp C34S (lane 2, after a space), co-expressed with Prp C34A (lane 3) and co-expressed with wild-type Prp (lane 4). In lanes 1-3 L27MycHis$_6$ remains un-cleaved, while in the presence of wild-type Prp in lane 4, it exhibits the characteristic 1kDa downward shift that represents specific N-terminal cleavage between P1 Phe 9 and P1' Ala 10. This was verified by Edman degradation.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]
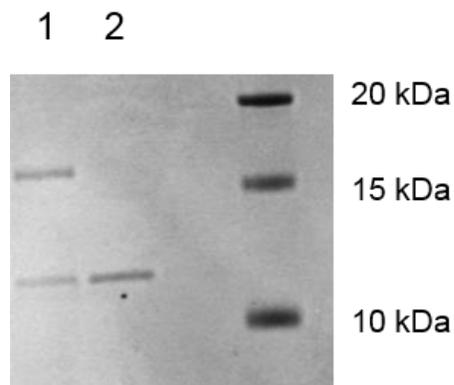
Figure 9. Prp C34A co-purifies with L27MycHis$_6$

Coomassie-stained 16.5% Tris-Tricine gel (Biorad) of L27MycHis$_6$ co-expressed with Prp C34A (lane 1). Lane 2 contains only purified Prp. The top band in lane 1 is Ni-NTA purified *S. aureus* L27MycHis$_6$. The identity of the co-purifying lower band in lane 1 was confirmed to be Prp by mass spectrometry.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

The original identification of Prp was based on the cleavage of the major capsid and scaffolding proteins of bacteriophage 80α in *S. aureus*. We therefore confirmed that Prp was able to carry out the previously observed N-terminal cleavage of these two phage structural proteins upon co-expression in *E. coli* (Fig. 10). The small differences in electrophoretic mobility of these two proteins were shown to result from the expected sequence-specific cleavages by determining the total mass of the processed proteins by MALDI-MS and further confirmed by LC-MS of tryptic peptides. Prp thus serves an additional function in *S. aureus*, as the prohead protease for 80α and related staphylococcal phages whose capsid and scaffold proteins share the conserved N-terminal cleavage motif.
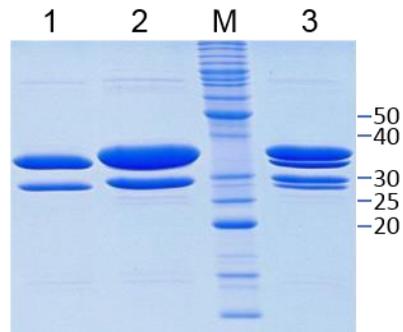
Figure 10. Prp cleavage of bacteriophage 80α scaffold and major capsid proteins.

Experiment performed by Keith Manning in the Dokland lab at University of Alabama Birmingham. Capsid-like structures formed upon expression of gp46 and gp47 in *E. coli* from plasmid pPD2 in the absence (lane 2) and presence (lane 1) of pEW15, expressing Prp, were purified on sucrose gradients and resolved by SDS-PAGE. Lane 3 shows the two samples mixed together, in order to better illustrate the small shift in electrophoretic mobilities of the two proteins following cleavage. M is the BenchMark protein ladder (Life Technologies), with selected sizes indicated next to the gel. A MALDI spectrum of the purified capsid fraction run in lane 1 showed major peaks at 21,704 Da and 35,042 Da, corresponding to the masses of cleaved gp46 (theoretical mass 21,701 Da) and cleaved gp47 (35,063 Da), respectively. Tryptic fragments of these two protein bands were also analyzed by LC-MS to confirm that they corresponded to the expected cleaved proteins.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

Image reproduced with permission

Chapter 5

Effects of Prp and L27 Mutant Expression on Ribosome Assembly

Both L27 and Prp are essential in *S. aureus*. One common approach is to examine the effects of mutations in essential genes by over-expressing mutant alleles and assess the resulting phenotypes.

We investigated the effects of mutant Prp by examining the growth of strains overexpressing catalytically inactive Prp C34A compared to wild-type Prp, using a *S. aureus* T7 expression system (D'Elia, 2006). Expression from the plasmids in *S. aureus* strain SA178RI was induced for 16 hours at 32° C and growth was compared to the same strain with empty vector. Overproduction of Prp C34A led to a marked growth defect compared to overproduction of the wild-type Prp (Fig. 11). (Note that the expression strain still contains native L27 and Prp.)
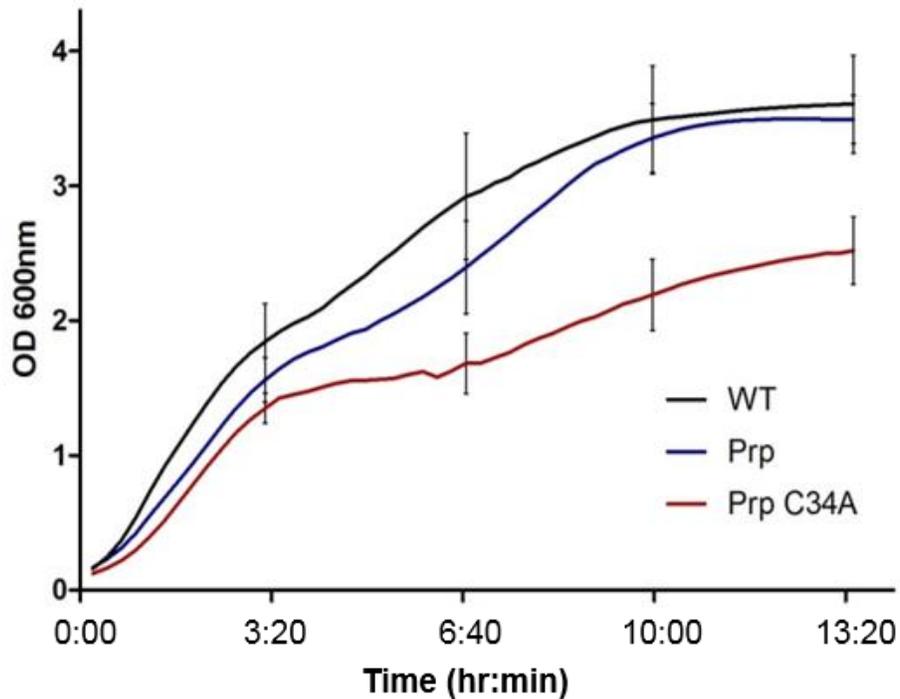
Figure 11. Growth effects of Prp C34A over-expression

Growth curves of SA178RI containing empty T7 expression vector pG164 (WT), or derivatives expressing wild-type Prp (pEW25) or Prp C34A (pEW26). Bacteria were grown in the presence of inducer and appropriate antibiotics in 96-well format at 32º C. Starting culture density was between 0.12 and 0.15 absorbance at 600nm, and absorbance was measured every 15 minutes for 16 hours using a BioTek Microplate reader. Error bars represent a 95% CI, n=3.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

The role of Prp in *S. aureus* ribosome assembly was investigated by examining ribosomal profiles from cells overexpressing either wild-type Prp or inactive Prp C34A. Overexpression of Prp C34A had a very significant effect on ribosomal subunit profiles, resulting in a 50% decrease in mature 70S ribosomes. There was also an accumulation of what appeared to be pre-50S assembly intermediates (Fig. 12), suggesting that proper ribosome assembly is dependent on L27 cleavage. However, the dominant negative phenotype of Prp C34A could have been a result of catalytically inactive Prp binding to L27 without cleaving, and thereby sequestering it from participation in ribosomal assembly. L27 has a known role in 50S subunit assembly and the ribosomal subunit pattern in the strain overexpressing Prp C34A resembled what has been observed with a partial depletion of L27 from the large subunit (Wower, 1998). The effect of wild-type Prp overexpression on ribosomal subunit proportions did not deviate significantly from empty vector.
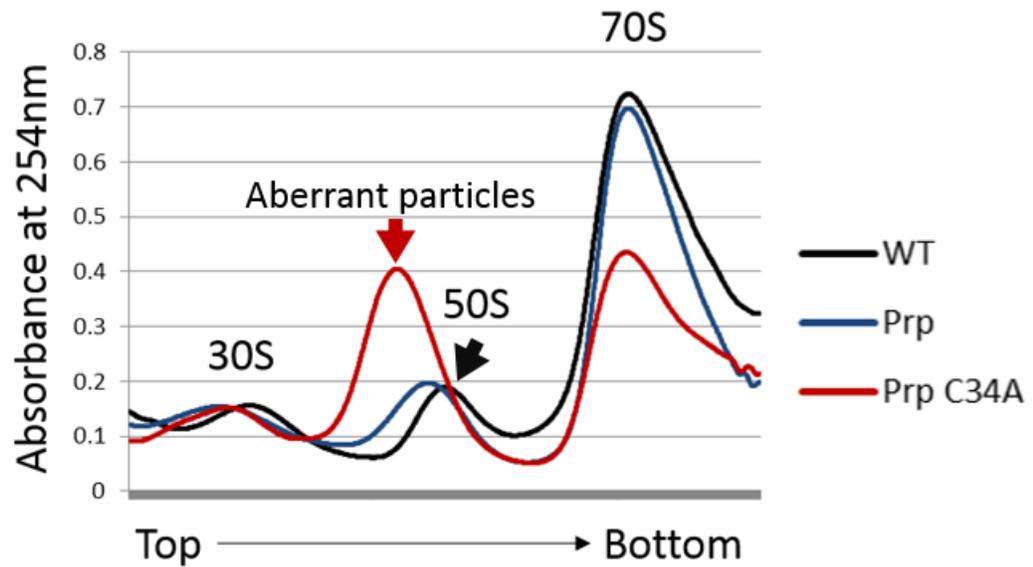
Figure 12. Ribosome assembly profile for strains expressing Prp and dominant negative Prp C34A

Typical sucrose gradient profiles of ribosomal subunits are shown with the positions of expected subunits indicated. The accumulation of what appears to be a pre-50S subunit in the Prp C34A strain indicates that a ribosomal assembly bottleneck has occurred.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

Variant forms of L27 were also overexpressed to determine their effects on cell growth and ribosomal subunit distribution: wild-type, pre-cleaved (deletion of L27 residues 2-9) or un-cleavable (L27 F8A-F9A, giving motif MLKLNLQAAA; Fig. 13). In contrast to the strong dominant negative effect of Prp C34A, overexpression of the mutant forms of L27 showed little effect on cell growth (a mild defect was seen with the un-cleavable mutant; Fig 14) and yielded ribosomal subunit profiles similar to cells expressing empty vector (Fig. 15). The presence of native levels of L27 and Prp in this strain likely masks any effects of these L27 N-terminal cleavage mutants in ribosomal assembly. We also investigated the effect on *S. aureus* growth of overexpressing the L27 N-terminal peptide that would be released upon cleavage (MLKLNLQFF). The growth curve was qualitatively and quantitatively different from that of L27 overexpression, exhibiting a pronounced lag in the middle of the exponential growth phase before resuming growth (Fig. 16). Preliminary examination of the ribosomal subunit profile of this strain showed no anomalies. How the overexpressed peptide alone is able to exert this striking effect on *S. aureus* growth is an open question. It is possible that the cleaved peptide acts as a competitive inhibitor of Prp, blocking L27 cleavage, or that it has another role in the cell, perhaps in a regulatory capacity.
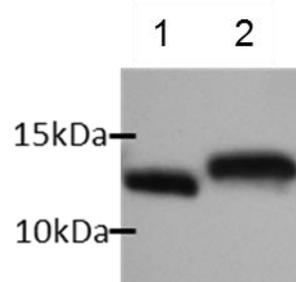
Figure 13. L27 F8A-F9A is un-cleavable

S. aureus L27 and an L27 cleavage motif mutant in which F8 and F9 were replaced with alanines were C-terminally tagged with Myc and His_6 and expressed in S. aureus strain SA178RI. The Western blot shows tagged proteins post purification probed with anti-Myc antibodies. Lane 1, wild-type L27; lane 2 L27 F8A-F9A. The characteristic 1kDa difference between cleaved and un-cleaved is evident.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

Image reproduced with permission
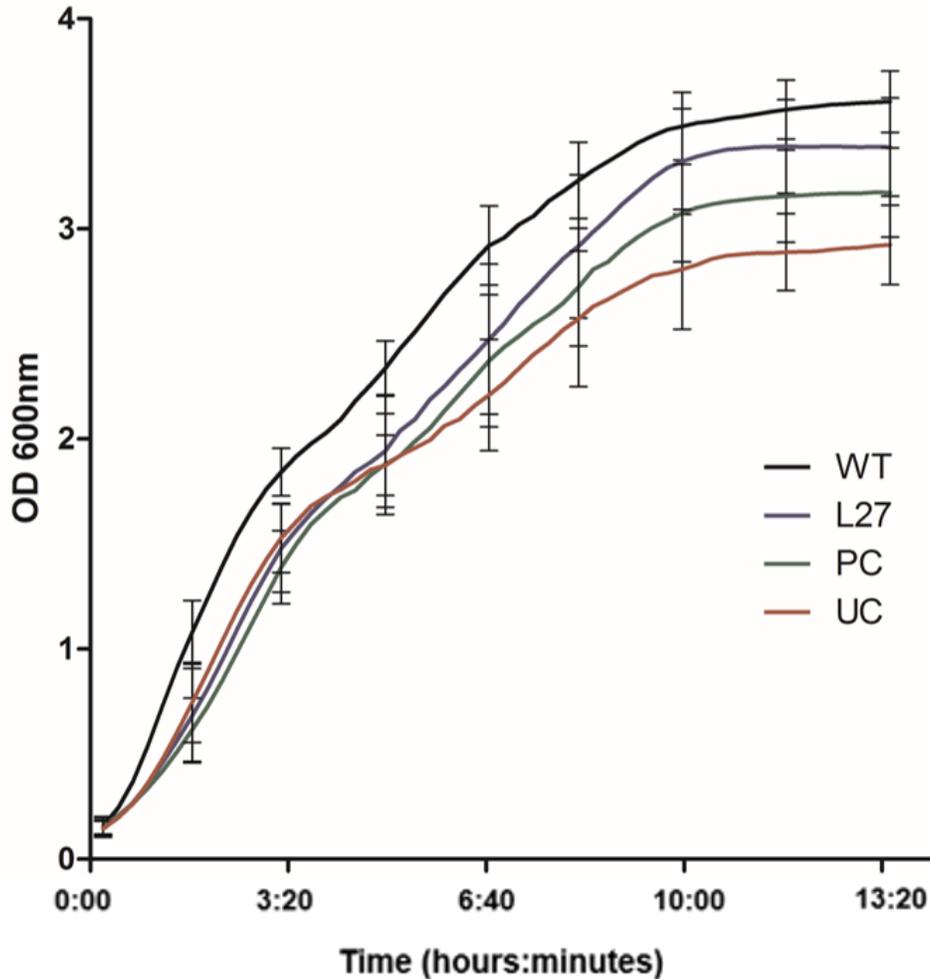
Figure 14. Cell growth following overexpression of L27 mutants

S. aureus strains overexpressing L27 (L27), pre-cleaved L27 (PC), un-cleavable L27 F8A-F9A (UC) or containing empty vector (WT) were grown under inducing conditions in a BioTek Microplate Reader at 32 degrees C until stationary phase. Optical Density at 600nm was measured once every 15 minutes. Error bars represent one standard deviation above and below the mean, n=3. There is a significant difference in growth between L27 and UC from t= 10 hours to 1=13:30.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

Image reproduced with permission
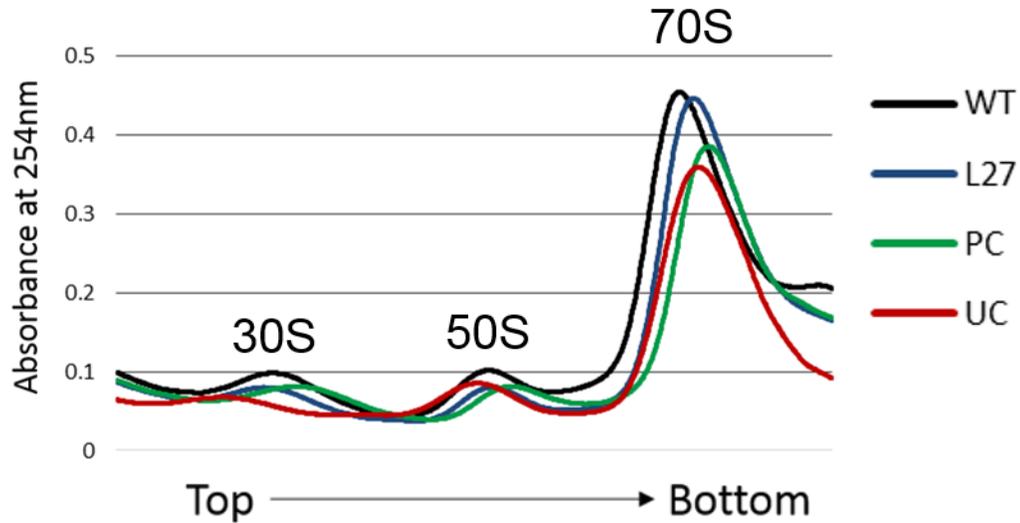
Figure 15. Polysome analysis for strains expressing L27 and L27 mutants

Typical sucrose gradient profiles are shown. The effect of L27 mutant overexpression on ribosomal subunit proportions did not deviate significantly from empty vector. *S. aureus* strains overexpressing L27 (L27), pre-cleaved L27 (PC), un-cleavable L27 F8A-F9A (UC) or containing empty vector (WT) were grown under the same conditions as those in Figure 12.
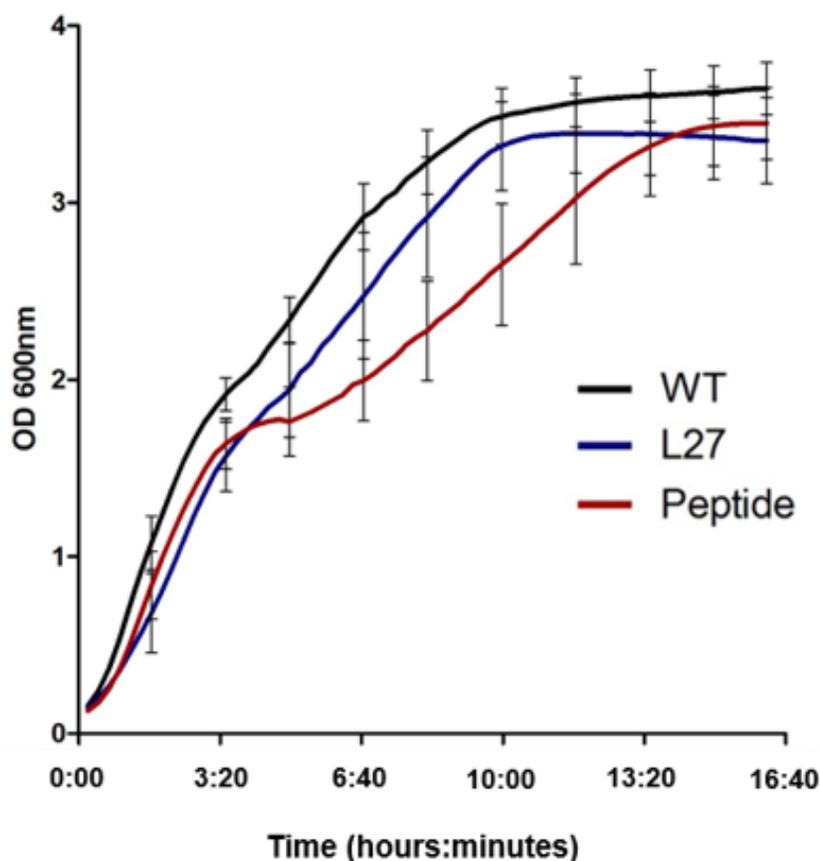
Figure 16. Growth effects of overexpressing the peptide MLKLNLQFF

S. aureus strains overexpressing L27 (L27), L27 residues 1-9 as a peptide (Peptide) or containing empty vector (WT) were grown as above. Expression of the peptide caused a deformation in the growth curve at t=3:20 that is not seen in WT and L27. Growth is significantly different between L27 overexpression and peptide overexpression from t=8:20 to t=10:30. Error bars represent one standard deviation above and below the mean, n=3.

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

We further characterized the aberrant ribosomal particle produced in the presence of Prp C34A to determine the stage of assembly at which they were trapped. The composition of both wild type and aberrant 50S particles were analyzed by mass spectrometry (LC MS/MS). We chose this approach because sucrose gradient fractions are a complex mixture of both ribosomal proteins and co-migrating cellular protein complexes with a similar sedimentation coefficient. Furthermore, these cells also contain wild-type ribosomes. LC MS/MS of tryptic peptides allowed us to confirm the presence of specific ribosomal proteins in these fractions, while examination of the MS1 spectra of the peptides from individual ribosomal proteins allowed us to determine the relative amounts of several specific proteins within each sample. Chromatograms for the precursor masses of two tryptic peptides from all large subunit ribosomal proteins detected were generated from both particle types. Peptide peak areas for identical peptides in both samples were calculated using conventional tryptic digest mass spectrometry MS 1 spectra (Table 3).

We were able to compare peptide abundances for all detected r-proteins to those of two early assembly proteins, L3 and L4 (Figure 17; Table 3). The ratio of L3 and L4 to each other and to the vast majority of other ribosomal proteins detected remained the same, within a standard deviation above and below, in both the 50S and aberrant particles, consistent with the inherent stoichiometry of ribosomal proteins. However, the internal ratios of later assembly proteins like L16 and L25 to early assembly proteins L3 and L4 were noticeably diminished in the aberrant particles. The standard by which a particular protein was judged as sub-stoichiometric with respect to L3 or L4 was a mean ratio to those proteins of 0.5 or lower. The sub-stoichiometric proteins, from greatest

amount detected to least, were L2, L14, L29, L18, L20, L16, L27 and L25. Not noted in figure 17 are proteins for which only one tryptic peptide was measured: L24, L32, L33 (sometimes referred to as L33a) and L36 (see Table 3). We did not detect L9, L28, L34 and L35.

The loss of L16, L27 and L25 is consistent with a ribosome trapped at a late stage of assembly. R-proteins L16, L18, L25, L27, L33, and L36 are all situated near the PTC and the central protuberance (see Fig. 18), which are thought to be the final regions assembled (Chen, 2013; Jomaa, 2013; Li, 2013). There were losses of some unexpected r-proteins, however. Ribosomal protein L20 is not considered a late assembly protein in the current assembly maps, and is situated in the interior of the central body of the ribosome. Its decrease is consistent with data previously generated on *E. coli* ribosomes assembled without L27 (Wower, 1998). However, the *E. coli* aberrant 50S subunits formed without L27 present in that study also failed to contain L21, which was not significantly absent from our aberrant 50S particles. L2 and L14 are situated near the interior of the ribosome and far away from the catalytic site. L20 is extremely interior to the ribosome and L29 is situated over the exit tunnel, one of the first regions thought to be assembled (Figure 18).
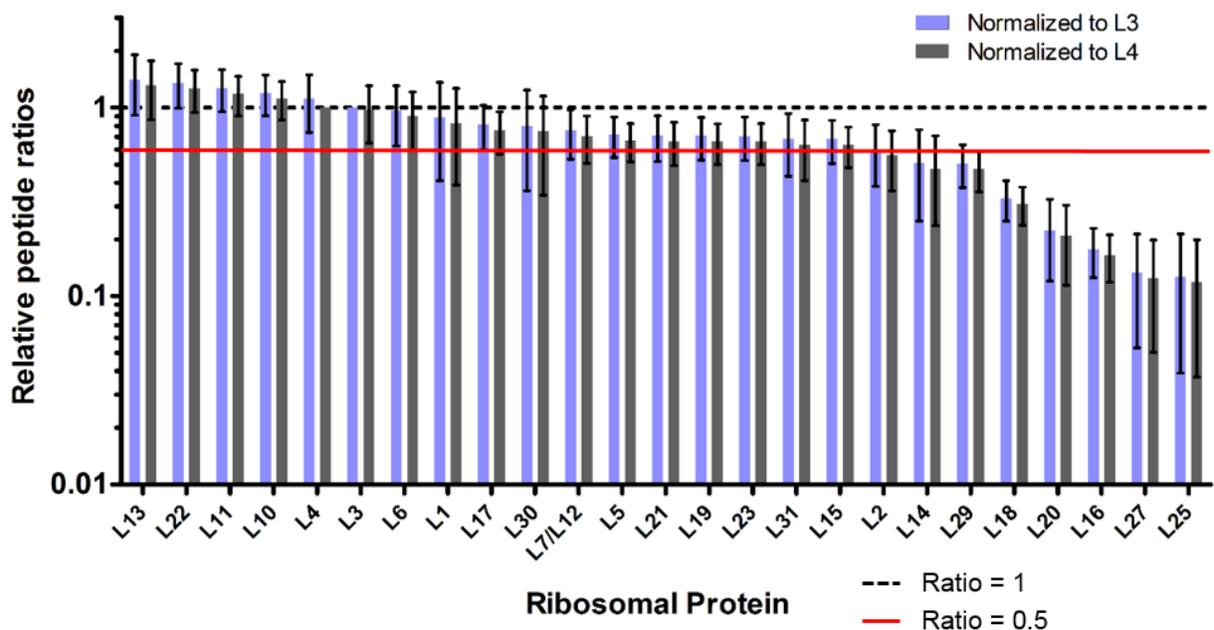
Figure 17. Ribosomal peptide ratios reflect stoichiometric differences in the r-proteins in aberrant vs wild type 50S subunits

Relative change in abundance of ribosomal proteins in aberrant particles. Precursor ion spectra for two tryptic peptides were assessed for each ribosomal protein represented. The absolute areas of each tryptic peptide were compared to those of an early assembly protein within their sample, e.g. the value of (area of peptide 1 from ribosomal protein x / area of ribosomal peptide 1 from L3) was calculated separately for wild type and aberrant particles. Those values were then compared (aberrant value / wild type value) to assess whether the aberrant value represented a significant change compared to that of the wild type. The black dotted line represents a ratio of 1, while the red continuous line represents a ratio of 0.5.

In this graph, the mean of four peptide comparisons for each ribosomal protein (six for L16) is depicted with error bars representing one standard deviation above and below the mean. Most values are extremely close, resulting in error bars that encompass the ratio of 1, displaying no significant change. The peptide values that have decreased the most are those of L2, L14, L18, L20, L16, L27 and L25. Not depicted: L24, L32, L33, and L36, for which only one representative peptide could be detected in both wild type and aberrant particle samples. Those comparisons can be found on Table 3. Not detected: L9, L28, L34 and L35.

Image adapted from Wall et al 2015; used with permission
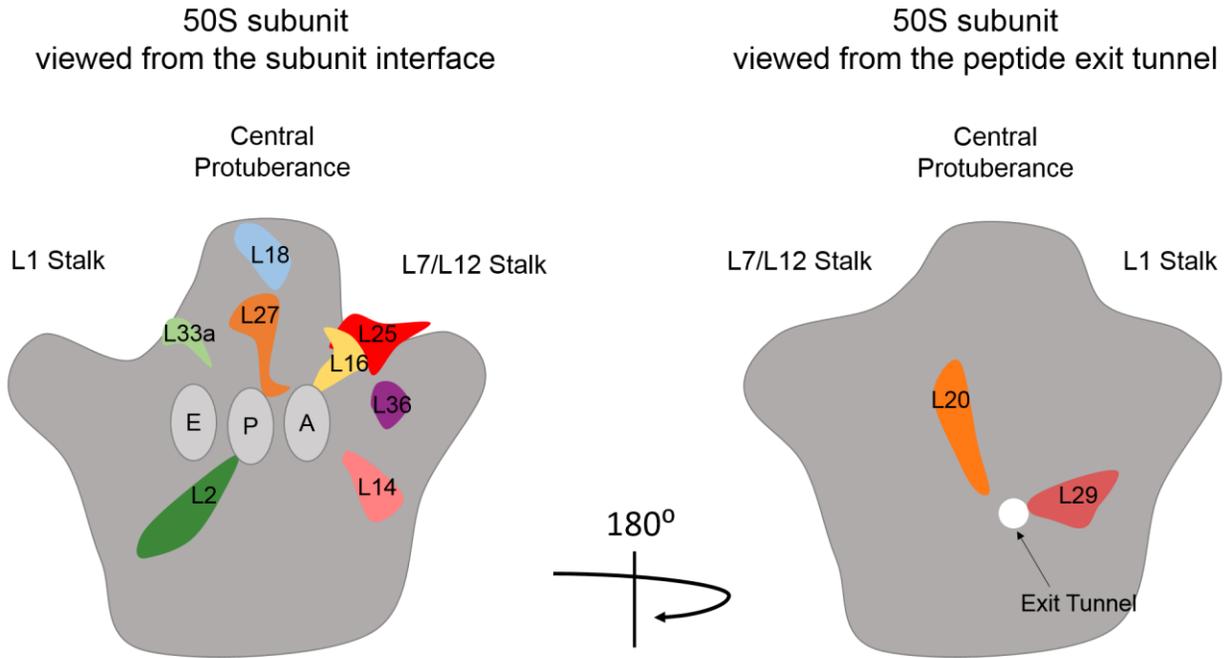
Figure 18. Approximate 2D locations of large subunit ribosomal proteins that were lost in conspicuous amounts in the presence of the Prp C34A mutant.

Taking into consideration the "semi-quantitative" nature of our data, r-proteins L2, L14, L16, L18, L20, L25, L27, L29, L33a and L36 were represented in MS1 peptide spectra as having their ratio to early proteins L3 or L4 decreased by at least two fold in the aberrant particle versus a wild type 50S particle.

These particles are most likely incomplete assemblies of the 50S subunit. Ribosomal proteins L9, L28, L34 and L35 were not detected in either sample. However, there are legitimate precision concerns with this type of analysis for ribosomal peptides. Conventional tryptic digest LC MS/MS is particularly problematic for small, highly negatively or positively charged proteins due to their distribution of tryptic residues. Only one peptide was found from each L24, L32, L33 and L36 (Table 3), so they are not included in Figure 17. There are very few predicted detectable tryptic peptides from ribosomal proteins L28, L34, L35 and the small negative protease Prp. Prp was detected in neither sample by mass spectrometry, but was detected in aberrant particles by immunoblot (see below and Fig. 19). There is only one measurable tryptic peptide produced in Prp predicted digestion; even in the mass spectrometry performed on the strong co-purified PrpC34A band (Fig. 9) Prp spectra were only detected in 3 scans, though of the proteins detected (mostly keratin contamination) it represented the only protein that could produce a band of that molecular weight.

Rabbit polyclonal antibodies were raised to both L27 and Prp to allow immunodetection by Western blot. L27 was detected in aberrant particles as an imbalanced doublet that appears to represent large amounts of un-cleaved L27 and only minor amounts of cleaved L27, while wild type particles only contained the smaller cleaved L27 band. In addition, Prp was present in the aberrant particles and absent in normal 50S subunits (Fig. 19). It seems that the bound complex of un-cleaved L27 and PrpC34A was assembled into many of the ribosomes in the strain overexpressing PrpC34A, which would explain their arrest in development.
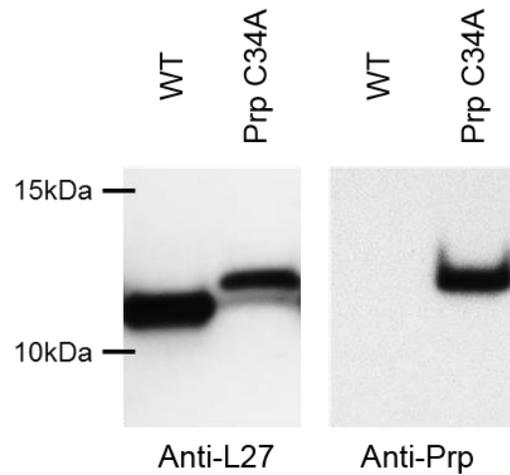
Figure 19. Prp and un-cleaved L27 are assembled into aberrant 50S particles

Detection of L27 and Prp in ribosomal particle fractions by Western blot. Aberrant and wild type 50S particle fractions were concentrated and run out on a 16.5% Tris-Tricine gel for maximal separation of the small proteins. The proteins were electroblotted onto PVDF and probed with custom antibodies to L27 and Prp. Both blots show wild type particles on the left (WT) compared to aberrant particles produced in the presence of PrpC34A on the right (PrpC34A).

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

Image reproduced with permission

Chapter 6

L27 Deletion and Complementation

The phenotype displayed when mutant Prp is expressed, as described in the previous chapter, was consistent with a role for Prp and L27 cleavage in ribosome assembly. Interpretation of those results is complicated, however, by the presence of wild type Prp and L27. In order to more accurately assess the roles for Prp and L27 cleavage, it was necessary to be able to examine the effects of L27 mutations directly. For this it was necessary to create a complementation system that would allow us to toggle between wild-type L27 and one of our mutants of interest. Further, we sought to determine whether separate expression of the post-cleavage peptide would have an effect.

Creation of a complementation system in *S. aureus* required a fairly complicated strain construction regime. Strain SA178RI carries a chromosomal copy of the gene for T7 RNA polymerase under an IPTG-inducible promoter. This strain was transformed with pEW27, a vector that expresses Sau L27 from a T7 promoter. The resulting strain, ST256, was subjected to allelic exchange to knock out the chromosomal copy of L27 using plasmid pEW68. A spectinomycin resistance cassette was substituted for chromosomal L27 while maintaining the cells in the presence of IPTG, allowing pEW27 to complement

via ectopic expression of L27. All candidates that had resolved pEW68 were also spectinomycin resistant and IPTG dependent. IPTG dependence indicated that L27 must be expressed from pEW27 to support growth. <u>This strain thus represents the first conclusive genetic evidence that L27 is essential in *S. aureus*</u>. This new L27 complementation/deletion strain was named ST360 (Figure 20).
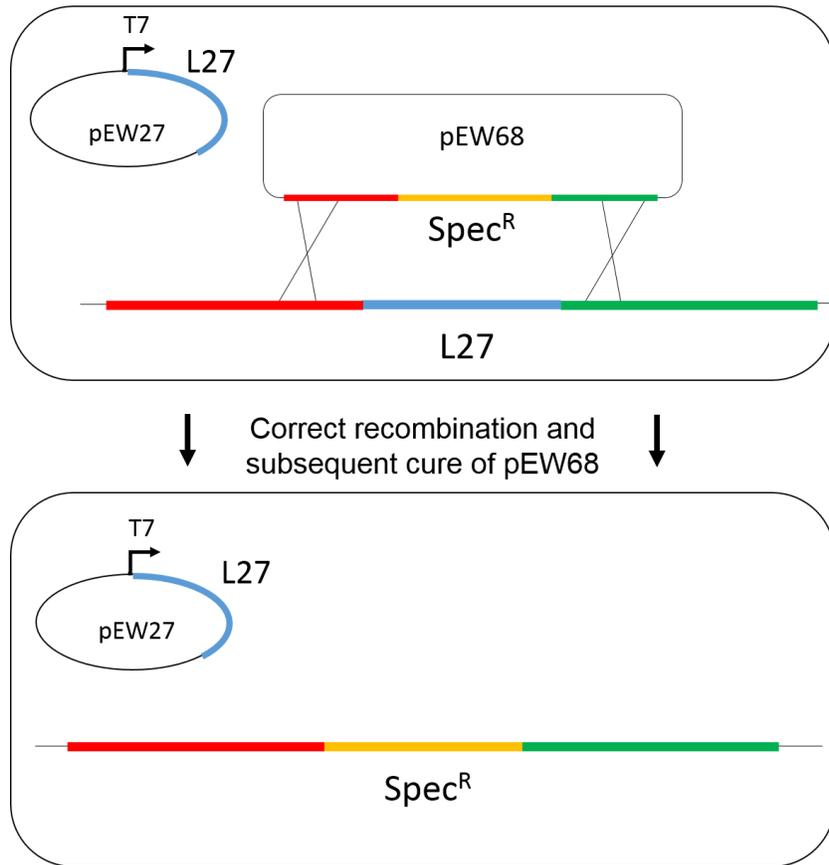
Figure 20. ST360 creation - L27 complementation with pEW27 allowed pEW68 to exchange chromosomal L27 with a spectinomycin resistance cassette

In order to replace chromosomal L27 with a spectinomycin resistance cassette, all cultures had to be grown in the presence of 1mM IPTG to induce expression of L27 from pEW27. This produced sufficient amounts of L27 so that chromosomal L27 could be replaced. ST360 is completely IPTG dependent, because it can no longer produce L27. This demonstrates definitively for the first time that L27 is essential in *S. aureus*.

A compatible plasmid containing a kanamycin resistance cassette and an arsenite-inducible promoter, pT104 (Liu, 2004), was introduced to ST360 as the vector for mutant L27 expression. We first performed a control experiment to show that the presence of arsenite alone will induce enough L27 from that promoter to allow growth. Plasmid pEW72 is a pT104 derivative expressing wild type Sau L27 and it is able to complement the L27 chromosomal deletion in the presence of 5 μM NaAsO4. There is no leaky expression from the promoter; in the absence of IPTG and arsenite, there is no growth of that strain. That was a great day. The pT104 backbone was used to construct plasmids pEW73-75, which expressed variant forms of L27 to test their ability to complement an L27 deletion (Fig. 21).

Plasmid pEW73 expresses "pre-cleaved" L27 (Δ 2-9) that is missing the N-terminal cleavage motif, making it effectively equal to the "short" Eco L27. We used Edman degradation on purified pre-cleaved L27 with a C-terminal His$_6$ tag (pEW76) to determine that this protein loses its N-terminal Met when produced in *S. aureus*, precluding any obstruction of the peptidyl transferase center. It is essentially identical to the studied *E. coli* or *T. thermophilus* N-terminus of L27 that has been visualized at the PTC in the ribosomes of those organisms (Polikanov, 2014).
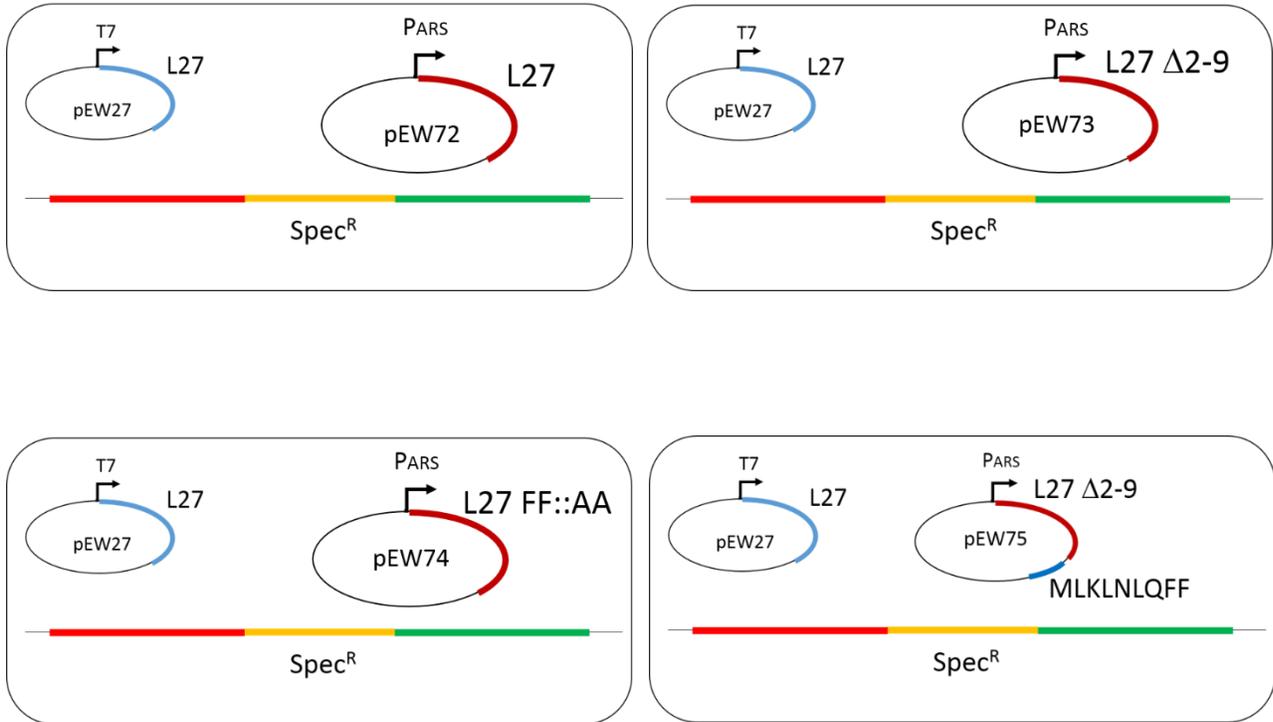
Figure 21. Transformation of ST360 with pT104-derived plasmids pEW72-75 allowed complementation analysis with mutant forms of L27

Each pT104-derived plasmid contains a tightly controlled arsenite-inducible promoter, pictured on the right in each cell with labeled P_{ARS} promoter. Plasmid pEW72 produces wild type L27 with its native cleavage motif (MLKLNLQFF'ASKK…). Plasmid pEW73 produces L27 without the cleavage motif, what we call "pre-cleaved" L27 (M'ASKK…). Plasmid pEW74 produces L27 with a motif that is not a Prp substrate (MLKLNLQAAASKK…). Plasmid pEW75 produces pre-cleaved L27 from one ORF, and then the post-cleavage peptide (MLKLNLQFF) from another ORF under the same promoter but with a separate RBS.

Plasmid pEW74 expresses Sau L27 F8A, F9A, an un-cleavable mutant of Sau L27. We confirmed that this protein with N-terminal residues MLKLNLQAAA was no longer cleaved by Prp via Western blot, as mentioned in the previous chapter and Figure 13. We also found that although Prp binds tightly to the normal L27 motif, we were not able to pull down native Prp when this tagged un-cleavable protein was expressed and purified from *S. aureus* (Wall, Johnson, Peterson and Christie, unpublished results). This seems to indicate that Prp can no longer bind the FF::AA mutant, suggesting an important role for the Phe residues in enzyme/substrate binding.

Plasmid pEW75 expresses pre-cleaved L27 and, under the same promoter but using a separate downstream RBS, the post-cleavage released peptide "MLKLNLQFF." Small peptides have been shown to have a role as hormones in *Enterococcus fecaelis*, a related bacterium that also encodes long L27. We sought to determine whether the peptide had an effect when separated from its original source or could compensate for pre-cleaved L27.

Of these plasmids, only pEW72 expressing wild type L27 could complement an L27 chromosomal deletion in ST360. All strains were spotted onto solid media containing 1mM IPTG, no inducer, or 5 μM NaAsO$_2$. All grew on IPTG (pEW27-dependent growth), all died without any inducer (no leakage from either promoter) and none but the strain containing pEW72 survived on arsenite (Fig. 22, bottom). This indicates that not only is L27 cleavage essential, it also must be tightly regulated. The peptide may yet have a role

in cellular processes, but that role is insufficient to promote *S. aureus* growth without regulated L27 cleavage.

Growth patterns of complementation strains

Growth curves were measured for the strains containing only IPTG-inducible L27 and those that also contained an arsenite-inducible alternative wild type L27, pre-cleaved L27 and un-cleavable L27, mentioned previously. The first growth curve (Fig. 23) is a comparison between cells supplied with IPTG to allow L27 production versus those depleted of L27 by washing cells and then inoculating them into media without IPTG. IPTG-induced L27 production allowed cells to complete at least six doublings in under nine hours. L27 depletion caused an arrest in growth after just over four doublings in ten hours. The continued growth during L27 depletion is likely due to the presence of previously assembled, functioning ribosomes in the initial inoculum. The first doubling results in half the residual functional ribosomes, one fourth in the second doubling, one eighth in the third doubling, one sixteenth in the fourth – this seems to be the point at which growth can no longer continue at anywhere near the normal rate.

Approach: substitute pre-cleaved and un-cleavable versions of L27 under a different inducible promoter and check for complementation by spotting onto media containing either inducer

P_{T7} L27

pEW27

1) IPTG-induced L27 production

L27::Spec^R

S. aureus chromosome

P_{ars} L27*

pEW72-75

2) L27 deletion using Spectinomycin resistance cassette

3) NaAsO$_2$-induced L27 mutant production

AsO$_2$-inducible L27 variants:

	IPTG Induction (pEW27)	Ø	AsO2 Induction (pEW72-5)
S. aureus L27 (MLKLNLQFF'A…)	●		○
Pre-Cleaved L27 (M'ASKK…)	●		
Un-Cleavable L27 (MLKLNLQAAASKK…)	○		
Pre-Cleaved L27 (M'ASKK…) + Peptide (MLKLNLQFF)	●		

Figure 22. Complementation schematic; only wild type L27 can complement the chromosomal L27 deletion

Top – outline of the complementation system created in SA178RI for this study.

Bottom – Growth of cells in the presence of different L27 mutants. On the left there is a description of the L27 allele carried on the arsenite-inducible plasmid, forming horizontal rows in this array. On the right, the three columns are photographs taken of solid culture media containing 1 mM IPTG, no inducer, or 5 µM NaAsO$_2$, pictured in that order. In each row the same amount of each described strain has been spotted onto the different media. Only the strain containing wild type L27 under P$_{ARS}$ can grow on arsenite media.

The arsenite-inducible complementation strains were measured after induction of the mutant L27 allele. In Figure 24, arsenite-induced wild type L27 completed eight doublings in 24 hours, with six doublings occurring in the first ten hours. This is slightly slower than cells supplied IPTG, almost certainly because the T7 promoter under IPTG control produces more L27 than does the arsenite inducible promoter. However, cells supplied with pre-cleaved L27 or un-cleavable L27 from the arsenite promoter managed to double only five times in 24 hours, with the first four doublings taking place in about seven to eight hours, and the fifth doubling taking place around eighteen hours after inoculation. The growth pattern displayed in conditions of L27 deprivation and in the presence of pre-cleaved or un-cleavable L27 expression was nearly identical. These results suggest that supplying pre-cleaved or un-cleavable L27 does not result in the production of new active ribosomes, as is most probably the case with L27 depletion from the cell.
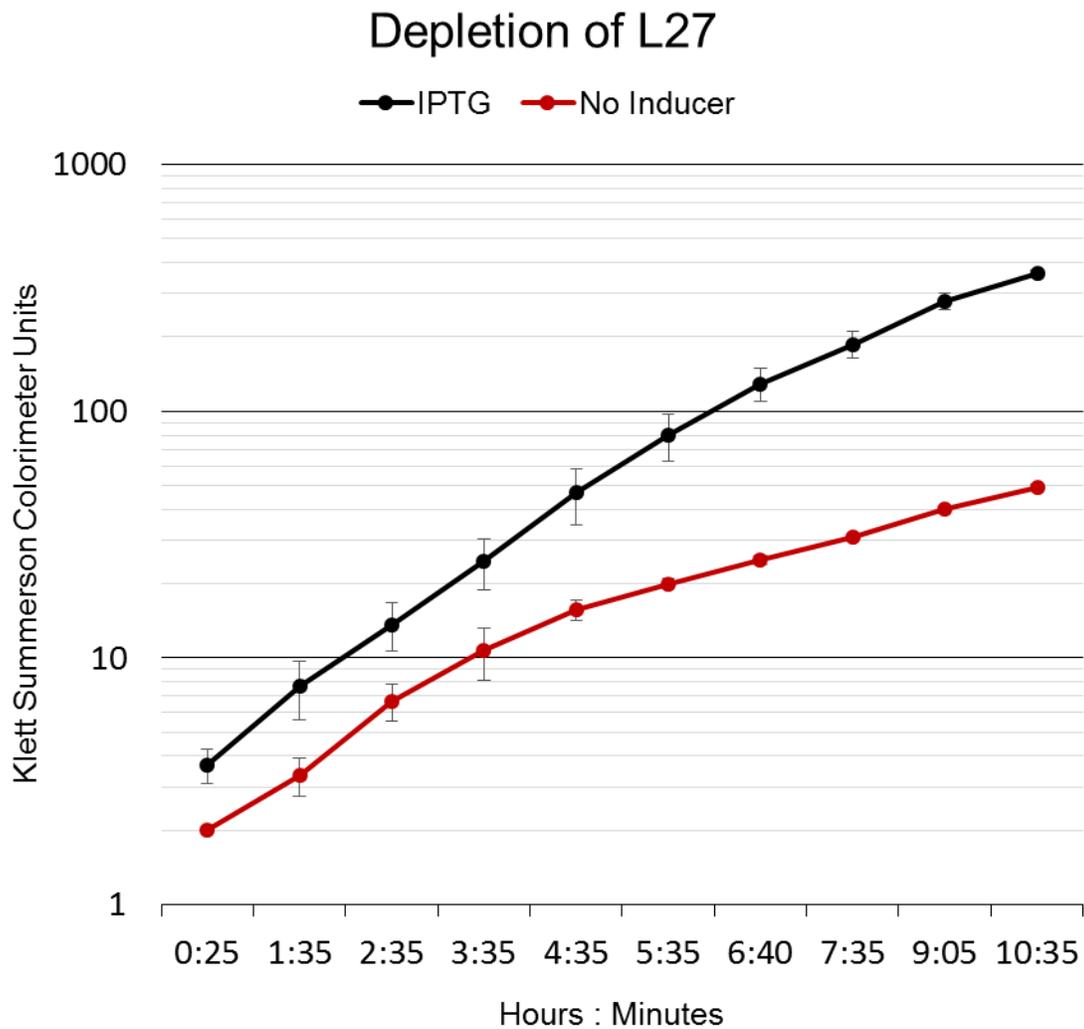
Figure 23. L27 depletion results in a severe growth deficit

Cells were first grown to Klett=30 in IPTG media. A 0.5 mL aliquot of the exponential culture was harvested and washed twice with media that did not contain IPTG. These cells were resuspended in 0.5 mL BHI broth and then inoculated into 9.5 mL BHI with or without IPTG. These cultures were shaken at 200 rpm, 37° C and were measured at hour time points for ten hours. The final sample measurements are above K=300 in the presence of IPTG, and around K=50 without IPTG. Error bars represent one standard deviation above and below the mean, n=3.
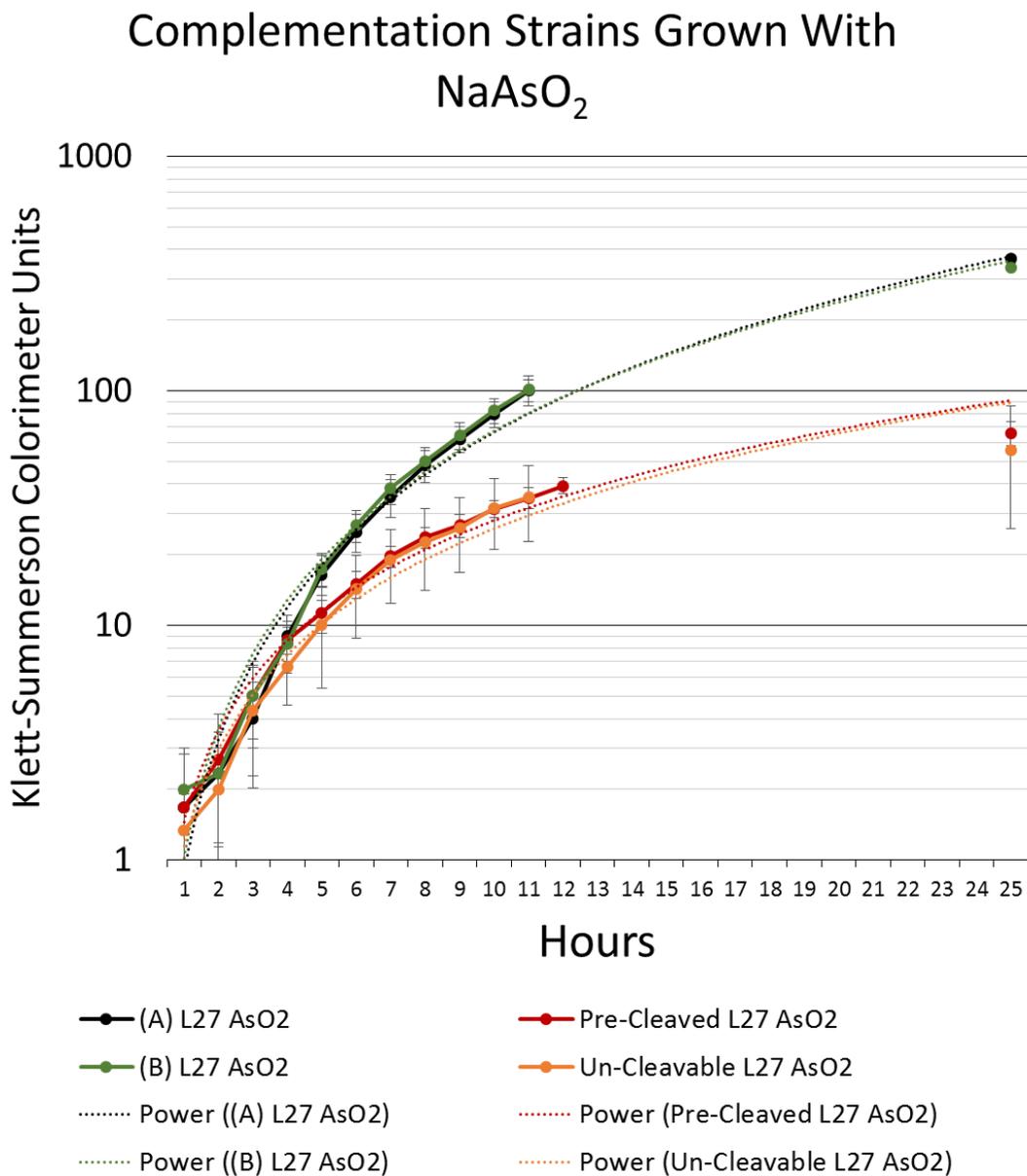
Figure 24. Growth curves of complementation strains grown in arsenite with trendlines

Cells were first grown to Klett=30 in IPTG media. A 0.5 mL aliquot of the exponential culture was harvested and washed twice with media that did not contain IPTG. These cells were resuspended in BHI broth containing 5 μM arsenite (AsO2) and then inoculated into 9.5 mL arsenite BHI broth. These cultures were shaken at 200 rpm, 37° C and were measured at hour time points for 12-13 hours, and then one measurement was taken at 24 hours. The "Power" trendlines represent the best-fit exponential equation for the data sets of each experiment.

Error bars represent one standard deviation above and below the mean, n=3.

Chapter 7

Molecular Modeling and Mechanistic Predictions

Existing structures of Prp family members demonstrate that Prp forms its own completely new clan of cysteine proteases, dubbed clan CR by the MEROPS database which catalogs and classifies peptidases, their substrates and their inhibitors. Within clan CR, *Staphylococcus aureus* Prp is now the prototype member of Family C.108 (cysteine proteases, group number 108). It is as yet unknown how Prp binds its substrate, so loop refinement, docking and molecular dynamics were undertaken to explore possible models for Prp-substrate interaction.

Structural modeling through protein sequence homology has become an extremely important tool in deciphering hitherto unknown functions in conserved domains. This modeling attempt was relatively conservative, only modeling two small loops onto an existing crystal structure, docking a short peptide and then subjecting that modeled complex to a short round of molecular dynamics simulation. The mechanism and necessary requirements known for cysteine proteases informed the modeling attempt. As discussed in Chapter 4, cysteine proteases are hydrolases that use nucleophilic attack on the carbonyl carbon in the scissile bond. This is achieved via a catalytic dyad, His and Cys, or a catalytic triad, His, Cys and Asp or in some cases Asn. There is a residue in the

Prp motif that is usually Asp or Asn, three residues prior to the catalytic Cys, but the role of that residue in catalysis is not yet known. For the purpose of this modeling experiment we assumed that Prp relies on the catalytic dyad because those residues are perfectly conserved, while the other residue differs.

In the mechanism of catalysis for a cysteine protease with a catalytic dyad (Fig. 25), the imidazole group of the His polarizes the sulfhydryl of the Cys and abstracts a proton, producing the ionic pair imidazolium and thiolate. The thiolate ion attacks the carbonyl carbon of the substrate scissile bond, producing a tetrahedral intermediate with an oxyanion. The nitrogen in the tetrahedral intermediate then abstracts a proton from the imidazolium moiety. This breaks the bond between the nitrogen and the enzyme-substrate complex, forming an amino terminus, which allows that peptide chain to dissociate. The carbon-oxygen double bond is restored and a covalent acyl-enzyme intermediate is formed on the catalytic Cys. A water molecule attacks the carbonyl carbon of the thioester bond, which forms another tetrahedral intermediate containing an oxyanion. This intermediate is hydrolyzed and the C-terminus of the substrate is released as a free acid. The formation of an oxyanion during the tetrahedral intermediate stage occurs twice in this mechanism, which typically requires stabilization via a pocket in the enzyme with available amide nitrogens as hydrogen bonding partners.
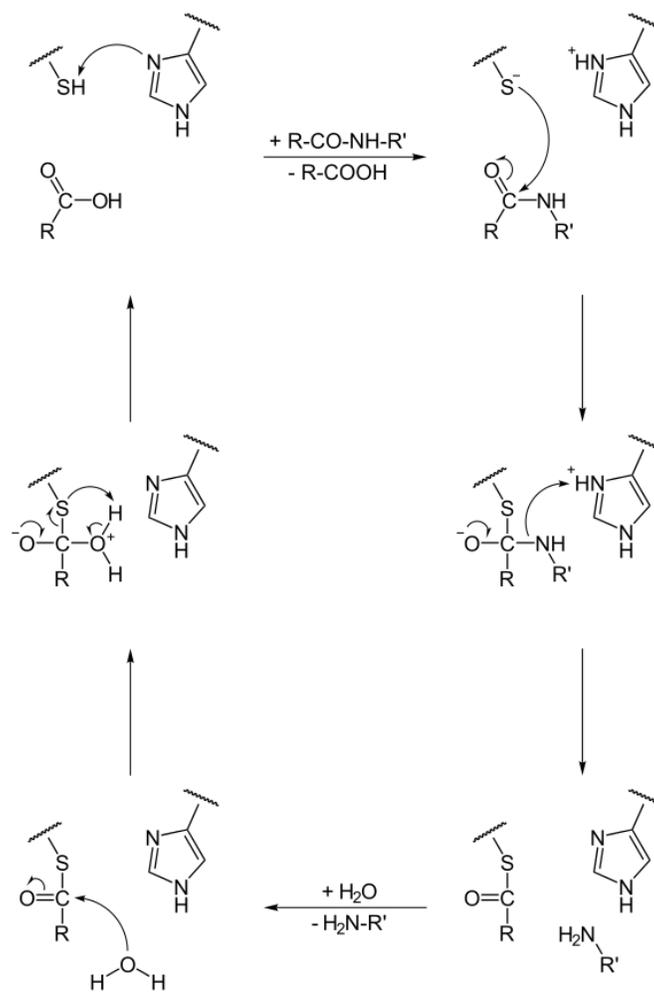
Figure 25. Mechanism of a cysteine protease containing a His/Cys catalytic dyad

Cysteine proteases make use of the fact that thiolate is a strong nucleophile, and that the imidazole group of His can readily abstract a proton from the Cys sulfhydryl group to produce a thiolate moiety. This mechanism involves the formation of an oxyanion tetrahedral intermediate, a covalent acyl-enzyme complex, and then another oxyanion tetrahedral group before hydrolysis of the covalent substrate-enzyme complex is achieved.

Public domain image from Wikipedia (http://en.wikipedia.org/wiki/Cysteine_protease).

From this mechanism we know that the sidechains of the catalytic dyad must be in close enough proximity to achieve ionization, and that a required feature of the enzyme would be a region that could stabilize the oxyanion of the tetrahedral intermediate.

Modeling began by examining existing structures of Prp in the PDB. These include homologs from *Thermatoga maritima* (1S12), *Streptococcus pneumoniae* (2IDL), *Streptococcus mutans* (2G0I) and *Staphylococcus aureus* (2P92). They all share the unique two layer α/β sandwich dimeric structure of pseudo twofold rotational symmetry. It is not yet known whether the dimer could contain strict twofold rotational symmetry because the flexible loop conformations are likely quite variable in solution. Each monomer is composed of two helices in contact with one side of a five stranded β-sheet. Peptide chain topology from N to C terminus is $\beta\beta\alpha\beta\beta\alpha\beta$. The fold conservation is remarkable, given average sequence identities of 20-30 percent. Figure 26 shows conserved residues mapped onto the 2IDL backbone.

This fold was completely novel in 2005 when the *T. maritima* structure was solved, so much so that even though the authors conjectured that the conserved residues formed a catalytic center, the closest structural homologs they could identify were a hydratase and a gene of unknown function from yeast, leading them to pass on functional speculation (Shin, 2005). The *Streptococcus* structures were solved by crystallographic consortia and remain unpublished. By the time of writing, the paper about *S. aureus* 2P92 had just been published, in which the authors speculate that because L21, Prp and L27

exist in an operon, it would make sense for them to all form a complex related to ribosome assembly, with Prp as the acidic "glue" between alkaline L21 and L27. They demonstrated that Prp is a dimer in solution, validating our identical finding (Wall, Johnson, Peterson and Christie, in prep). They did not discern that the protein contained the catalytic center characteristic of a cysteine protease, even though they had literally thousands of sequences at their disposal and did not discuss the absence in the crystal structure of a 10 amino acid loop that fell in between the completely conserved His and Cys residues (Chirgadze, 2015).

A striking feature of these structures is their inferred lack of catalytic capacity. The catalytic dyad must at some point be solvent exposed to accept substrate and the His and Cys must face each other in such a way that the His can abstract a proton from the sulfhydryl group of the Cys. No structure available has these features. Though the His and Cys are sometimes in close proximity in the structures, as in 2IDL (Fig. 26), the rotation of their sidechains is influenced by the flexible loop that lies between them, orienting them in opposite directions. This might have been the confounding factor in the failure to identify this protein as a cysteine protease.

Molecular Modeling

The PDB file of the incomplete *Staphylococcus aureus* Prp dimer (2P92) was analyzed in SYBYL-X 2.1, a graphical interface program for molecular modeling. It was immediately apparent that one of the catalytic residues (His 22) was missing from an

active site on chain A, and that the orientation of the existing His 22 imidazole group would not be conducive to catalysis as it does not face the sulfhydryl group of catalytic Cys 34. For this reason, chain B and all the co-crystallizing waters were deleted from 2P92 so that modeling could be performed on monomer chain A. The incomplete *S. aureus* Prp structure 2P92 Cys 34 was cycled through a library of rotamers and a conformation was chosen in which the sulfhydryl group was solvent-exposed. From this point, the loop between Gly 21 and Ile 32 could be modeled in such a way that the chain conformation would reflect the new pro-catalytic conformation of Cys 34. One loop model of the appropriate residues was built on this structure using a Modeller 9.12 (Sali, 1995) script written by Dr. Hardik Parikh (see Methods). The model now included all the residues present in chain A of *S. aureus* Prp, without chain interruption. This structure file was then subjected to a loop refining script written by Dr. Parikh using Modeller to explore different possible conformations for the loop missing from the crystal structure (Gly 21 to Ile 32). 50 models were created and checked for pro-catalytic conformations. A plausible active model, number 12, was chosen based on the conformation of the relevant atoms of its catalytic residues (Cys: sulfur 254 and His: nitrogen 159 or nitrogen 161). In model 12, these atoms were only about 4 Å apart. This distance would allow hydrogen bond formation and eventually proton abstraction from Cys 34, which is mandatory for catalysis in the mechanism of most cysteine proteases. In addition, this model gave the most open conformation of the upper loop, allowing greater solvent access to the catalytic center for the purpose of docking.

.

Figure 26. Existing Prp structures share conserved residues and a non-catalytic conformation of the active site

Highly conserved residues according to the sequence alignment (MUSCLE using BLOSUM62 cost matrix.) are labeled and mapped in red onto the grey ribbon backbone (top left) and grey cpk depiction of *S. pneumoniae* Prp dimer (PDB ID 2IDL, top right). This structure represents the entire protein without sequence gaps. The catalytic pair His 22 and Cys 34 are shown in atomic representation on both models, colored by atom type. They are oriented in a non-catalytic conformation (see ribbon structure) and obviously not solvent-exposed (see cpk structure). Several of the conserved residues appear to be involved in substrate stabilization near to the catalytic center.

Accession numbers *S. aureus* (ABQ49493) PDB ID 2P92, *S. mutans* (WP_024782247) PDB ID 2G0I, *S. pneumoniae* (NP_358607) PDB ID 2IDL, *T. maritima* (NP_229256) PDB ID 1S12.

- 81 -

Model 12 was dimerized using 2P92 as a template and docking was attempted in Gold 5.2, but it became apparent that a four residue loop (residues 60-63), which would likely be mobile in solution, had a closed conformation that partially obscured the active site of the enzyme. For this reason, the loop refine script was again utilized in Modeller to allow the four residue loop to adopt a conformation that did not obscure the active site. 15 models were generated and the one that allowed greatest solvent exposure of the catalytic site was chosen.

This refined model was dimerized in SYBYL because the crystal structure depicts a dimer and Prp has been found to dimerize in solution (Wall, Johnson, Christie and Peterson, in prep). The sulfhydryl hydrogen of the active site cysteine on chain A was replaced with a lone pair in SYBYL and the active site histidine was protonated for the purpose of docking (Figure 27). The dimer model was read into Gold, and docking was attempted with the blocked peptide Acetyl (Ac)-Phe-Phe-Ala-Ser-Amide (Am), which represents the P2-P2' substrate residues as they would appear within a peptide. Docking was achieved using a distance constraint of 3-2 Å from the catalytic sulfur to the carbonyl carbon of the peptide bond between Phe and Ala. This distance was used to adequately represent the beginning of the nucleophilic attack that would form the covalent intermediate between enzyme and substrate. The sidechains of the peptide substrate fit well into the hydrophobic pockets surrounding the enzyme catalytic center, so this substrate conformation was used as a scaffold constraint for the docking of a longer portion of the cleavage motif, Ac-Asn-Leu-Gln-Phe-Phe-Ala-Ser-Am. This was the largest peptide that was docked due to accuracy constraints on free peptide length. Residues

Met 1, Leu 2, Lys 3 and Leu 4 from the full L27 cleavage motif do not appear in the model, though Leu 4 is quite conserved.

The overall docked peptide seemed to fit well, but the target scissile peptide bond in this model was angled such that the carbonyl carbon to be attacked was not in direct alignment with the thiolate nucleophile on the enzyme. Because this enzyme-substrate model was not perfectly positioned for catalysis, molecular dynamics simulations were used to "shake" the docked model into a more realistic catalytic conformation.

Molecular dynamics (MD) simulations were performed in NAMD (Phillips, 2005). Model 12_4_6 dimer with docked substrate Ac-NLQFFAS-Am was solvated with water and 50 mM NaCl ions, and subsequently 25,000 frames were collected for a total of 200 ns at 37ºC. Frames from the early portion of this simulation showed improved steric interactions between enzyme and docked substrate – the MD simulation readjusted the target carbonyl so that nucleophilic attack from the enzyme thiolate could be possible. Frames from this re-positioned substrate enzyme complex were captured and are depicted in Figure 28. As the MD simulation continued, the docked substrate began to loosen from the enzyme starting at its N-terminus, which seemed to be less tightly bound than the FAS region that was buried deeper in the enzyme active site. It may be that the earlier residues in the cleavage motif (MLKL) have a large role in stabilization by binding to other regions of Prp. Further modeling work or structural determination is necessary to provide a clearer picture of how Prp binds its substrate.
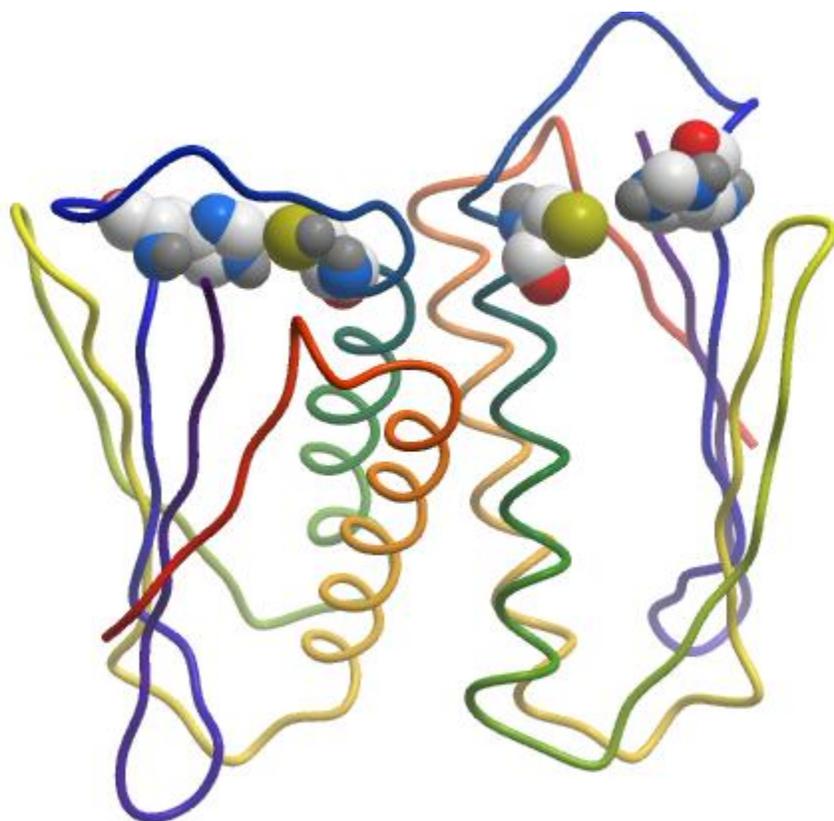
Figure 27. Ribbon model of *S. aureus* Prp with active site residues in catalytic conformation

The missing residues of PDB ID 2P92 are modeled in a loop that arcs over the catalytic center. This right active site in this model contains the pro-catalytic ionization of active site Cys and His residues, while the left active site residues are in pre-ionization conditions (both shown in atomic representation with cpk coloration). Ribbon coloration is blue (N-terminus) to red (C-terminus). The catalytic dyads are now oriented correctly and solvent exposed so that the enzyme/substrate reaction that begins the protease mechanism could be possible.
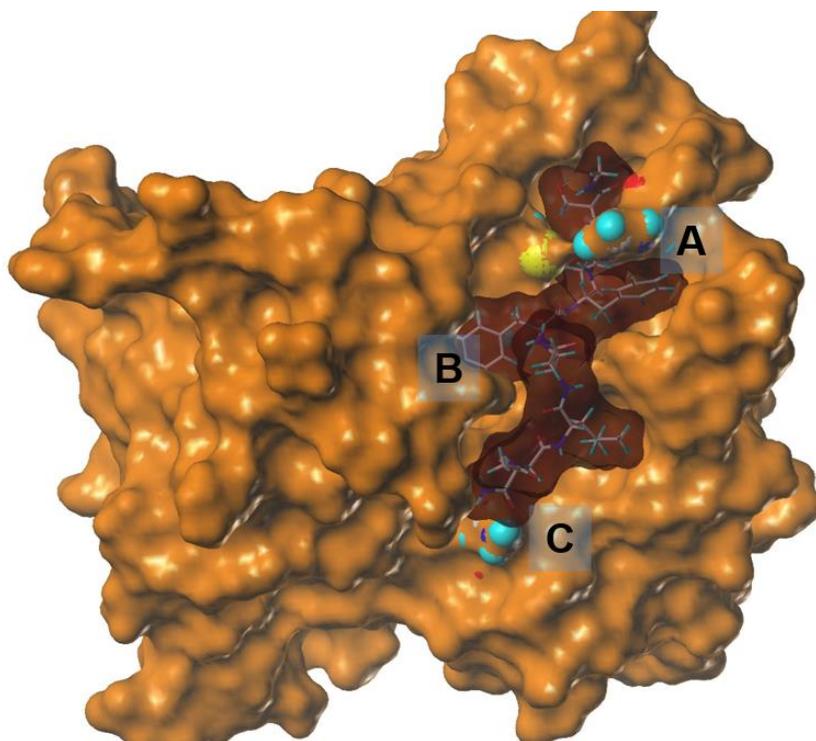
A) Possible π-π stacking interaction between substrate P1 Phe and catalytic His22 of protease

B) Protease dimer interface seems to form a hydrophobic pocket to accommodate P2 Phe sidechain

C) Possible hydrogen bonding interaction between substrate P5 Asn sidechain amide and that of Asn46 of protease

D) Conserved Ser38 hydroxyl in protease seems to hydrogen bond with substrate amide backbone, perhaps supporting the split conformation of the P2 and P1 Phe sidechains
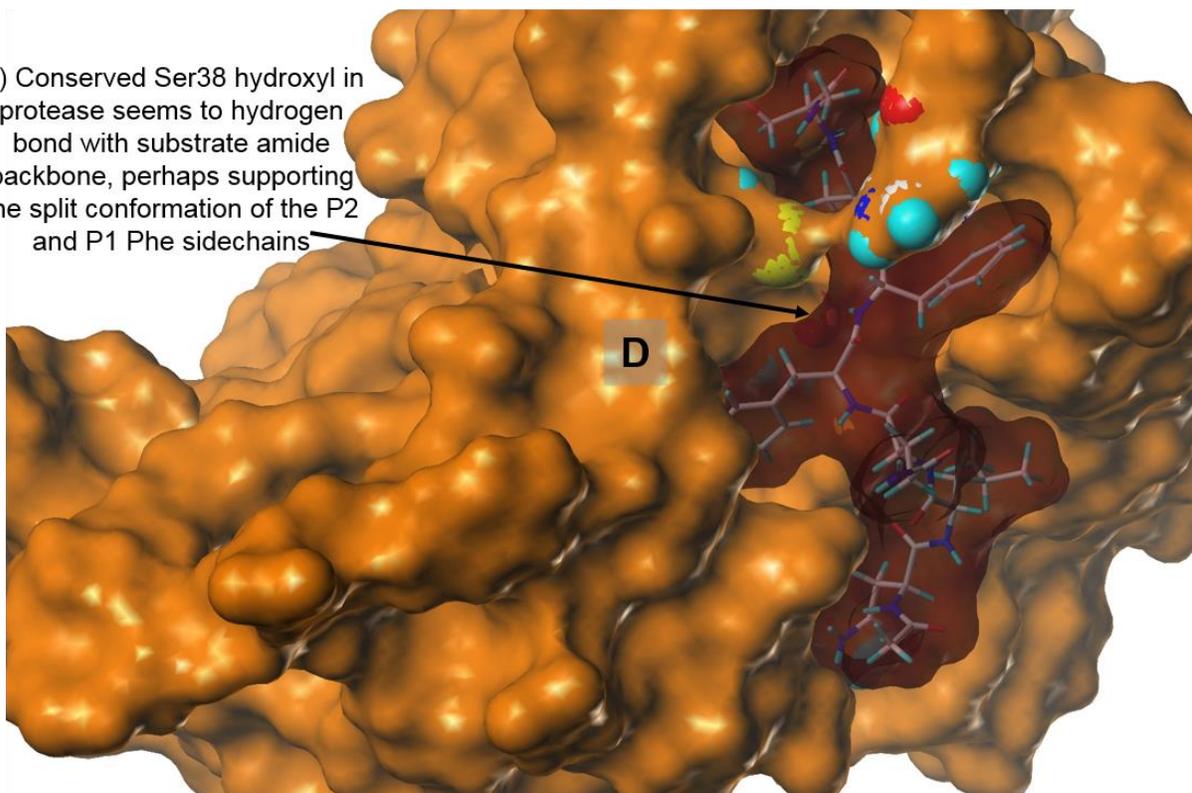
Figure 28. Surface model of *S. aureus* Prp docked with peptide Ac-NLQFFAS-Am

Model Analysis

This model seems to predict a pi-pi stacking interaction of catalytic His 22 on the enzyme with the P1 Phe on the substrate (Fig. 28, 29). This is interesting because the intrinsic mobility of His 22, derived from its co-conservation with preceding Gly 21, would be problematic for the maintenance of an active catalytic center. If the substrate P1 Phe indeed stacks with His 22, this would probably confer stability to the active conformation of the enzyme in the presence of its appropriate substrate. In agreement with this hypothesis, the overwhelming majority of known L27 cleavage motifs have Phe as the P1 residue. Rare exceptions exist, however - Leu in *C. difficile* (cleavage motif MNLQLL'A), and Cys in *Clostridium acetobutylicum* (extremely variant cleavage motif of INLSLC'A). Indeed, the first member of this enzyme family that was identified as an active phage prohead protease, Prp in Streptococcus phage Cp-1, is published as cleaving an unrecognizable motif PVLEGARINH'A, which still has a ring moiety as the P1 sidechain. [NB: It seems possible that this cleavage site identification is a mistake - slightly downstream the sequence DFQKH'A seems a more likely candidate.] However, the sequence of Prp in phage Cp-1 is extremely divergent from consensus. These disparities probably cannot be explained until there are more Prp structures from phage and bacteria with variant cleavage motifs.

The sidechain of the conserved Ser 38 residue on Prp seems to be a hydrogen bonding partner for the amide nitrogen between P2 Phe and P1 Phe in the substrate backbone. This residue is universally conserved among all Prps as a Ser or Thr spaced

four residues after the catalytic Cys, making it the next rung on the alpha helix that faces the catalytic center. The completely conserved hydroxyl moiety seems to orient the amide nitrogen of the substrate peptide backbone in a way that promotes the split between the P2 and P1 Phe aromatic sidechains.

Two residues directly below the small loop on the right opposite the catalytic site seem to have quite an important role. This loop contains a highly conserved glycine at position 65, which is adjacent to Gly 21, the completely conserved hinge that seems to promote His 22 movement. These two flexible glycines and the amide following Gly 65 create a fluctuating pocket that appears important for oxyanion stabilization during the tetrahedral intermediate stages of the mechanism of this protease (Fig. 30). The amide hydrogens would provide excellent attraction and support to the oxyanion to both stabilize its position and minimize its chance of erroneous bond formation. Further, they lie in a direct line passing through the substrate opposite but level with the catalytic Cys 34, positioning them perfectly to cradle the oxyanion after the linear nucleophilic attack. The fact that the oxyanion pocket seems so flexible is likely in keeping with the sequence specificity of this enzyme. In Figure 30, the amide hydrogens are not all in a perfect cradling position relative to the substrate, but this could be an artifact of the original structure having been crystallized in the absence of ligand. Preliminary NMR studies have found evidence of massive structural rearrangements that occur upon substrate binding (Wall, Scarsdale, Peterson and Christie, un-published).
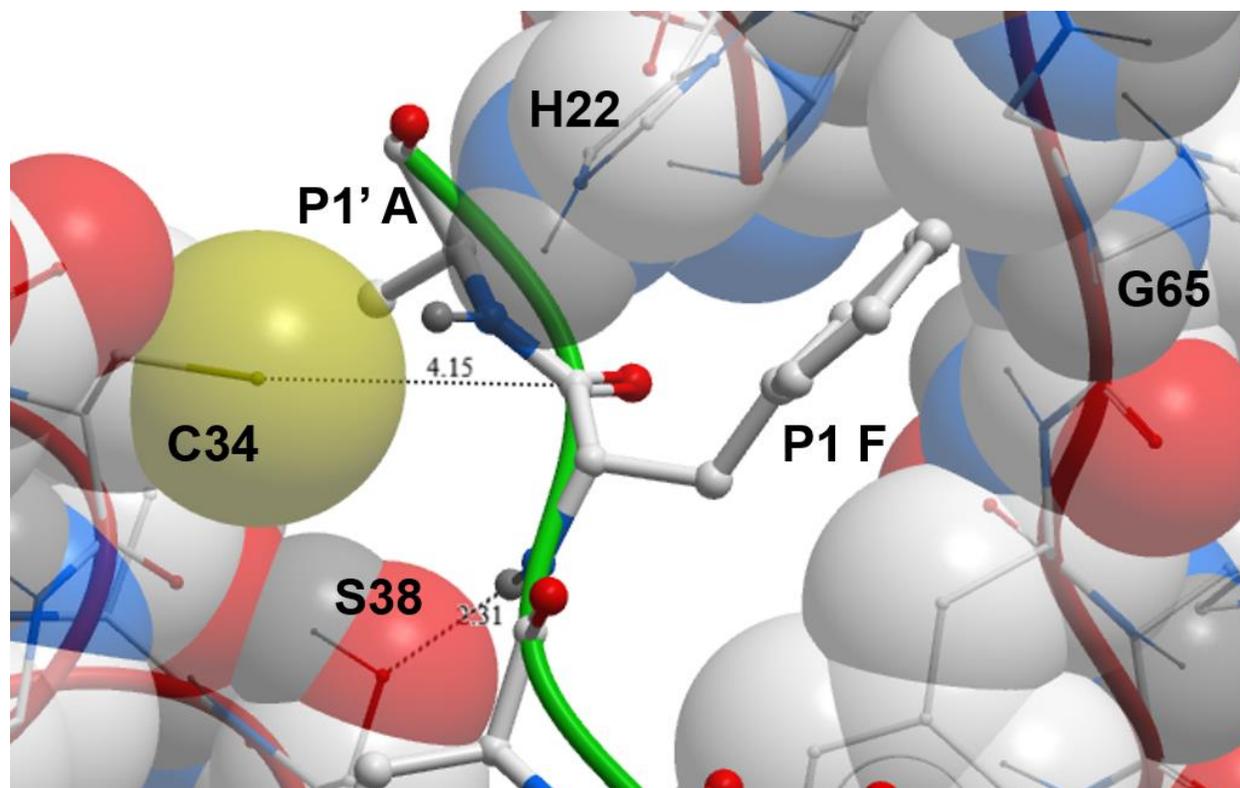
Figure 29. Nucleophilic attack by Cys34 and substrate stabilization by Ser38

The substrate is depicted with a green ribbon backbone, while Prp is modeled in semi-transparent cpk representation with a red ribbon backbone. The thiolate ion of Cys 34 attacks the carbonyl carbon of the scissile bond between P1 Phe and P1' Ala in the substrate. The completely conserved hydroxyl one helix turn below the catalytic Cys 34, Ser 38 in this model, appears to stabilize the amide nitrogen between the P2 Phe and P1 Phe by hydrogen bonding. This seems to allow the P1 Phe to form π-π stacking interactions with catalytic His 22.
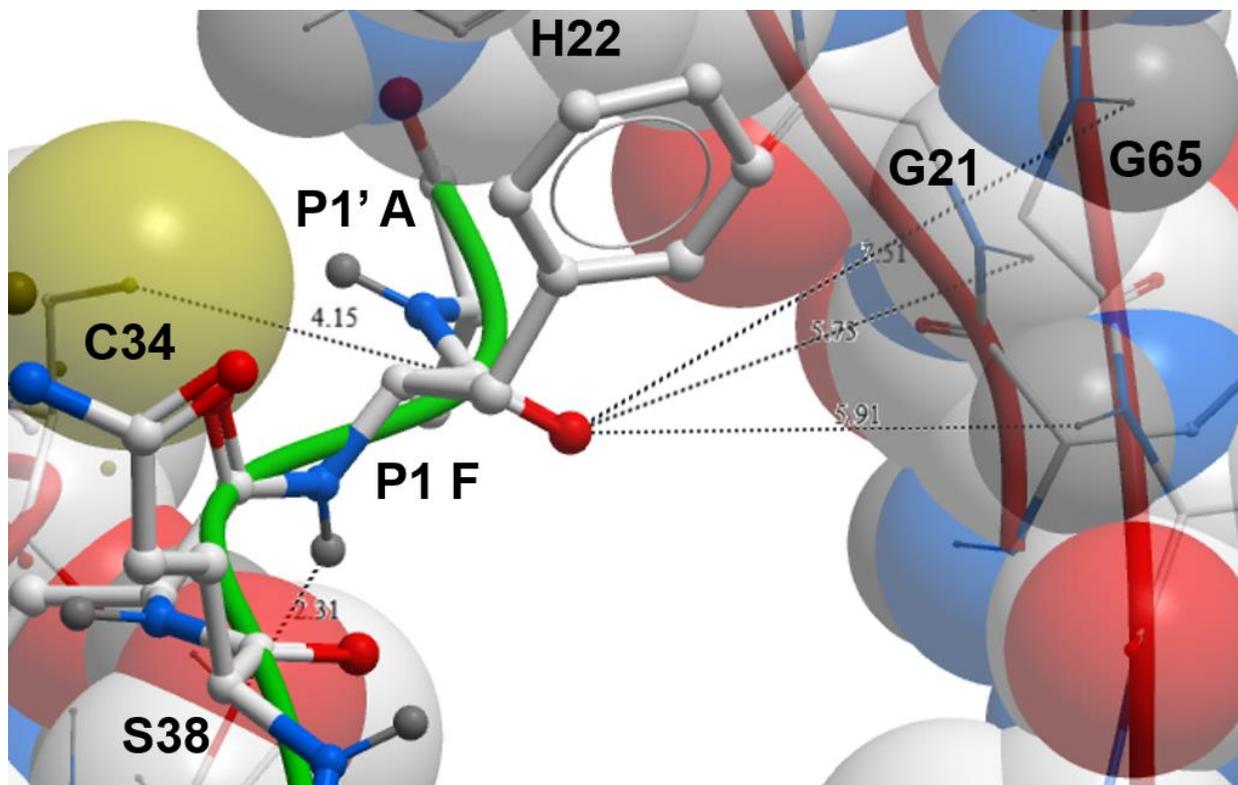
Figure 30. Putative oxyanion hole formed by amides of Gly 21, Gly 65 and the non-conserved subsequent residue His 66

Though they are not oriented perfectly in this model, the amide nitrogens of Gly 21, Gly 65 and His 66 (not residue specific) likely flex in solution to support the oxyanion formed during Prp catalysis. Gly 21 and Gly 65 are completely conserved, indicating intrinsic flexibility necessary in this region. These two residues lie in the same plane as the nucleophilic attack, making their location ideal to receive the reactive oxyanion moiety.

This model also appears to reveal the purpose of a previously unexplained Asn 46 conserved in the Firmicute Prps about 10 Å away from the catalytic site. This Asn would be an ideal hydrogen bonding partner for the conserved P5 Asn in the substrate cleavage motif. Preliminary work showed that Prp mutant N46A was significantly reduced in its ability to co-purify with tagged L27, suggesting a deficiency in binding. This Prp single mutant could still cleave its substrate, but kinetic studies should be performed to see whether the K_m of this mutant enzyme is significantly shifted and inhibiting catalysis.

The P4 Leu in the cleavage motif fit well into a large hydrophobic patch on the exterior of the Prp subunit that was solvent-exposed in the crystal structure. It also appears that the initial residues of the N-terminal cleavage motif (MLKL) would continue this pattern, filling in other hydrophobic patches on the Prp exterior that extend to the bottom rear of the other monomer.

Model verification via Prp assay

In order to test the predictions that our model makes regarding residues important to substrate binding and catalysis and ultimately to search for inhibitors, it was necessary to design a high throughput-capable continuous fluorescent assay. This assay monitors Prp activity in the presence of a quenched fluorescent peptide containing the cleavage motif. In this type of assay, when the substrate is cleaved the fluorophore is released from its quenching partner and fluorescence readings at the fluorophore emission wavelength

increase with time. There can be low level stable background fluorescence, but the majority of fluorescence should only be evolved in the presence of active Prp.

I designed the Prp assay and completed preliminary tests that demonstrate that fluorescence is only produced in the presence of active Prp. The amino acid sequence of the substrate was selected to maximize its hydrophilic residue content, aiming for the greatest possible peptide solubility. The peptide is stable and did not self-cleave or change fluorescence over time when monitored in solution. The assay allows us to determine Prp kinetics, and the effects of any mutants on the Km or Kcat.

Prp was purified as an N-terminal His$_6$-SUMO fusion protein and checked for activity using this assay. Fluorescence increase only occurred in the presence of wild type Prp fusion proteins. Mutation of the catalytic Cys 34 to Ala renders Prp incapable of catalysis (see Fig. 8A and Fig. 8B). In the presence of Prp C34A, no fluorescence increase was observed (Fig. 32). We also tested a Ser mutant of C34, in an attempt to discern whether the presence of a conserved Asp/Asn in the Prp catalytic motif suggested that the protease contained a catalytic triad. Though this mutant had been tested previously via co-expression with tagged L27 and Western blotting and found to be inactive (Fig. 8B), we wanted to determine whether it could demonstrate any activity in vitro. The C34S active site mutant was also incapable of catalysis, indicating that the highly conserved Asp residue (Asp 31) is not sufficient to constitute a His-Ser-Asp catalytic triad in this protease. Of note, Prp C34A and C34S are both known to bind substrate tightly without cleaving, which demonstrates that binding is probably not

sufficient to increase substrate fluorescence. This increases confidence in the validity of the Prp assay.

We are in the process of fine-tuning the Prp assay. A Prp kinetics profile will be prepared and then the effects of various mutations on Prp kinetics will be tested. Mutants have been made in residues that are either highly conserved or predicted to be important by their conservation and by their positions in the model (Wall, Johnson, Peterson and Christie, unpublished work).
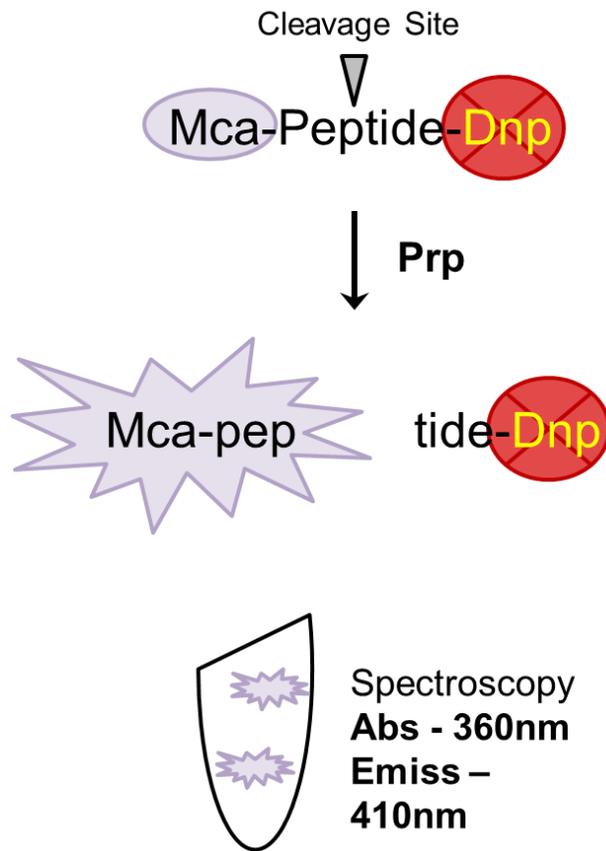
Figure 31. Quenched fluorescent peptide assay mechanism

A properly designed peptide substrate with a fluorophore on one terminus and a quencher on the other will not fluoresce until it is cleaved. Monitoring fluorescence evolution at the appropriate wavelength allows performance of a continuous assay, useful in determining enzyme kinetics.
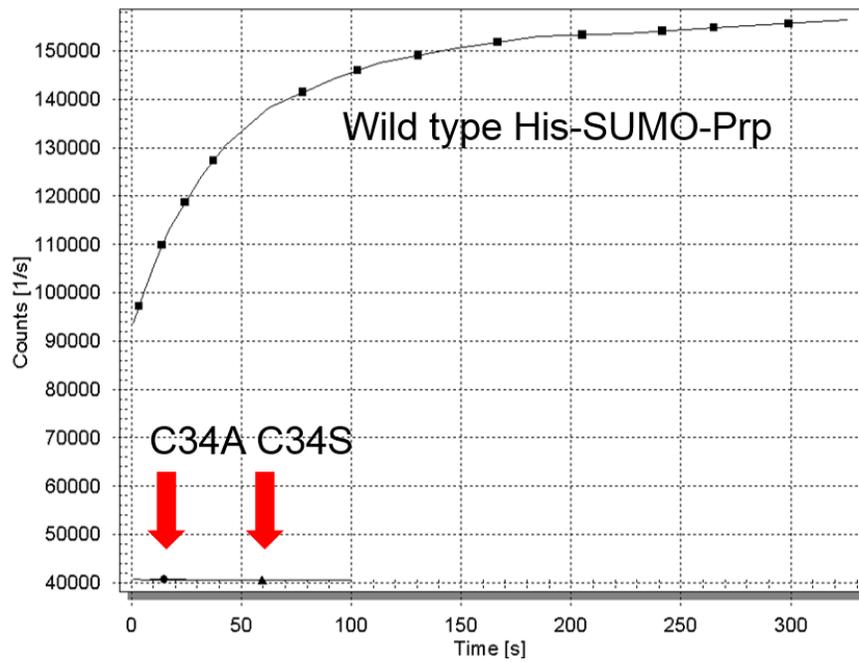
Figure 32. Only active tagged Prp causes fluorescence increase in the presence of quenched peptide substrate

Wild type Prp, mutant Cys 34 Ala and mutant Cys 34 Ser were assayed as N-terminal His-SUMO fusions. Only wild-type Prp cleaved the peptide substrate to release the fluorophore and caused relative fluorescence increase.

Chapter 8

Discussion

In this study, we have found that ribosomal protein L27 in Firmicutes and related bacteria has an N-terminal extension when compared to *E. coli*. This is important because the extension occludes N-terminal residues that have been well studied in *E. coli* L27 and found to be vital to A and P-site tRNA stabilization. We demonstrated that the extension must be cleaved away for *S. aureus* to survive and that the protease that performs the cleavage has a dominant negative effect when expressed as a catalytically inactive mutant. Overexpression of the inactive mutant gives rise to an aberrant pre-50S particle that contains some amount of the inactive protease in complex with un-cleaved L27, which would cause steric interference with the addition of later r-proteins like L16. This L27/Prp bound pair seems to also have blocked some L27 incorporation, as the mass spectrometry data suggest that L27 is depleted from ribosomes on the whole. It seems from these data, however, that a non-functional, likely extremely deleterious pro-protein form of L27 can be assembled into the 50S. The reason that this is possible is as yet undiscovered.

Importantly, not only is this L27 processing step essential, it is apparently highly regulated, because "pre-cleaved" L27 Δ 2-9 cannot complement an L27 deletion. This raises questions of whether the function of this process is related to ribosome biogenesis, regulation of translation, or perhaps both. The ribosomal biogenesis arrest produced in the presence of dominant negative Prp C34A is reminiscent of other studies of inactive ribosome assembly GTPase chaperones that bind but cannot dissociate.

Ribosomes are ancient machines. By far the largest group of genes in the Last Universal Common Ancestor (LUCA) are associated with translation (Fox, 2010). The ribosome is thought to have been nearly fully formed in LUCA, indicating that most of its development had occurred prior to that point in time, 3.5 billion years ago. While ribosomal structure and function are thought to be generally conserved across all domains of life, there are differences in composition that impact ribosomal size and function between eubacteria, archaea and eukaryotes and that are currently being exploited for therapeutic purposes. This new work demonstrates a novel twist in the story of bacterial ribosomes. **Our findings indicate that there are aspects of the basic biology of the ribosome in *S. aureus* and related bacteria that differ substantially from that of the *E. coli* ribosome.** It remains to be seen whether these differences can be a target for selective antimicrobial therapies.

Ribosomal biology has been intensively studied in the Gram negative eubacterium *E. coli*, but is far less well understood in other bacteria. The information is largely limited to studies in *E. coli* and data from available crystal structures, of which there are only

three for bacteria: *E. coli*, *Thermus thermophilus* and *Deinococcus radiodurans*. The latter two bacteria are hardly representative of the mesophile pathogens that we must attempt to target with antimicrobials. Nevertheless, the structure that is most often used to model ribosomal antibiotics is that of *T. thermophilus*, because it crystallizes readily. This is understandable, but at the same time we must acknowledge how this limits our view of ribosomal biology. There are good reasons to suspect that large differences exist in the ribosomes of many bacteria, especially those that contain the L27 N-terminal extension.

Work has been done in both *E. coli* (Eco) and Firmicute model organism *Bacillus subtilis* (Bsu) to understand the importance of each ribosomal protein to the function of the cell. Of fundamental note in these studies: the essential large subunit ribosomal proteins differ! There are some that are essential in Eco but dispensable in Bsu, including L22, L23, and L28. Those that are essential in Bsu but dispensable in Eco are L21, L24, L27, and L30 (Akanuma, 2012). Ribosomal protein complements also differ between bacteria. In a bioinformatic study of 995 bacterial genomes, 44 of 56 annotated r-proteins were found to be ubiquitous. Non-universal r-proteins L19p, L31p, L34p, L36p, L9p and S16p were missing from only one to three genomes examined. The rarest non-universal bacterial ribosomal proteins were determined to be L7ae, L25p, L30p, S21p, S22p and S31e, found in only a few groups (Yutin, 2012). It is not clear whether bacteria that encode L7ae (as noted, a homolog of Eukaryote and Archaeal L7) ever actually incorporate it into ribosomes. Though the gene for L7ae is present on the chromosome of *B. subtilis*, the L7ae protein was not detected in a proteomic study of fully assembled *B. subtilis* ribosomes (Lauber, 2009).

When r-proteins in each bacterial genome were concatenated and a phylogeny was drawn from this information, three major groups dubbed "megaphyla" emerged as deeply related clades of bacteria. Our interest was drawn to megaphylum III, which contains Fusobacteria, Mollicutes (aka Tenericutes), and Firmicutes. The Synergistetes were fairly closely related on a single offshoot of the tree (Yutin, 2012). All of these bacteria except one clade of Mollicutes contain the L27 N-terminal extension and Prp.

Mollicutes contain the smallest genomes of all bacteria that can be grown in axenic culture. *Mycoplasma genitalium* represents the iconic minimal "free-living" cellular genome at about 580,000 base pairs in length. Fully twenty-five percent of its genome is devoted to encoding genes related to translation. Further, the essential genes of this organism have been elucidated (Glass, 2006). Mollicutes can therefore be used as a reference for a postulated minimal essential translation apparatus (Grosjean, 2014). Four main subgroups of Mollicutes have been identified - Spiroplasma, Pneumoniae, Hominis and *Acholeplasma Aneroplasma Candidatus* phytoplasma (AAP). Of the culturable members of these groups, only the Hominis group lacks N-terminally extended L27 and the concomitant Prp homolog. If these minimal bacteria encode this system which requires what could be considered two "extra" pieces, it would be a strong argument for that system's importance. The translational apparatus of the Mollicute groups also demonstrates that some of the apparently disposable genes belong to the class that modify ribosomal RNA and ribosomal proteins - of which Prp certainly would be one. Why

then, are the L27 N-terminal extension and Prp retained? There must be some reason for even the smallest of bacterial cells to contain this cleavage event.

As the most discarded or shuffled translation genes are often rRNA and r-protein modification enzymes, it follows logically that ribosomal protein post-translational modifications differ widely. L27 has been mentioned extensively, but Bsu, unlike Eco, does not methylate its L3, doubly methylates its L16 and does not methylate its L33 (Lauber, 2009; Arnold, 1999). Many ribosomal biologists are starting to believe that the modifications themselves are not as important as the interaction of the modification enzyme with the ribosome at a given point. Frequently, if the nucleotide to be modified is mutated such that it cannot be methylated, lack of methylation at that site has no discernable effect on the function of the ribosome (Williamson, 2008). Simply the association of the modification enzyme at the right point in assembly may promote conformational changes in both enzyme and substrate that could chaperone the formation of complex assemblies (James R. Williamson, personal communication). The data gathered thus far regarding the ability of inactive PrpC34A to inhibit ribosomal assembly supports this hypothesis.

Given this information it seems impossible that all bacterial ribosomes function in an identical manner and they may even have significant conformational differences. Ribosomal protein processing is an unexplored mechanism that may provide new insights into the regulation of protein synthesis or possibly a new ordered pathway of ribosome assembly.

Ribosome Assembly

Almost all studies of ribosome assembly have been performed in *E. coli*, with only a few recent studies in *B. subtilis*. Ribosome assembly is a complex process, especially in the case of the large subunit. The production of rRNA is tightly coupled to growth rate and highly regulated. The ribosomal proteins, which must be produced in a concerted, stoichiometric manner, are often self-regulating, and many serve as repressors for the operon in which they are encoded (Shajani, 2011). Ribosomal proteins must bind their site on the nascent rRNA while it is being transcribed and simultaneously folded. This process is intermittently assisted by the binding and release of myriad rRNA modification enzymes, GTPases, and at least 20 to 30 other chaperones. Though it has often been modeled as an extremely hierarchical and step-wise progression, the sum of these parallel events is a largely stochastic and statistical process that results in many convergent pathways to assembly (Williamson, 2008; Sykes, 2010).

The in vitro assembly of a functional 50S subunit of the *E. coli* ribosome by Nierhaus and Dohme in 1974 was met with deserved acclaim, but the conditions under which assembly was performed were harsh and nonphysiological (Nierhaus, 1974). Nevertheless, the ordered, linear Nierhaus Map of 50S subunit assembly has been cited by those studying bacterial ribosomes ever since. In vitro large subunit assembly consists of three kinetic intermediates: a 33S particle, a 41S to 43S particle and a 48S particle. The in vivo assembly intermediates discovered are of similar sedimentation rates, but

their ordered introduction of ribosomal proteins is quite different. For instance, late 33S or 41S in vitro particles contain L2, leading to its classification as an early-binding protein (Nierhaus, 1974). However, *E. coli* in vivo 32S and 43S large subunit precursors were found not to contain L2, indicating that it was a middle to late-binding protein (Nierhaus, 1991). Study of the assembly of the 50S subunit is particularly problematic, especially when compared to the well-characterized 30S subunit. Many types of local secondary structure occur during transcription of the 23S rRNA, and it is not yet understood how all of these are chaperoned and coaxed into an active complex. One of the roles r-proteins seem to perform is to effect the gradual decrease of rRNA flexibility, guiding the formation of single stranded secondary structures until an active particle is achieved. There are also many RNases involved in assembly, most of which are not understood. The in vitro work also certainly does not account for the truly massive cadre of ribosomal GTPases, methylases and other possible chaperones that are highly conserved, some of them across all domains of life (Shajani, 2011).

Recently, investigators at The Scripps Research Institute, Michigan State and McMaster University have employed the increased precision of modern mass spectrometry to analyze large subunit in vivo assembly intermediates and determine their protein content. Current methods of stable isotopic labeling and quantitative mass spectrometry (qMS) can serve to analyze both rRNA and proteins and check for modification of those components and to examine their relative order of addition to the ribosome (Williamson, 2008).

To inventory and quantitate ribosomal proteins from in vivo assembly intermediates, two cultures of *E. coli* were grown, one in ^{14}N media, the other in ^{15}N media. All the ribosomal subunits and assembly intermediates gathered from the ^{14}N culture were fractionated after sucrose gradient centrifugation. Each fraction was spiked with the same amount of exclusively fully assembled 70S ribosomes gathered from the ^{15}N culture. In this way, the spectrum of each partial assembly ^{14}N peptide could be compared in absolute area and intensity with the spectrally shifted fully assembled ^{15}N 70S standard peptide (Chen, 2013).

Isotopic pulse labeling studies involve cells that are harvested from normal media and grown in 50% ^{15}N media for various time points. When these fractions are compared to a fully ^{15}N labeled 70S ribosome standard, lighter, unlabeled peptide peaks represent r-proteins incorporated before the pulse and 50% labeled peaks represent r-proteins incorporated after the pulse (Sykes, 2010). This method was used to distinguish "on pathway" intermediates from degradation products by how quickly the label is incorporated into all proteins in a particular particle, as degradation products would be less likely to incorporate label over time. Particle labeling kinetics can be further subdivided into peptide labeling kinetics. The ribosomal proteins that are labeled fastest are the ones that were first translated after the pulse. It has already been shown that free ribosomal protein pools are not large for early assembly proteins, so the majority of those have already been incorporated into a particle without label and require a second ribosomal generation to become labeled. In contrast, late assembly proteins are produced

and added to existing unlabeled assembly intermediates, demonstrating faster label uptake (Chen, 2013).

These isotopic labeling studies in *E. coli* revealed a more complex assembly landscape for the 50S subunit than previously imagined. 50S assembly is more diffuse and nuanced than 30S assembly, which forms a less complex particle that consists chiefly of the decoding center. Two regions in the large 50S subunit must be assembled with great accuracy – the exit tunnel and the PTC, compared to only one region in the smaller 30S subunit. Assembly of the 50S subunit takes twice as long as the 30S, and appears to start at the "back" - the region on the opposite side of the PTC, above the exit tunnel. There is an intermediate step in which r-proteins are inserted in various areas all over the surface of the particle, and assembly then ends with construction of the central protuberance and the PTC. Large subunit r-protein labeling kinetics were compared and grouped temporally. With the exception of the proteins that bind to the 5' end of the 23S rRNA transcript (L20, L21, L22 and L24), large subunit proteins bind every region of the rRNA at every stage of assembly. This study led to the re-classification of L27 as a late-binding protein, though Nierhaus found it to be middle-binding in vivo in 1991. In addition to r-protein temporal assembly designations, this study uncovered previously examined and and hitherto unknown GTPases and chaperones that likely assist in the assembly of the large subunit (Chen, 2013). What is known about the roles that ribosomal chaperones play will be discussed later in this work.

When in vivo isotopic labeling studies were performed in *B. subtilis* at Michigan State University and McMaster University in collaboration with J. R. Williamson at Scripps, an RbgA (large subunit assembly GTPase not present in *E. coli*) depleted cells produced 45S particles somewhat similar in r-protein composition to the aberrant particle in this work. Both particles lost staggering amounts of L16, L25 and L27 and some of L18, all proteins resident on or near the central protuberance. There were some notable differences, however: our particle was conspicuously deficient in L29 and had normal levels of L30, while theirs had these ratios reversed. While their particle was deficient in L27 it had normal levels of L20, yet we lost significant amounts of L20 in our L27-deficient particle. As previously mentioned, L27 deletion in *E. coli* depletes the large subunit of L20 and L21, yet L21 was within normal range for our particle while L20 was significantly depleted. Many of these discrepancies might be corrected with increased accuracy in our measurement of our r-protein peptides, which will require some type of peptide labeling for qMS. Taking into account the limited scope of data that we have, it seems likely that we have trapped an earlier assembly intermediate than the one studied by this group, simply because many more r-proteins intrinsic to the interior of the ribosome were deficient in our particle compared to theirs (Jomaa, 2013).

When the *B. subtilis* 45S particles were subjected to cryo-electron microscopy (cryo-EM), it was found that the central protuberance and PTC were the least mature regions, with disorder in the sites that should be open to accept tRNAs. These particles were incompetent for translation, but could be matured when both RbgA and L16 were reintroduced. The authors of this study theorized that the final RbgA effect, which seems

to allow L16 incorporation and support other late r-protein association, was a checkpoint on PTC formation, only permitting the active site of the ribosome to mature as long as the rest of the ribosome was assembled correctly (Jomaa, 2013).

As a further illustration of the confusion in the literature surrounding ribosome assembly in *B. subtilis*, a paper published immediately before that of Jomaa *et al* 2013 used post growth isobaric tagging methods for qMS (as opposed to in vivo isotope labeling) on *B. subtilis* r-proteins from 45S subunits in the same RgbA depletion strain utilized by Jomaa *et al.* The study by Li *et al.* in 2013 measured extremely dissimilar ribosomal protein ratios and significantly different cryo-electron microscopy images, with some particles from their 45S pool showing absolutely no density for the CP, even though their aberrant particles arose from the same strain background with similar ribosome purification methods. I noted only three substantial methods deviations that might account for these differences. The first difference, as mentioned above, was the use of in vitro isobaric tag labeling instead of in vivo stable isotope incorporation. In vitro labeling is reported to be less accurate because it relies on successful conjugation reactions between the isobaric tags and the peptides rather than assured label incorporation via growth with an isotope tag as the sole source of nitrogen. Further, degradation products cannot be distinguished from on-pathway intermediates using isobaric tag labeling (Chen, 2013). Second, wild type 50S control subunits in the Li *et al.* 2013 study were collected from lab strain *B. subtilis* 168, not the RgbA depletion strain background in RgbA replete conditions as in Jomaa *et al.* Third and most interestingly, the 45S immature subunits collected from RgbA depleted cells in the Li *et al* study were subjected to an extra

purification step involving the addition of large amounts of wild-type 30S subunits to the 45S fraction and incubating for ten minutes at 37º C, allowing subunits to associate if possible, and then subjecting that particle mixture to another sucrose gradient fractionation to re-isolate the 45S subunits. Ostensibly this would clear the 45S peak of any contaminating mature 50S particles, or for that matter, any particles that retained functional vestiges of the intersubunit bridges required for stable association with the 30S, by moving those complexes to the 70S peak.

The particles reported by Li *et al.* 2013 showed measurable loss of L3, L15 and L13, setting them at odds with the data in this work and that of Jomaa *et al*. These proteins have not been reported to be involved in intersubunit bridging, so their disappearance is mysterious. The fact and extent of the losses of L18 and L20 that we observed disagree with both Jomaa and Li. There was a group of proteins that were consistently lost or could not be detected in all of the studies, this work included: L28, L16, L27, L33, L35 and L36. While our methods of quantification were primitive compared to the other studies, the consistency of our results with other more precise methods in identification of these proteins was reassuring. There are very likely well-defined groups of r-proteins added at different stages of assembly, but the interconnectedness and dynamic nature of the ribosome significantly complicates interpretation of these groups. Later proteins stabilize earlier proteins and all participate in massive rRNA rearrangements during ribosome biogenesis. The loss of just one protein can cause major conformational shifts tens of angstroms distant from the location of the deficit. It is clear that much work remains to be done using qMS, improved cryo electron microscopy resolution and hitherto unexplored

avenues to uncover the complexities of bacterial ribosome assembly, especially in organisms that are not *E. coli*.

The type of in vivo isotopic labeling study performed in Chen *et al*, 2013 and Jomaa *et al*, 2013 has worked well with *B. subtilis*, yet in its current form the procedure cannot be performed with *S. aureus,* which cannot grow well (if at all) in minimal media. The use of stable isotope labeled amino acids in culture (SILAC) reagents to label r-proteins in *S. aureus* grown in defined media, while theoretically possible, is prohibitively expensive. It would cost about one thousand dollars per liter of media to incorporate only a few labeled amino acids, while complete amino acid labeling comparable to the nuanced effects achieved in the above studies would cost multiple thousands of dollars per liter. Other chemical labeling methods such as isobaric tagging for relative and absolute quantification (iTRAQ) tags must be applied to *S. aureus* ribosomal samples after their collection and before mass spectrometric analysis, rather than labeling in vivo. iTRAQ tags fragment at offset masses, shifting a peptide to a higher mass than its counterpart in the control sample, so peaks may be compared. As mentioned previously, these methods are less accurate than pulse labeling with stable isotopes. It is probably a great deal more practical for extensive ribosome assembly studies in Firmicutes to be performed in *B. subtilis*, which thrives in minimal media and is also easier to lyse given its elongated cell morphology compared to spherical *S. aureus*.

Ribosome friends – ribosomal GTPases and chaperones

In order to understand possible functions of the Prp/long L27 cleavage system, it is helpful to examine existing ribosome associated proteins known to be involved in assembly or ribosomal regulation and their possible correlations with Prp/long L27. There are many, many documented proteins that have been observed to interact with the ribosome in various ways; they have been reviewed extensively (Iost, 2013; Goto, 2013; Wilson, 2007). For the purposes of this work I will only discuss ribosome biogenesis proteins that are conserved within the Mollicutes species that contain Prp and long L27. As previously mentioned, Mollicutes can be considered the model for a minimal translation apparatus due to their small genomes (Grosjean, 2014), making retention of the Prp/long L27 system in many clades seem all the more significant. We can compare and contrast the ribosome-associated chaperones and regulators present in Mollicutes with and without Prp/long L27 to provide focus on what might be different and important to the ribosomes and ribosomal pathways in our particular group of bacteria. Perhaps the pattern of ribosomal chaperone conservation within this group will eventually give clues to the reason that the Prp/long L27 system is conserved.

GTPases involved in assembly

Ribosomes require a number of trans-acting factors that use GTP or ATP to engage in energy consuming structural rearrangements that are required for correct rRNA folding during maturation. These enzymes generally have one conformation that has high ribosomal affinity when bound to GTP, and a radically different conformation after GTP hydrolysis that has considerably less affinity for the ribosome (Goto, 2013). RgbA, previously discussed in the section on ribosomal assembly, contains a GTPase domain

that is conserved across all three forms of life and is involved in a late step of 50S particle maturation. Depletion of RgbA leads to decreased 70S ribosomes relative to subunits, disorder in the central protuberance of the 45S particle produced and loss of late binding r-proteins, such as L16, L28, L27, L36 and L33a (Jomaa, 2013; Li, 2013). Recently, it has been demonstrated that a function of RbgA seems to be to properly position r-protein L6 during assembly, as mutations in RbgA that caused ribosome assembly defects were rescued by compensatory mutations in L6. Though they did not affect cellular growth rate, these L6 mutants still caused some disturbance in 50S assembly. From this and the locations of the L6 mutants on a well-conserved L6/rRNA interface, the authors theorize that RgbA associates with L6 and helix 97 in order to properly position them, and/or to allow incoming late assembly proteins to bind (Gulati, 2014).

This GTPase is universal, and there is so far no recognized ribosomal GTPase differentially conserved in Prp/long L27 Mollicutes vs other Mollicutes. It is therefore less likely that RgbA is involved in the pathway created by Prp/L27, but understanding its role serves to highlight the function of ribosomal chaperones that reconfigure the assembling structure in favor of proper protein placement. Prp might have a similar function, associating with the ribosome by binding to the L27 N-terminal extension, allowing conformational changes to occur around it or effecting them directly and then cleaving away the small peptide in order to dissociate from the assembling ribosome. Indeed, this reaction requires no ATP or GTP, possibly making it an energy efficient way to solve a kinetic problem that occurs in rRNA folding or r-protein incorporation.

DEAD-box helicases

DEAD-box helicases are widely conserved throughout all domains of life and essential in eukaryotes; in complex cells their function is to chaperone single stranded RNA and thus facilitate almost every aspect of RNA metabolism and transport. In contrast, the functions of studied bacterial homologs are poorly understood, and the proteins are usually only essential under cold shock conditions. In *E. coli* and Firmicutes DEAD-box helicases are highly associated with ribosome biogenesis and RNA metabolism. The protein known as DeaD or CsdA in *E. coli* has been shown to aid in the assembly of the 50S subunit. Strains with *deaD* deletions grow slowly, form aberrant 40S precursors under normal growing conditions, and the majority of their mature 50S subunits are defective in translation, indicating that DeaD might have a role in the final steps of ribogenesis. At low temperatures DeaD co-sediments with the 50S subunit, suggesting that it might have a role in preventing undesirable rRNA secondary structure that occurs under those conditions. In Firmicutes the DeaD homologs are also non-essential at 37ºC, but become essential under cold shock conditions (Iost, 2013). Studies of Firmicute DeaD effects on ribogenesis are sparse, but DeaD deletions were analyzed for ribosomal anomalies in *Listeria monocytogenes* to understand its psychrotolerance mechanisms. These studies demonstrated that the deletion caused a loss of 70S subunits and accumulation of a 40S precursor at 16ºC. Further, the authors demonstrated via complementation that DeaD binds the 50S subunit and fulfills its role as a chaperone via its C-terminus, as the 40S precursor could be ablated by complementing with wild-type DeaD but not a C-terminal deletion mutant (Netterling, 2012). Interestingly, even though Mollicutes show very low overall conservation of DEAD-box helicases, the Prp/long L27

group universally retains a protein highly related to DeaD, while the Hominis group that lacks Prp/long L27 does not (Grosjean, 2014).

Ribonucleases

Bacterial rRNA is produced as a precursor molecule that must be trimmed by Rnase III or a functional homolog to begin the formation of mature 23S, 16S and 5S molecules. In contrast to 16S rRNA, which after Rnase III cleavage begins as 17S rRNA and undergoes fairly extensive processing, in *E. coli* precursor 23S rRNA loses only a few bases during maturation, seven from the 5' end and eight from the 3' end. Yet, it was recently shown that retention of the eight nucleotides on the 3' end caused a significant defect in ribosome assembly (Gutgsell, 2012). Processing of 23S rRNA occurs via Mini Rnase III or alternatively Rnase J1, J2, Ph and YhaM in *B. subtilis* (Redko, 2010). *Mycoplasma genitalium* primarily utilizes Rnase R for maturation of its 23S rRNA. In this system Rnase R performs various precise methylation-sensitive cleavages of rRNA and executes efficient tRNA processing (Lalonde, 2007). Mollicutes have Rnase III homologs, but no Rnase Ph or YhaM. They also do not contain Rnase A, B, D, E, G, HI, HII or T, but Mollicutes that can be grown in axenic culture uniformly retain Rnase R, J1, J2, P, and Nrna, a ribonuclease specific to very short RNAs. However, Prp/long L27 group Mollicutes retain another ribonuclease, Rnase Y, that is not present in Mollicutes without the L27 cleavage system (Grosjean, 2014). Rnases can be involved in the processing of RNAs to maturation and also in the RNA metabolism of the cell. Regulation of certain Rnases can stabilize or destabilize existing transcripts, allowing for global cellular response to stimulus.

Ribonuclease M5 is responsible for 5S rRNA maturation in low G-C Gram positive bacteria that do not encode the RNase E/G homolog that performs that role in the proteobacteria. Ribonuclease M5 and 5S rRNA maturation have been shown to be nonessential in *B. subtilis* (Condon, 2001; Allemand, 2011), yet Prp/long L27 group Mollicutes retain this processing enzyme while Hominis group Mollicutes without Prp/long L27 do not (Grosjean 2014). Further study reveals that while it contains an expected conserved domain with resemblance to an N-terminal topoisomerase-primase found in DnaG homologs (NCBI cd01027, with relationship to pfam domain 13155), the M5 nuclease C-terminal domain DUF4093 is of unknown function. This DUF is conserved in Firmicutes, Mollicutes in the groups that contain Prp/long L27, Fusobacteria and Spirochaetes. The domain has not been found to be conserved in Synergistes, however, which have Rnase E homologs instead of ribonuclease M5, making retention of this DUF not perfectly correlated with the presence of Prp or Prp/long L27. Yet, the interesting dichotomy of Mollicute Ribonuclease M5 conservation in those that contain Prp/L27 versus its loss in those that do not suggests marked differences in their rRNA maturation pathways.

Nutrient stress response regulators

Apart from ribosome assembly, it seems theoretically possible that the amino acid starvation response could cause down-regulation of Prp, causing buildup of long L27 and arresting synthesis of new ribosomes. Bacteria are known to regulate production of their ribosomes in response to nutrient limitation in part by controlling the transcription of rRNA. In *E. coli* this is a well-studied phenomenon called the stringent response and its main

effectors are RelA, SpoT and pyrophosphoryated GTP or GDP, abbreviated (p)ppGpp. (p)ppGpp uses protein DksA as a cofactor to exert widespread changes in transcriptional regulation by binding to RNA polymerase (RNAP). The stringent response also affects translation and replication. This phenomenon has been intensively studied for over forty years, yet the exact mechanism (p)ppGpp uses to exert its effect remains uncertain, and its binding site on RNAP remains unknown (Wolz, 2010).

Recent work has uncovered that proteobacteria like *E. coli* are distinct from Firmicutes like *S. aureus* in their (p)ppGpp production pathway and its regulatory targets. There is no homolog of DksA in Firmicutes, no indication that (p)ppGpp interacts directly with RNA polymerase, and concomitantly the region before the transcriptional start for rRNA operons does not contain the discriminator sequences that would decrease their transcription, as in the *E. coli* (p)ppGpp/DksA/RNAP response system. Instead of RelA ((p)ppGpp production) and SpoT ((p)ppGpp hydrolysis), *S. aureus* contains a fusion protein with hydrolase and synthase domains called Rsh and two small (p)ppGpp synthases often called RelP and Rel Q (Wolz, 2010). It appears that the function of RelP/Q is to produce small amounts of background (p)ppGpp during exponential phase, which causes the hydrolase domain of Rsh to be essential for consistent low level detoxification. Rsh responds to amino acid stress while RelP/Q are more responsive to cell wall stress or alkaline conditions (Geiger, 2014). Much work remains to be done to understand the (p)ppGpp regulon in Firmicutes, including response regulator CodY and virulence up-regulation under stress conditions, which are beyond the scope of this dissertation.

In the Hominis group of Mollicutes that do not contain Prp/long L27, there is no obvious (p)ppGpp synthase homolog to be found anywhere in the genome, so it is possible that this GTP nucleotide regulatory system does not even exist in that group of bacteria. In contrast, the Prp/long L27 Mollicutes uniformly retain the bifunctional Rsh protein, but do not retain the smaller RelP/Q (p)ppGpp synthases (Grosjean, 2014). Studies are needed that check L27 cleavage status in stationary phase or other stress conditions to determine whether the Prp/long L27 system responds to various types of cell stress by abolishing L27 cleavage.

Chapter 9

Remaining Questions and Future Directions

<u>Is Prp involved in the formation of the PTC or is L27 cleavage immediate?</u>

At this point we cannot tell if L27 is only cleaved after incorporation into ribosomes (i.e., perhaps Prp must transport un-cleaved L27 to the PTC, then cleave and recycle). Even though we have performed immunoblotting experiments in *E. coli* in which L27 Myc His$_6$ is completely cleaved during co-expression with Prp, it could be that Sau L27 was first incorporated into the *E. coli* ribosome prior to its cleavage. The *E. coli* ribosome has previously been shown to assemble foreign versions of L27, such as that of *Aquifex aeolicus,* which is comparatively C-terminally extended as opposed to N-terminally extended (Maguire, 2001). The *E. coli* strain expressing Sau L27 Myc His$_6$ without Prp present was also prone to extremely slow growth after protein induction and had to be grown at 30^0 C to retrieve even small amounts of expressed protein. It is possible that *E. coli* expressing Sau L27 sickened as it began to assemble the toxic N-terminally extended protein.

I think it is justified to speculate that L27 cleavage is not immediate since production of L27 without its cleavage motif ("pre-cleaved" L27 Δ2-9) cannot complement an L27

deletion. Ostensibly it is possible that pre-cleaved L27 is not folded correctly as a result of the lack of the extension and cleavage event or that Prp must bind L27 on the cleavage motif in order to transport it to the ribosome. In either case, L27 Δ2-9 will not be present in the 50S subunit, and this will likely cause an aberrant assembly intermediate identical to ones collected from L27-depleted cells. Ribosomal subunit analysis and subsequent mass spectrometry of r-proteins are the next experiments planned using the L27 complementation system.

It is important to note that pre-cleaved L27 produces a protein identical to that found in mature ribosomes, indicating that its normal function of tRNA stabilization should be intact. Either this protein is assembled and that is insufficient for growth or it is not assembled due to folding issues or a lack of binding to a Prp chaperone. If the protein is assembled and results in assembly aberrations or otherwise nonfunctional ribosomes (the growth arrest indicates that no new functional ribosomes are produced in its presence), then we can reasonably suspect that the event of Prp cleavage has an essential effect on ribosome formation, likely at or near the PTC.

If it is true that Prp must bind un-cleaved L27 to transport it to the ribosome, ribosome analysis of the strain expressing only un-cleavable L27 might show low incorporation of un-cleaved L27 in the 50S subunit. This L27 mutant is unable to co-purify native Sau Prp, which seems to indicate a considerable binding deficiency compared to the reciprocal effect seen when catalytically inactive PrpC34A or C34S co-purifies with tagged native L27. If Prp only cleaves L27 after it has been assembled into the ribosome, then both the

strain expressing un-cleavable L27 and that expressing pre-cleaved L27 should contain the mutant protein in their 50S ribosomal subunit or in a semi-equivalent assembly intermediate. While it seems formally possible that un-cleaved L27 does not fold properly without the cleavage event, we have already shown that un-cleaved (native cleavage motif) L27 can be assembled into the ribosome when bound to Prp. The western blots of aberrant 50S particles produced during expression of PrpC34A showed that the entire Prp/L27 complex could be stably associated with the pre-50S subunit.

When ribosome analyses are performed on the complementation strains there is a distinct set of possible results:

- If both pre-cleaved and un-cleavable L27 can be assembled into particles, Prp is not necessary to traffic L27 to the ribosome.
  Prp plays some later role, as if the cleavage event was a chance for Prp to associate with the assembling ribosome and thus aid maturation.

- If pre-cleaved L27 is assembled and un-cleavable L27 is not, cleavage might be necessary so that L27 can assemble, but Prp is probably not responsible for trafficking L27 to the ribosome.

- If un-cleavable L27 is assembled in any amount and pre-cleaved L27 is not (low affinity Prp association to un-cleavable L27 may allow this situation to occur) Prp could be implicated as a necessary protein for L27 incorporation into the ribosome.

- If neither un-cleavable nor pre-cleaved L27 assembles, either both proteins are defective in folding (a prospect that I see as unlikely), or Prp could be responsible for L27 traffic to the ribosome.

My current pet theory is that the L27 N-terminus forms a temporary place holder during assembly or is involved in Prp traffic to the ribosome. After Prp has associated with the assembling 50S at a key step or checkpoint, causing conformational changes that allow proper ribosome maturation, it is released when cleavage occurs. This would finally make sense of why an essential PTC element is made as a poisonous, inactive pro-protein. If active PTC formation is dependent on conformational changes associated with Prp binding, any interruption or abrogation of Prp binding would be deadly and it would not matter that the PTC was obstructed by un-processed L27.

What role does Prp play in phage capsid development?

In Spilman *et al.* 2012, our group reported the results of co-expressing the capsid and scaffold of phage 80α without their cleavage motifs (as if those motifs had already been cleaved away, as in the "pre-cleaved" L27 mutant). Compared to the ordered pro-capsid like structures produced when wild-type scaffold and capsid were co-expressed, the Dokland lab found that the cryo-electron micrographs (cryo-EM) of pre-cleaved capsid and scaffold co-expression products contained disregulated sheets of capsomers called polyheads (Spilman, 2012). Further, recent work in our group has demonstrated that the N-terminal cleavage motif of 80α scaffold is completely essential to the viability of the

phage and if it is absent, nothing resembling a capsid structure is observed under cryo-EM. In apparent contrast to this result, 80α capsid mutants in both the un-cleavable and the pre-cleaved form seem to have no effect on the viability of the phage (Wall, Klenow, Manning, Dokland and Christie, unpublished). This leads to the obvious question of the role of Prp in 80α capsid assembly.

All known tailed DNA phages assemble their capsid structural proteins into a rounded prohead or procapsid before a maturation event, often DNA packaging, causes the rounded structure to expand, thin, and angularize to form a mature capsid. The herpesviruses and many phages contain a maturation protease that aids in this transition, as mentioned in the introduction to Chapter 2. The role of the prohead protease has so far been observed to be removal of the assembly core or scaffolding proteins. In phage HK97, the N-terminal "delta" domain of the capsid protein is involved in intrinsic scaffolding activity, and is essential for assembly, but is not present in the final particle, having been cleaved away by the maturation protease (Oh, 2014). In the case of HK97, the protease must interact with the delta domain to be incorporated into the procapsid, where it allows the structure to assemble and then removes scaffolding proteins (Duda, 2013). In an alternative strategy with a similar theme, herpesviruses, phage P2 and phage λ encode an independent scaffold protein that has an N-terminal protease domain and a C-terminal capsid-binding domain; upon assembly, the scaffold cleaves itself away. In the case of λ, smaller protease subunits are made by internal initiation within the larger protease/scaffold fusion reading frame and these co-assemble with the full scaffold (Aksyuk, 2011).

In the case of 80α, we have determined one method by which scaffold interacts with capsid. At the scaffold C-terminus, conserved basic residues appear to form salt bridges with acidic residues in a helix/helix interaction with α2 in the capsid protein, which is located on the interior wall of the procapsid in cryo EM reconstructions (Wall, Klenow, Dokland and Christie, unpublished). If the N-terminal cleavage motif on scaffold is responsible for shuttling Prp to the nascent procapsid aggregate where it performs a crucial step in assembly, this could explain why absence of the scaffold motif obliterates procapsid formation.

In contrast, the function of the N-terminal motif on capsid remains unclear, especially as its absence or abrogation of its cleavage seems to have little to no effect on phage viability. Questions remain about the role of Prp in capsid assembly. It may be the fact that it is a dimer in solution allows it to bind two proteins at the same time before cleaving. Perhaps it works by binding two scaffold or two capsid proteins at the same time, promoting nucleation and aggregation before it exits by cleaving. This seems plausible with scaffold, but seems to be impossible for capsid as the motif-less and un-cleavable forms have no phenotype.

In order to determine whether Prp cleavage of scaffold is dependent upon procapsid assembly, we need to co-express C-terminally His$_6$ tagged scaffold with Prp in *E. coli*. Previous work involving scaffold and Prp was always done in the presence of the capsid protein. If Prp is truly an assembly/maturation protease, it seems unlikely that it will

immediately cleave scaffold in an un-assembled state. Further, if Prp does not cleave scaffold when they are co-expressed, it seems likely that Prp will co-purify with its tagged un-cleaved substrate, as it has done previously. From such a complex we could also ascertain the stoichiometry of Prp to scaffold by size exclusion chromatography and determining the relative protein band intensities in SDS PAGE.

<u>Why do some Thermotogales and Spirochaetes have a Prp homolog but not an N-terminal extension on L27?</u>

A bioinformatic analysis of Thermotogales and Spirochaete genomes was undertaken in the hope of revealing a new target for Prp. Even without N-terminally extended L27 present, Prp is still encoded in the same genomic position, between L21 and L27, which leads to unexplored questions of its regulation. Prp homologs in Thermotogales tend to have differing characteristics, such as a shorter flexible loop between the catalytic histidine and cysteine, and neutral to alkaline PIs where most Prp family members in Firmicutes are acidic. To address this, I performed numerous Blast-P searches in the Thermotogales and Spirochaetes with different L27 motifs, which have natural variation, to try to discern plausible motifs that might be hiding in other proteins.

Flagellar biosynthesis protein FlhB has emerged as a possible target for Prp in both the Thermotogales and Spirochaetes. FlhB contains an N-terminal cleavage motif that closely matches the L27 cleavage motif present in those Thermotogales that contain long L27 (Figure 33). The cleavage motif on FlhB can be either at the extreme N-terminus, as it is in L27, or as far into the protein as the 50th or 60th residue. The cleavage motif is always outside the conserved functional domain annotation. FlhB is a component of the

"export gate" of the apparatus responsible for the construction of the flagellum, and concomitantly export of proteins through the type III secretion system. It is a transmembrane protein that joins a large complex of other membrane proteins to create the gated channel through which flagellar structural proteins must pass (reviewed in Minamino, 2014). In support of this hypothesis, the *Bacillus subtilis* FlhB contains an almost identical 9-10 amino acid N-terminal motif to that in *B. subtilis* L27. Other bacteria that contain long L27 with one cleavage motif contain FlhB with a slightly different motif, showing a change from D to N, F to Y or an S, T or N instead of an A after the P1 F. It is known that there is some flexibility in Prp substrate specificity, because more subtle point mutations in the L27 cleavage motif can be tolerated. Specifically, a mutant form of L27 containing an S in place of A at the P1' site was still a substrate for *S. aureus* Prp and was cleaved appropriately, as were P4 Q::I and a substitution of Y for the P1 F (Wall and Christie, unpublished). It is normal for *S. aureus* Prp to cut three slightly different motifs: that of *S. aureus* L27, and those of phage 80α scaffold and major capsid protein. It seems probable that Prp affinity for each motif might be different, affecting the catalytic capacity for Prp on each target.
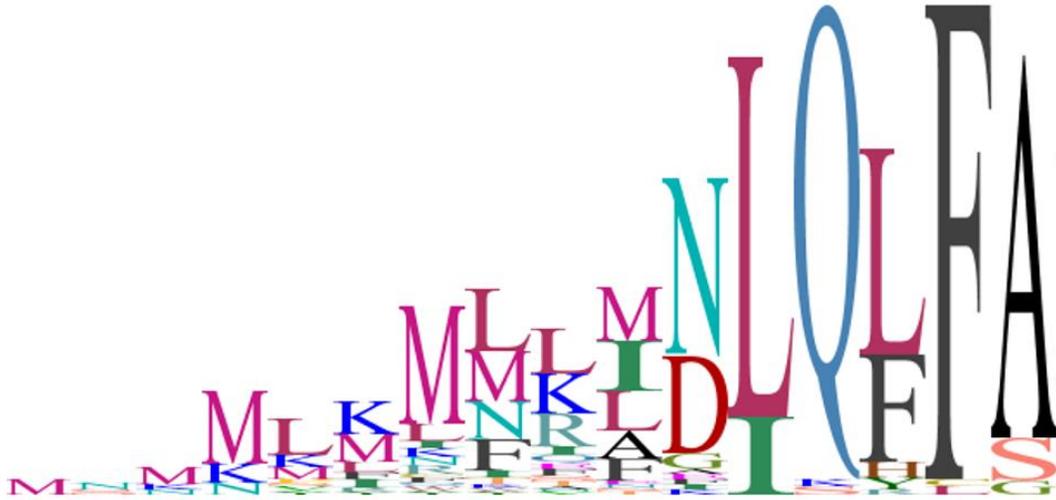
L27 cleavage logo
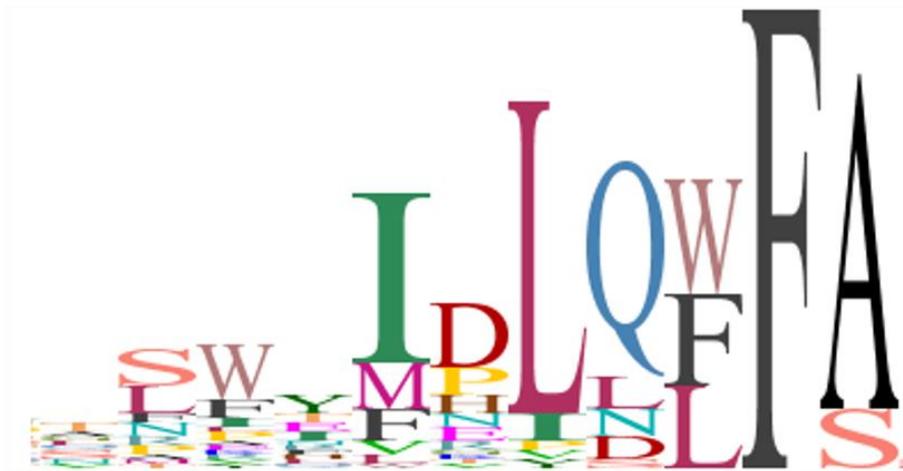
FlhB N-terminal logo from Prp-containing
Spirochaetes and Thermotogales

Figure 33. L27 cleavage logo compared to FlhB N-terminal sequence logo

Hidden Markov Model (HMM) logo representing alignments of N-terminally extended L27 and FlhB homologs in bacteria that contain Prp.

<u>Is the role of Prp primarily that of an assembly chaperone?</u>

All the data so far lead us to believe that the major role of this protease is in assembly of large complex macromolecular structures. It seems that these large assemblies have to be both chaperoned and regulated in their production, something we already know of the 50S subunit. The first piece of evidence that led us to the discovery of the protease was its involvement in the processing of phage capsid and scaffold components. Many Prp homologs are found on phage genomes, paired with the cleavage motif on their capsid structural proteins. Some bacteria have the cleavage motif on the N-terminus of their FlhB homolog, a transmembrane protein involved in the export gate required for flagellar biosynthesis and type III secretion systems. These phage and FlhB Prps apparently have little or nothing to do with ribosomes, but might have evolved from those that did.

Given this information, it would seem fitting to categorize our Prps by their substrates, which could lead us to a greater understanding of Prp/substrate interactions. In this way we might be able to design antibiotics directed against ribosomal Prps, which would be a completely novel target for antibiotic design, helping to fill the empty drug pipeline. If it is found that N-terminal FlhB cleavage is essential for motility in Spirochaetes, perhaps Prp-directed anti-motility drugs that stop progression of the bacteria deeper into tissues will be good adjuvant therapies against syphilis, Lyme disease, and relapsing fever. Indeed, bacteria might develop resistance much more slowly to what is probably a non-lethal drug.

How well does our model predict Prp function or substrate specificity?

Our model brings with it a range of predictions about the Prp mechanism that can be explored using the fluorescent assay we have developed. Mutations to be tested include G21A, H22A, D31A, S38A and N46A. Gly 21 is completely conserved; it seems to act as a hinge that allows free movement of catalytic His 22. We predict a kcat decrease in mutant G21A as increased rigidity around His 22 could impair its ability to come into close contact with Cys 34. His 22 may only adopt a catalytic conformation when substrate fills binding site. Mutant H22A should result in a complete loss of function, confirming the predicted catalytic dyad. Asp 31 is almost completely conserved and in our model its carboxylate group could interact with two amide nitrogens in the C terminus of a substrate in a non sequence-specific manner. If this is the case, D31A should produce a Km deficit. If Asp 31 also hydrogen bonds with His 22 when the protease adopts a catalytic conformation, then there would also be a decrease in kcat. Residue 38 is always either a Ser or Thr, which seems to indicate the importance of a hydroxyl group in the region below the catalytic dyad. This hydroxyl appears to hydrogen bond with the substrate amide nitrogen between the P2 Phe and the P1 Phe, possibly stabilizing their split conformation. Mutant S38A should therefore result in a Km deficit. Asn 46 is highly conserved in a group of bacteria closely related to *S. aureus*. It appears to form an amide-amide hydrogen bonding interaction with the conserved Asn of the target cleavage sequence. Mutant N46A would therefore reduce the Km of Prp for its substrate.

The assay can also be used to test Prp affinity for variant cleavage motifs. In early attempts to develop an un-cleavable L27, mutant cleavage motifs were generated with

sequences Q7I (MLKLNL**I**FFA), F9Y (MLKLNLQF**Y**A), and A10S (MLKLNLQFF**S)**. Prp was capable of cleaving all these proteins when they were expressed in *S. aureus* with a C-terminal Myc-His₆ tag for Western blot detection. Not much can yet be known about the affinity of Prp for these variant motifs, but there is apparently some leeway in the specificity of Prp for the cleavage motif (Wall and Christie, unpublished work).

<u>What work might be done to characterize existing Prps by what they cleave?</u>

I believe this question is key. If a Prp antibiotic were developed, theoretically it could kill all of the friendly Firmicute bacteria in the human gut. This is completely unacceptable and would only pave the way for antibiotic-associated diarrhea and evolution of resistance. Prp antibiotics would need to be as specific to small groups of bacteria as possible. L27 cleavage motifs vary between bacterial families. It would be extremely important to discern structural differences in the enzyme that binds MLKLNLQFF'A in *S. aureus* versus the enzyme that binds MLMNLQFF'S in *Lactobacillus plantarum* or MLNMNLQLL'A in *Clostridium difficile*. If significant structural differences could be found, then perhaps these antibiotics could be tailored to provide more targeted treatment.

Molecular modeling is not a perfect approach to this problem, but it is fast, cheap and might lead to more specific hypotheses that can help to justify expensive wet work. The Deacon Active Site Profiler is a program that compares known structures of active sites, notes the conserved and differing residues in the region surrounding the catalytic center and creates sequence string "functional site signatures" based on this information. The program then queries the protein databases in NCBI and sorts proteins into different signature groups that have similar active sites, but have variations in surrounding

residues (Fetrow, 2006). Ideally, we could use this information to make homology models of each Prp to its closest inferred structural relative and then check which residues and side chains land near the active site. These data, when paired with their target cleavage motifs on L27, might yield substrate specificity interaction patterns within Prps that cleave similar motifs.

Docking and molecular dynamics could begin to shed light on how *S. aureus* Prp might or might not bind disparate cleavage motifs. For example, if favorable docking and molecular dynamics results cannot be achieved between this *S. aureus* model and the peptide substrate from the *C. difficile* L27 motif, this could suggest that Prp might not bind the *C. difficile* L27 very well, providing hope for the idea of differential targeting in genera-specific active sites.

To sufficiently verify this type of modeling, we must collect wet data suggesting that *S. aureus* Prp does not bind the *C. difficle* motif. Fortunately, the assay we developed is ideal for testing this kind of question quickly. We would simply order the C. diff cleavage motif peptide and compare it to the Sau L27 substrate in a fluorescent competition assay. If the C. diff peptide cannot compete at a measurable level with the native substrate, that would point toward disparate binding/active site families within the Prp clade that might be differentially targeted. To confirm, we could also perform isothermal titration calorimetry (ITC) with Sau Prp and other motifs to check their actual binding constants. These data could demonstrate that Sau Prp binds the C. diff motif too well, making it less likely that genera-specific antibiotics could be developed.

In either case, given the widespread nature of this phenomenon, intense biochemical and structural study of the different Prps and their substrates is important, and will likely lead not only to better understanding of Firmicute biology, but also to desperately needed novel drug interventions.

<u>Given its importance, what is the best way to design antibiotics that target Prp?</u>

Protease inhibitor drugs are currently used for treatment of coagulation disorders, hypertension, HIV infection, cancer and diabetes. These drugs are effective but are also known for their side-effects, often due to problems with specificity. Even when the drugs are specific, some of the enzymes that the inhibitors target also play a role in myriad pathways within the mammalian cell, and their complete biological functions are not fully elucidated (Deu, 2012).

In the case of Prp, at least one exact biological function and the consequences of abrogation of that function are known. Specifically, we know that if *S. aureus* L27 does not get cleaved, the bacterium dies. Inhibition of this enzyme should kill those bacteria that employ it to cleave L27. For those bacteria that have the cleavage motif on FlhB instead of L27, it is possible that their flagellar motility could be impaired. However, the importance of the cleavage motif to flagellar assembly is only theoretical at present. Prp also represents a unique protein fold, with no significant structural homology to any studied human proteins, which should make it a truly worthwhile drug target.

The current most pressing problem with protease inhibitors as drugs is that they often lack specificity. There are many of the seven types of protease in the human body that all work using the seven conserved mechanisms, so off-target effects of a particular protease inhibitor can have wide-ranging physiological consequences. For this reason, protease inhibitors that covalently modify the active site of a cysteine or serine protease using a reactive "warhead" are often deleterious to humans. One would hope that Prp is small enough and cleaves such a specific sequence that it could still be targeted in this way, because covalent modification is a very potent mechanism of inhibition. However, it is important to consider other avenues of drug design that would allow for greater specificity.

Competitive inhibition via substrate analog is theoretically possible. In this approach to inhibiting Prp activity, a non-hydrolyzable peptidomimetic "dummy substrate" could be developed containing modified residues, beta-linked amino acids or D-amino acids. These molecules would be required to bind Prp with an affinity orders of magnitude greater than that of the native target to be practically considered as drugs. There are few data on the human administration of peptidomimetics containing D-amino acids and beta linkages, but early trials suggest that they are not intrinsically toxic (Vlieghe, 2010). Peptidomimetics have not yet been widely used as drugs, largely due to their high lability in the human gastrointestinal tract that precludes oral administration, and their high rate of clearance and degradation in the bloodstream. These factors, while not absolutely insurmountable, make this avenue of drug design the most problematic (Weinstock, 2012).

It appears that Prp can take many non-catalytic conformations, as evidenced by the fact that none of the four Prp crystal structures that exist would be capable of catalysis due to the orientation of their active site residues. In the Prp family, the catalytic His is always preceded by a Gly, forcing it onto a flexible loop that appears to be able to swing outward, away from the catalytic Cys. This likely has functional relevance in substrate binding and specific catalysis. It remains to be seen whether Prp substrate binding relies on induced fit or conformational selection - this information would aid in the design of possible conformational inhibitors. It may be that a tight-binding inhibitor could simply stabilize extant non-catalytic conformations of the enzyme. For example, the right inhibitor could sterically hinder the movement of the active site histidine side chain, forcing it away from the catalytic cysteine, as its conformation in structures PDB ID 2G0I and 2IDL.

In a hybrid approach, if it is discovered that hydrolysis of the acyl-enzyme intermediate is the rate limiting step in the Prp mechanism, it may be possible to stabilize that covalent enzyme-substrate complex. This is called uncompetitive inhibition and it would prevent regeneration of the Prp catalytic center by preventing Prp-L27 dissociation. In this way the enzyme active site remains covalently modified by its native substrate without the introduction of a reactive warhead drug that could have off-target effects (Deu, 2012). Our studies of catalytically inactive PrpC34A overexpression demonstrated the capacity of a non-covalent L27-PrpC34A complex to radically inhibit *S. aureus* growth and ribosome production, so this approach to Prp inhibitor design seems promising.

Further investigation of Prp kinetic properties, including discovery of the rate limiting step, are required and ongoing.

In light of the rise in antibiotic resistance in *S. aureus* and other Gram-positive pathogens, such as *Clostridium difficile* and *Enterococcus faecium*, and the paucity of novel antibiotics in the development pipeline (Gould,2009; Hughes *et al.*, 2014), the findings reported here lay the foundation for the development of new therapeutic approaches that interfere with protein synthesis by targeting the processing of L27.

References

Akanuma, G., Nanamiya, H., Natori, Y., Yano, K., Suzuki, S., Omata, S., . . . Kawamura, F. (2012). Inactivation of ribosomal protein genes in bacillus subtilis reveals importance of each ribosomal protein for cell proliferation and cell differentiation. *Journal of Bacteriology, 194*(22), 6282-6291. doi:10.1128/JB.01544-12 [doi]

Aksyuk, A. A., & Rossmann, M. G. (2011). Bacteriophage assembly. *Viruses, 3*(3), 172-203. doi:10.3390/v3030172

Anderson, K. L., Roberts, C., Disz, T., Vonstein, V., Hwang, K., Overbeek, R., . . . Dunman, P. M. (2006). Characterization of the staphylococcus aureus heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *Journal of Bacteriology, 188*(19), 6739-6756. doi:10.1128/JB.00609-06

Anderson, K. L., Roux, C. M., Olson, M. W., Luong, T. T., Lee, C. Y., Olson, R., & Dunman, P. M. (2010). Characterizing the effects of inorganic acid and alkaline shock on the staphylococcus aureus transcriptome and messenger RNA turnover. *FEMS Immunology and Medical Microbiology, 60*(3), 208-250. doi:10.1111/j.1574-695X.2010.00736.x; 10.1111/j.1574-695X.2010.00736.x

Arnold, R. J., & Reilly, J. P. (1999). Observation of escherichia coli ribosomal proteins and their posttranslational modifications by mass spectrometry. *Analytical Biochemistry, 269*(1), 105-112. doi:S0003-2697(98)93077-9 [pii]

Auerbach, T., Bashan, A., Harms, J., Schluenzen, F., Zarivach, R., Bartels, H., . . . Yonath, A. (2002). Antibiotics targeting ribosomes: Crystallographic studies. *Current Drug Targets.Infectious Disorders, 2*(2), 169-186.

Ban, N., Beckmann, R., Cate, J. H., Dinman, J. D., Dragon, F., Ellis, S. R., . . . Yusupov, M. (2014). A new system for naming ribosomal proteins. *Current Opinion in Structural Biology, 24*, 165-169. doi:10.1016/j.sbi.2014.01.002 [doi]

Bashan, A., & Yonath, A. (2008). Correlating ribosome function with high-resolution structures. *Trends in Microbiology, 16*(7), 326-335. doi:10.1016/j.tim.2008.05.001

Becher, D., Hempel, K., Sievers, S., Zuhlke, D., Pane-Farre, J., Otto, A., . . . Hecker, M. (2009). A proteomic view of an important human pathogen--towards the quantification of the entire staphylococcus aureus proteome. *PloS One, 4*(12), e8176. doi:10.1371/journal.pone.0008176; 10.1371/journal.pone.0008176

Belousoff, M. J., Davidovich, C., Zimmerman, E., Caspi, Y., Wekselman, I., Rozenszajn, L., . . . Yonath, A. (2010). Ancient machinery embedded in the contemporary ribosome. *Biochemical Society Transactions, 38*(2), 422-427. doi:10.1042/BST0380422; 10.1042/BST0380422

Castro, H. C., Abreu, P. A., Geraldo, R. B., Martins, R. C., dos Santos, R., Loureiro, N. I., . . . Rodrigues, C. R. (2011). Looking at the proteases from a simple perspective. *Journal of Molecular Recognition : JMR, 24*(2), 165-181. doi:10.1002/jmr.1091 [doi]

Cech, T. R. (2000). Structural biology. the ribosome is a ribozyme. *Science (New York, N.Y.), 289*(5481), 878-879.

Chaudhuri, R. R., Allen, A. G., Owen, P. J., Shalom, G., Stone, K., Harrison, M., . . . Charles, I. G. (2009). Comprehensive identification of essential staphylococcus aureus genes using transposon-mediated differential hybridisation (TMDH). *BMC Genomics, 10*, 291-2164-10-291. doi:10.1186/1471-2164-10-291; 10.1186/1471-2164-10-291

Chirgadze, Y. N., Clarke, T. E., Romanov, V., Kisselman, G., Wu-Brown, J., Soloveychik, M., . . . Chirgadze, N. Y. (2015). The structure of SAV1646 from staphylococcus aureus belonging to a new `ribosome-associated' subfamily of bacterial proteins. *Acta Crystallographica.Section D, Biological Crystallography, 71*(Pt 2), 332-337. doi:10.1107/S1399004714025619 [doi]

Colca, J. R., McDonald, W. G., Waldon, D. J., Thomasco, L. M., Gadwood, R. C., Lund, E. T., . . . Mankin, A. S. (2003). Cross-linking in the living cell locates the site of action of oxazolidinone antibiotics. *The Journal of Biological Chemistry, 278*(24), 21972-21979. doi:10.1074/jbc.M302109200

Condon, C., Brechemier-Baey, D., Beltchev, B., Grunberg-Manago, M., & Putzer, H. (2001). Identification of the gene encoding the 5S ribosomal RNA maturase in bacillus subtilis: Mature 5S rRNA is dispensable for ribosome function. *RNA (New York, N.Y.), 7*(2), 242-253.

D'Elia, M. A., Pereira, M. P., Chung, Y. S., Zhao, W., Chau, A., Kenney, T. J., . . . Brown, E. D. (2006). Lesions in teichoic acid biosynthesis in staphylococcus aureus lead to a lethal gain of function in the otherwise dispensable pathway. *Journal of Bacteriology, 188*(12), 4183-4189. doi:10.1128/JB.00197-06

Deu, E., Verdoes, M., & Bogyo, M. (2012). New approaches for dissecting protease functions to improve probe development and drug discovery. *Nature Structural & Molecular Biology, 19*(1), 9-16. doi:10.1038/nsmb.2203 [doi]

Duda, R. L., Oh, B., & Hendrix, R. W. (2013). Functional domains of the HK97 capsid maturation protease and the mechanisms of protein encapsidation. *Journal of Molecular Biology, 425*(15), 2765-2781. doi:10.1016/j.jmb.2013.05.002 [doi]

Fetrow, J. S. (2006). Active site profiling to identify protein functional sites in sequences and structures using the deacon active site profiler (DASP). *Current Protocols in Bioinformatics / Editoral Board, Andreas D.Baxevanis ...[Et Al.], Chapter 8*, Unit 8.10. doi:10.1002/0471250953.bi0810s14 [doi]

Fox, G. E. (2010). Origin and evolution of the ribosome. *Cold Spring Harbor Perspectives in Biology, 2*(9), a003483. doi:10.1101/cshperspect.a003483; 10.1101/cshperspect.a003483

Frank, J., & Agrawal, R. K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature, 406*(6793), 318-322. doi:10.1038/35018597 [doi]

Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., . . . Venter, J. C. (1995). The minimal gene complement of mycoplasma genitalium. *Science (New York, N.Y.), 270*(5235), 397-403.

Fuchs, S., Pane-Farre, J., Kohler, C., Hecker, M., & Engelmann, S. (2007). Anaerobic gene expression in staphylococcus aureus. *Journal of Bacteriology, 189*(11), 4275-4289. doi:10.1128/JB.00081-07

Fuchs, S., Zuhlke, D., Pane-Farre, J., Kusch, H., Wolf, C., Reiss, S., . . . Engelmann, S. (2013). Aureolib - a proteome signature library: Towards an understanding of staphylococcus aureus pathophysiology. *PloS One, 8*(8), e70669. doi:10.1371/journal.pone.0070669; 10.1371/journal.pone.0070669

Garcia-Vallve, S., Simo, F. X., Montero, M. A., Arola, L., & Romeu, A. (2002). Simultaneous horizontal gene transfer of a gene coding for ribosomal protein l27 and operational genes in arthrobacter sp. *Journal of Molecular Evolution, 55*(6), 632-637. doi:10.1007/s00239-002-2358-5

Geiger, T., & Wolz, C. (2014). Intersection of the stringent response and the CodY regulon in low GC gram-positive bacteria. *International Journal of Medical Microbiology : IJMM, 304*(2), 150-155. doi:10.1016/j.ijmm.2013.11.013 [doi]

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A.,3rd, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods, 6*(5), 343-345. doi:10.1038/nmeth.1318; 10.1038/nmeth.1318

Glass, J. I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., . . . Venter, J. C. (2006). Essential genes of a minimal bacterium. *Proceedings of the National Academy of Sciences of the United States of America, 103*(2), 425-430. doi:10.1073/pnas.0510013103

Goto, S., Muto, A., & Himeno, H. (2013). GTPases involved in bacterial ribosome maturation. *Journal of Biochemistry, 153*(5), 403-414. doi:10.1093/jb/mvt022 [doi]

Gould, I. M. (2009). Antibiotic resistance: The perfect storm. *International Journal of Antimicrobial Agents, 34 Suppl 3*, S2-5. doi:10.1016/S0924-8579(09)70549-7; 10.1016/S0924-8579(09)70549-7

Gulati, M., Jain, N., Davis, J. H., Williamson, J. R., & Britton, R. A. (2014). Functional interaction between ribosomal protein L6 and RbgA during ribosome assembly. *PLoS Genetics, 10*(10), e1004694. doi:10.1371/journal.pgen.1004694 [doi]

Gutgsell, N. S., & Jain, C. (2012). Gateway role for rRNA precursors in ribosome assembly. *Journal of Bacteriology, 194*(24), 6875-6882. doi:10.1128/JB.01467-12 [doi]

Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., . . . Yonath, A. (2001). High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell, 107*(5), 679-688. doi:S0092-8674(01)00546-3 [pii]

Hughes, D., & Karlen, A. (2014). Discovery and preclinical development of new antibiotics. *Upsala Journal of Medical Sciences,* doi:10.3109/03009734.2014.896437

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics, 14*(1), 33-8, 27-8. doi:0263785596000185 [pii]

Iost, I., Bizebard, T., & Dreyfus, M. (2013). Functions of DEAD-box proteins in bacteria: Current knowledge and pending questions. *Biochimica Et Biophysica Acta, 1829*(8), 866-877. doi:10.1016/j.bbagrm.2013.01.012 [doi]

Ji, Y., Zhang, B., Van, S. F., Horn, Warren, P., Woodnutt, G., . . . Rosenberg, M. (2001). Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA. *Science (New York, N.Y.), 293*(5538), 2266-2269. doi:10.1126/science.1063566

Jomaa, A., Jain, N., Davis, J. H., Williamson, J. R., Britton, R. A., & Ortega, J. (2014). Functional domains of the 50S subunit mature late in the assembly process. *Nucleic Acids Research, 42*(5), 3419-3435. doi:10.1093/nar/gkt1295 [doi]

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., . . . Drummond, A. (2012). Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England), 28*(12), 1647-1649. doi:10.1093/bioinformatics/bts199; 10.1093/bioinformatics/bts199

Klein, E., Smith, D. L., & Laxminarayan, R. (2007). Hospitalizations and deaths caused by methicillin-resistant staphylococcus aureus, united states, 1999-2005. *Emerging Infectious Diseases, 13*(12), 1840-1846. doi:10.3201/eid1312.070629; 10.3201/eid1312.070629

Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., . . . Ogasawara, N. (2003). Essential bacillus subtilis genes. *Proceedings of the National Academy of Sciences of the United States of America, 100*(8), 4678-4683. doi:10.1073/pnas.0730515100

Kokkori, S., Apostolidi, M., Tsakris, A., Pournaras, S., Stathopoulos, C., & Dinos, G. (2014). Linezolid-dependent function and structure adaptation of ribosomes in a staphylococcus epidermidis strain exhibiting linezolid dependence. *Antimicrobial Agents and Chemotherapy, 58*(8), 4651-4656. doi:10.1128/AAC.02835-14 [doi]

Kreiswirth, B. N., Lofdahl, S., Betley, M. J., O'Reilly, M., Schlievert, P. M., Bergdoll, M. S., & Novick, R. P. (1983). The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature, 305*(5936), 709-712.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., . . . Danchin, A. (1997). The complete genome sequence of the gram-positive bacterium bacillus subtilis. *Nature, 390*(6657), 249-256. doi:10.1038/36786

Lalonde, M. S., Zuo, Y., Zhang, J., Gong, X., Wu, S., Malhotra, A., & Li, Z. (2007). Exoribonuclease R in mycoplasma genitalium can carry out both RNA processing and degradative functions and is sensitive to RNA ribose methylation. *RNA (New York, N.Y.), 13*(11), 1957-1968. doi:rna.706207 [pii]

Lauber, M. A., Running, W. E., & Reilly, J. P. (2009). B. subtilis ribosomal proteins: Structural homology and post-translational modifications. *Journal of Proteome Research, 8*(9), 4193-4206. doi:10.1021/pr801114k

Leach, K. L., Swaney, S. M., Colca, J. R., McDonald, W. G., Blinn, J. R., Thomasco, L. M., . . . Mankin, A. S. (2007). The site of action of oxazolidinone antibiotics in living bacteria and in human mitochondria. *Molecular Cell, 26*(3), 393-402. doi:10.1016/j.molcel.2007.04.005

Lecompte, O., Ripp, R., Thierry, J. C., Moras, D., & Poch, O. (2002). Comparative analysis of ribosomal proteins in complete genomes: An example of reductive evolution at the domain scale. *Nucleic Acids Research, 30*(24), 5382-5390.

Lee, C. Y., Buranen, S. L., & Ye, Z. H. (1991). Construction of single-copy integration vectors for staphylococcus aureus. *Gene, 103*(1), 101-105.

Li, N., Chen, Y., Guo, Q., Zhang, Y., Yuan, Y., Ma, C., . . . Gao, N. (2013). Cryo-EM structures of the late-stage assembly intermediates of the bacterial 50S ribosomal subunit. *Nucleic Acids Research, 41*(14), 7073-7083. doi:10.1093/nar/gkt423 [doi]

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., . . . Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry.B, 102*(18), 3586-3616. doi:10.1021/jp973084f [doi]

Maguire, B. A., Beniaminov, A. D., Ramu, H., Mankin, A. S., & Zimmermann, R. A. (2005). A protein component at the heart of an RNA machine: The importance of protein l27 for the function of the bacterial ribosome. *Molecular Cell, 20*(3), 427-435. doi:10.1016/j.molcel.2005.09.009

Maguire, B. A., Manuilov, A. V., & Zimmermann, R. A. (2001). Differential effects of replacing escherichia coli ribosomal protein L27 with its homologue from aquifex aeolicus. *Journal of Bacteriology, 183*(22), 6565-6572. doi:10.1128/JB.183.22.6565-6572.2001 [doi]

Martin, A. C., Lopez, R., & Garcia, P. (1998). Pneumococcal bacteriophage cp-1 encodes its own protease essential for phage maturation. *Journal of Virology, 72*(4), 3491-3494.

Mikulik, K., Bobek, J., Zikova, A., Smetakova, M., & Bezouskova, S. (2011). Phosphorylation of ribosomal proteins influences subunit association and translation of poly (U) in streptomyces coelicolor. *Molecular bioSystems, 7*(3), 817-823. doi:10.1039/c0mb00174k; 10.1039/c0mb00174k

Minamino, T. (2014). Protein export through the bacterial flagellar type III export pathway. *Biochimica Et Biophysica Acta, 1843*(8), 1642-1648. doi:10.1016/j.bbamcr.2013.09.005 [doi]

Neron, B., Menager, H., Maufrais, C., Joly, N., Maupetit, J., Letort, S., . . . Letondal, C. (2009). Mobyle: A new full web bioinformatics framework. *Bioinformatics (Oxford, England), 25*(22), 3005-3011. doi:10.1093/bioinformatics/btp493; 10.1093/bioinformatics/btp493

Netterling, S., Vaitkevicius, K., Nord, S., & Johansson, J. (2012). A listeria monocytogenes RNA helicase essential for growth and ribosomal maturation at low temperatures uses its C terminus for appropriate interaction with the ribosome. *Journal of Bacteriology, 194*(16), 4377-4385. doi:10.1128/JB.00348-12 [doi]

Nierhaus, K. H. (1991). The assembly of prokaryotic ribosomes. *Biochimie, 73*(6), 739-755. doi:0300-9084(91)90054-5 [pii]

Nierhaus, K. H., & Dohme, F. (1974). Total reconstitution of functionally active 50S ribosomal subunits from escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America, 71*(12), 4713-4717.

Nissen, P., Hansen, J., Ban, N., Moore, P. B., & Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science (New York, N.Y.), 289*(5481), 920-930. doi:8743 [pii]

O'Farrell, H. C., & Rife, J. P. (2012). Staphylococcus aureus and escherichia coli have disparate dependences on KsgA for growth and ribosome biogenesis. *BMC Microbiology, 12*, 244-2180-12-244. doi:10.1186/1471-2180-12-244; 10.1186/1471-2180-12-244

Oh, B., Moyer, C. L., Hendrix, R. W., & Duda, R. L. (2014). The delta domain of the HK97 major capsid protein is essential for assembly. *Virology, 456-457*, 171-178. doi:10.1016/j.virol.2014.03.022 [doi]

Olsson, M. H. M., Søndergaard, C.,R., Rostkowski, M., & Jensen, J. H. (2011). ***PROPKA3: Consistent treatment of internal and surface residues in empirical pK$_a$ predictions*** doi:10.1021/ct100578z

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry, 25*(13), 1605-1612. doi:10.1002/jcc.20084

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., . . . Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry, 26*(16), 1781-1802. doi:10.1002/jcc.20289 [doi]

Poliakov, A., Chang, J. R., Spilman, M. S., Damle, P. K., Christie, G. E., Mobley, J. A., & Dokland, T. (2008). Capsid size determination by staphylococcus aureus pathogenicity island SaPI1 involves specific incorporation of SaPI1 proteins into procapsids. *Journal of Molecular Biology, 380*(3), 465-475. doi:10.1016/j.jmb.2008.04.065; 10.1016/j.jmb.2008.04.065

Polikanov, Y. S., Steitz, T. A., & Innis, C. A. (2014). A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. *Nature Structural & Molecular Biology, 21*(9), 787-793. doi:10.1038/nsmb.2871 [doi]

Rawlings, N. D., Barrett, A. J., & Bateman, A. (2011). Asparagine peptide lyases: A seventh catalytic type of proteolytic enzymes. *The Journal of Biological Chemistry, 286*(44), 38321-38328. doi:10.1074/jbc.M111.260026 [doi]

Redko, Y., & Condon, C. (2010). Maturation of 23S rRNA in bacillus subtilis in the absence of mini-III. *Journal of Bacteriology, 192*(1), 356-359. doi:10.1128/JB.01096-09 [doi]

Rossi, F., Diaz, L., Wollam, A., Panesso, D., Zhou, Y., Rincon, S., . . . Arias, C. A. (2014). Transferable vancomycin resistance in a community-associated MRSA lineage. *The New England Journal of Medicine, 370*(16), 1524-1531. doi:10.1056/NEJMoa1303359 [doi]

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols, 5*(4), 725-738. doi:10.1038/nprot.2010.5; 10.1038/nprot.2010.5

Sagan, L. (1967). On the origin of mitosing cells. *Journal of Theoretical Biology, 14*(3), 255-274.

Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., & Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. *Proteins, 23*(3), 318.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning:A laboratory manual* (2nd ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Scherl, A., Francois, P., Bento, M., Deshusses, J. M., Charbonnier, Y., Converset, V., . . . Schrenzel, J. (2005). Correlation of proteomic and transcriptomic profiles of staphylococcus aureus during the post-exponential phase of growth. *Journal of Microbiological Methods, 60*(2), 247-257. doi:10.1016/j.mimet.2004.09.017

Selmer, M., Dunham, C. M., Murphy, F. V.,4th, Weixlbaumer, A., Petry, S., Kelley, A. C., . . . Ramakrishnan, V. (2006). Structure of the 70S ribosome complexed with mRNA and tRNA. *Science (New York, N.Y.), 313*(5795), 1935-1942. doi:10.1126/science.1131127

Shajani, Z., Sykes, M. T., & Williamson, J. R. (2011). Assembly of bacterial ribosomes. *Annual Review of Biochemistry, 80*, 501-526. doi:10.1146/annurev-biochem-062608-160432; 10.1146/annurev-biochem-062608-160432

Shin, D. H., Lou, Y., Jancarik, J., Yokota, H., Kim, R., & Kim, S. H. (2005). Crystal structure of TM1457 from thermotoga maritima. *Journal of Structural Biology, 152*(2), 113-117. doi:10.1016/j.jsb.2005.08.008

Shoji, S., Dambacher, C. M., Shajani, Z., Williamson, J. R., & Schultz, P. G. (2011). Systematic chromosomal deletion of bacterial ribosomal protein genes. *Journal of Molecular Biology, 413*(4), 751-761. doi:10.1016/j.jmb.2011.09.004; 10.1016/j.jmb.2011.09.004

Spilman, M. S., Damle, P. K., Dearborn, A. D., Rodenburg, C. M., Chang, J. R., Wall, E. A., . . . Dokland, T. (2012). Assembly of bacteriophage 80alpha capsids in a staphylococcus aureus expression system. *Virology, 434*(2), 242-250. doi:10.1016/j.virol.2012.08.031; 10.1016/j.virol.2012.08.031

Sykes, M. T., Shajani, Z., Sperling, E., Beck, A. H., & Williamson, J. R. (2010). Quantitative proteomic analysis of ribosome assembly and turnover in vivo. *Journal of Molecular Biology, 403*(3), 331-345. doi:10.1016/j.jmb.2010.08.005 [doi]

Tejedor, F., & Ballesta, J. P. (1986). Reaction of some macrolide antibiotics with the ribosome. labeling of the binding site components. *Biochemistry, 25*(23), 7725-7731.

Trobro, S., & Aqvist, J. (2008). Role of ribosomal protein L27 in peptidyl transfer. *Biochemistry, 47*(17), 4898-4906. doi:10.1021/bi8001874; 10.1021/bi8001874

Vlieghe, P., Lisowski, V., Martinez, J., & Khrestchatisky, M. (2010). Synthetic therapeutic peptides: Science and market. *Drug Discovery Today, 15*(1-2), 40-56. doi:10.1016/j.drudis.2009.10.009 [doi]

Voorhees, R. M., Weixlbaumer, A., Loakes, D., Kelley, A. C., & Ramakrishnan, V. (2009). Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nature Structural & Molecular Biology, 16*(5), 528-533. doi:10.1038/nsmb.1577; 10.1038/nsmb.1577

Wall, E. A., Caufield, J. H., Lyons, C. E., Manning, K. A., Dokland, T., & Christie, G. E. (2015). Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. *Molecular Microbiology, 95*(2), 258-269. doi:10.1111/mmi.12862 [doi]

Wang, H., Takemoto, C. H., Murayama, K., Sakai, H., Tatsuguchi, A., Terada, T., . . . Yokoyama, S. (2004). Crystal structure of ribosomal protein L27 from thermus thermophilus HB8. *Protein Science : A Publication of the Protein Society, 13*(10), 2806-2810. doi:10.1110/ps.04864904

Wang, J., Dasgupta, I., & Fox, G. E. (2009). Many nonuniversal archaeal ribosomal proteins are found in conserved gene clusters. *Archaea (Vancouver, B.C.), 2*(4), 241-251.

Wang, L., Trawick, J. D., Yamamoto, R., & Zamudio, C. (2004). Genome-wide operon prediction in staphylococcus aureus. *Nucleic Acids Research, 32*(12), 3689-3702. doi:10.1093/nar/gkh694

Wang, Y., & Xiao, M. (2012). Role of the ribosomal protein L27 revealed by single-molecule FRET study. *Protein Science : A Publication of the Protein Society, 21*(11), 1696-1704. doi:10.1002/pro.2149; 10.1002/pro.2149

Weinstock, M. T., Francis, J. N., Redman, J. S., & Kay, M. S. (2012). Protease-resistant peptide design-empowering nature's fragile warriors against HIV. *Biopolymers, 98*(5), 431-442. doi:10.1002/bip.22073 [doi]

Williamson, J. R. (2008). Biophysical studies of bacterial ribosome assembly. *Current Opinion in Structural Biology, 18*(3), 299-304. doi:10.1016/j.sbi.2008.05.001 [doi]

Wilson, D. N., & Nierhaus, K. H. (2007). The weird and wonderful world of bacterial ribosome regulation. *Critical Reviews in Biochemistry and Molecular Biology, 42*(3), 187-219. doi:10.1080/10409230701360843

Wilson, D. N., Schluenzen, F., Harms, J. M., Starosta, A. L., Connell, S. R., & Fucini, P. (2008). The oxazolidinone antibiotics perturb the ribosomal peptidyl-transferase center and effect tRNA positioning. *Proceedings of the National Academy of Sciences of the United States of America, 105*(36), 13339-13344. doi:10.1073/pnas.0804276105; 10.1073/pnas.0804276105

Wolz, C., Geiger, T., & Goerke, C. (2010). The synthesis and function of the alarmone (p)ppGpp in firmicutes. *International Journal of Medical Microbiology : IJMM, 300*(2-3), 142-147. doi:10.1016/j.ijmm.2009.08.017 [doi]

Wower, I. K., Wower, J., & Zimmermann, R. A. (1998). Ribosomal protein L27 participates in both 50 S subunit assembly and the peptidyl transferase reaction. *The Journal of Biological Chemistry, 273*(31), 19847-19852.

Wower, J., Hixson, S. S., & Zimmermann, R. A. (1989). Labeling the peptidyltransferase center of the escherichia coli ribosome with photoreactive tRNA(phe) derivatives containing azidoadenosine at the 3' end of the acceptor arm: A model of the tRNA-ribosome complex. *Proceedings of the National Academy of Sciences of the United States of America, 86*(14), 5232-5236.

Wower, J., Kirillov, S. V., Wower, I. K., Guven, S., Hixson, S. S., & Zimmermann, R. A. (2000). Transit of tRNA through the escherichia coli ribosome. cross-linking of the 3' end of tRNA to specific nucleotides of the 23 S ribosomal RNA at the A, P, and E sites. *The Journal of Biological Chemistry, 275*(48), 37887-37894. doi:10.1074/jbc.M005031200

Wower, J., Wower, I. K., Kirillov, S. V., Rosen, K. V., Hixson, S. S., & Zimmermann, R. A. (1995). Peptidyl transferase and beyond. *Biochemistry and Cell Biology = Biochimie Et Biologie Cellulaire, 73*(11-12), 1041-1047.

Xiao, M., & Wang, Y. (2012). L27-tRNA interaction revealed by mutagenesis and pH titration. *Biophysical Chemistry, 167*, 8-15. doi:10.1016/j.bpc.2012.04.003; 10.1016/j.bpc.2012.04.003

Yonath, A. (2005). Antibiotics targeting ribosomes: Resistance, selectivity, synergism and cellular regulation. *Annual Review of Biochemistry, 74*, 649-679. doi:10.1146/annurev.biochem.74.082803.133130

Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H., & Noller, H. F. (2001). Crystal structure of the ribosome at 5.5 A resolution. *Science (New York, N.Y.), 292*(5518), 883-896. doi:10.1126/science.1060089 [doi]

Yutin, N., PuigbÃ², P., Koonin, E. V., & Wolf, Y. I. (2012). Phylogenomics of prokaryotic ribosomal proteins. *Plos One, 7*(5), e36972. Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0036972

Appendices

Table 1. Plasmids and Bacterial Strains.

Plasmid	Description	Reference
pBADmycHisA	*E. coli* expression vector with arabinose-inducible promoter P_{BAD}; produces fusion proteins with C-terminal Myc and His_6 tags	Invitrogen
pET21a	*E. coli* T7 expression vector	Invitrogen
pG164	*S. aureus* T7 expression vector	D'Elia *et al*, 2006
pMAD	*E. coli/S.aureus* shuttle vector for allelic exchange	Arnaud, 2004
pRSFduet	*E. coli* T7 expression vector with RSF1030 replicon, compatible with pBADmycHisA	Novagen
pRW	pET21a derivative; N-terminal 6xHis SUMO tag added	Darrell Peterson
pT104	P_{Ars} arsenite-inducible expression in *S. aureus*	Liu, 2004
pTB	pET21a derivative; Added N-terminal 6xHis tag and TEV cleavage site	Darrell Peterson
pEW5	pBADmycHisA derivative; *S. aureus* L27 $MycHis_6$	This work
pEW7	pG164 derivative; *S. aureus* L27 $MycHis_6$	This work
pEW15	pRSFduet derivative; *S. aureus* Prp	This work
pEW21	pG164 derivative; *S. aureus* L27 Q7I $MycHis_6$	This work
pEW22	pG164 derivative; *S. aureus* L27 F9Y $MycHis_6$	This work
pEW23	pG164 derivative; *S. aureus* L27 A10S $MycHis_6$	This work
pEW25	pG164 derivative; *S. aureus* Prp	This work
pEW26	pG164 derivative; Prp C34A	This work
pEW27	pG164 derivative; *S. aureus* L27	This work

pEW28	pRSFduet derivative; *S. aureus* Prp C34A	This work
pEW29	pG164 derivative; pre-cleaved L27	This work
pEW34	pRW derivative; HisSUMO-ysxB	This work
pEW35	pRSFduet derivative; *S. aureus* Prp C34S	This work
pEW37	pG164 derivative; L27 F8A-F9A MycHis$_6$ (un-cleavable)	This work
pEW39	pRW derivative; HisSUMO-ysxBC34A	This work
pEW40	pRW derivative; HisSUMO-ysxBC34S	This work
pEW41	pG164 derivative; L27 F8A-F9A (un-cleavable)	This work
pEW43	pG164 derivative; MLKLNLQFF expression	This work
pEW50	pMAD derivative; allelic exchange vector for deletion of the L27 N-terminal extension	This work
pEW53	pTB derivative; HisTEVL27 in EcoRI/ XhoI	This work
pEW56	pRSFduet derivative; *S. aureus* PrpV40P	This work
pEW57	pRSFduet derivative; *S. aureus* PrpN46A	This work
pEW63	pTB derivative; PrpHis6x in NcoI/ XhoI	This work
pEW68	pMAD derivative; allelic exchange vector for deletion of L27 using a spectinomycin resistance cassette from pCN55	This work, Charpentier, 2004
pEW72	pT104 derivative; L27 expression in *S. aureus*	This work
pEW73	pT104 derivative; L27 Δ2-9 expression in *S. aureus*	This work
pEW74	pT104 derivative; L27 FF::AA expression in *S. aureus*	This work
pEW75	pT104 derivative; L27 Δ2-9 and L27 1-9 peptide expression in *S. aureus*	This work
pEW76	pG164 derivative; pre-cleaved L27 His6 expression in *S. aureus*	This work

Strains	Description	Source
Stellar Competent *E. coli* DH5α	*F–, endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80d lacZΔ M15, Δ (lacZYA - argF) U169, Δ (mrr - hsdRMS - mcrBC), ΔmcrA, λ–*	Clontech
E. coli BL21 DE3 RIL	*E. coli* B F- *ompT* hsdS(rB- mB-) *dcm*+ Tetr *E. coli* gal λ (DE3) *endA* Hte [*argU ileY leuW* Camr]	Agilent
S. aureus RN450	NCTC8325 cured of Φ11, Φ12 and Φ13	Novick, 1967
S. aureus RN4220	Restriction defective derivative of RN450	de Azavedo *et al.*, 1985
SA178RI	4220-derived CYL316 containing T7 RNA polymerase, Tet^R 2μg/μL	D'Elia *et al.*, 2006
ST256	SA178RI containing pEW27 which expresses L27 from T7 promoter in the presence of IPTG. Tet^R 2μg/μL, Chl^R 15μg/μL	This work
ST360	SA178RI L27::$Spec^R$ with complementing pEW27. This strain is IPTG-dependent. Tet^R 2μg/μL, Chl^R 15μg/μL, $Spec^R$ 250μg/μL	This work

Table 2. Primers arranged by plasmid construction
The annealing sequence for amplification is underlined and the 5' 15-16 bp homologous overhang required for the In-Fusion reaction is separated from the rest of the primer by an asterisk (*). Restriction sites used to linearize the plasmid prior to assembly are shown in bold in the sequence when they were reconstituted in the finished assembly. Double asterisks (**) indicates that these primers were not used for PCR but were annealed to form a ds DNA fragment that was assembled during the In-Fusion reaction.

Primer	Sequence 5'-3'
pEW5	
FIXEAW28	GTT CGG GCC CAA GCT* <u>TT TCT GCT ACT GCA TAT ACA GAA ACT TG</u>
EAW29	GAG GAA TTA ACC <u>ATG</u> * <u>TTA AAA TTA AAC TTA CAA TTC TTC GCA TCT A</u>
pEW7	
EAW32	GAA AGG AGG T**GG ATC C** * <u>ATG TTA AAA TTA AAC TTA CAA TTC TTC GCA TCT A</u>
EAW33	GCA GGA GCT C**GA ATT C** * TT <u>AAT GAT GAT GAT GAT GAT GGT CGA C</u>
pEW15	
EAW62	AGG AGA TAT ACC <u>ATG</u> *<u>ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA G</u>
EAW63	TTT ACC AGA CTC GAG* <u>TCA CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC</u>
pEW21	
EAW76	TAA GTT TAA TTT TAA CAT GGA TCC ACC TCC
EAW77	(ATT) TTC TTC GCA TCT AAA AAA GGG GTA AG
pEW22	
EAW78	GAA TTG TAA GTT TAA TTT TAA CAT GGA TCC ACC
EAW79	(TAT) GCA TCT AAA AAA GGG GTA AGT TCT ACA AAA AAC
pEW23	
EAW80	GAA GAA TTG TAA GTT TAA TTT TAA CAT GGA TCC ACC
EAW81	(TCA) TCT AAA AAA GGG GTA AGT TCT ACA AAA AAC GG
pEW25	
EAW83	GAA AGG AGG T**GG ATC C** * <u>ATG ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA GG</u>

EAW84	CA GGA GCT C**GA ATTC*** <u>TCA CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC AAT</u>
pEW26	
EAW85	TCC AGC (TGC) <u>AAC GAT* ATC ATG ACC ATA TTC ACC ATG GTC</u>
EAW86	ATC GTT (GCA) <u>GCT GGA* GCT TCA GCT GTA TTG TTT GGT AG</u>
pEW27	
EAW32	GAA AGG AGG T**GG ATC C** * <u>ATG TTA AAA TTA AAC TTA CAA TTC TTC GCA TCT A</u>
EAW87	GCA GGA GCT C**GA ATT** ***C** <u>TTA TTC AGC TAC TGC ATA TAC AGA AAC TTG TTT TTT G</u>
pEW28	
EAW62	AGG AGA TAT ACC <u>ATG *ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA G</u>
EAW85	TCC AGC (TGC) <u>AAC GAT* ATC ATG ACC ATA TTC ACC ATG GTC</u>
EAW86	ATC GTT (GCA) <u>GCT GGA* GCT TCA GCT GTA TTG TTT GGT AG</u>
EAW63	TTT ACC AGA CTC GAG* <u>TCA CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC</u>
pEW29	
EAW88	GAA AGG AGG T**GG ATC C** *ATG <u>GCA TCT AAA AAA GGG GTA AGT TCT ACA AAA AAC G</u>
EAW87	GCA GGA GCT C**GA ATT** ***C** <u>TTA TTC AGC TAC TGC ATA TAC AGA AAC TTG TTT TTT G</u>
pEW34	
FIXEAW101	AGA ACA GAT TGG AGG* C <u>ATG ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA GG</u>
EAW102	GTG CGG CCG CAA GCT* <u>TCA CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC AAT</u>
pEW35	
EAW62	AGG AGA TAT ACC <u>ATG *ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA G</u>
EAW105	TCC AGC (ACT) <u>AAC GAT* ATC ATG ACC ATA TTC ACC ATG GTC</u>
EAW106	ATC GTT (AGT) <u>GCT GGA* GCT TCA GCT GTA TTG TTT GGT AG</u>
EAW63	TTT ACC AGA CTC GAG* <u>TCA CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC</u>
pEW37	
EAW113**	GAA AGG AGG T**GG ATC C** ATG TTA AAA TTA AAC TTA CAA GCA GCA GCA TCT

EAW114**	AGA TGC TGC TGC TTG TAA GTT TAA TTT TAA CAT **GGA TCC** ACC TCC TTT C
EAW115	CAA GCA GCA GCA TCT * <u>AAA AAA GGG GTA AGT TCT ACA AAA AAC G</u>
EAW33	GCA GGA GCT C**GA ATT C** * TT <u>AAT GAT GAT GAT GAT GAT GGT CGA C</u>
pEW39	
FIXEAW101	AGA ACA GAT TGG AGG* C <u>ATG ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA GG</u>
EAW102	GTG CGG CCG CAA GCT* <u>TCA CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC AAT</u>
pEW40	
FIXEAW101	AGA ACA GAT TGG AGG* C <u>ATG ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA GG</u>
EAW102	GTG CGG CCG CAA GCT* <u>TCA CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC AAT</u>
pEW41	
EAW89	<u>GAA AGG AGG T**GG ATC C** * ATG TTA AAA TTA AAC</u>
EAW87	GCA GGA GCT C**GA ATT** *C <u>TTA TTC AGC TAC TGC ATA TAC AGA AAC TTG TTT TTT G</u>
pEW43	
EAW139**	GAA AGG AGG TGG ATC * ATG TTA AAA TTA AAC TTA CAA TTC TTC TAA *AA TTC GAG CTC CTG C
EAW140**	GCA GGA GCT CGA ATT* TTA GAA GAA TTG TAA GTT TAA TTT TAA CAT* GAT CCA CCT CCT TTC
pEW50	
EAW45	CGA TGC ATG **CCA TGG** *<u>TAT ACA GGA GGT GCA AAG TAT GTT TGC TAT TAT TG</u>
EAW145	CCC TTT TTT AGA TGC * <u>CAT CGG AAT GCA CCT CAC TTA TAA TTT AAT C</u>
EAW146	<u>GCA TCT AAA AAA GGG GTA AGT TCT ACA AAA AAC GG</u>
EAW48	GGC GAT ATC **GGA TCC** *<u>TTT ACC ACC GTC ACC GCC</u>
pEW53	
EAW163	CCA GGG ATC C**GA ATT***C <u>TTA AAA TTA AAC TTA CAA TTC TTC GCA TCT AAA AAA GG</u>
EAW164	GGT GGT GGT G**CT CGA***G <u>TTA TTC AGC TAC TGC ATA TAC AGA AAC TTG</u>
pEW56	
EAW62	<u>AGG AGA TAT ACC ATG</u> *<u>ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA G</u>

EAW167	AAA CAA T(GG) <u>AGC TGA</u>*AG CTC CAG CAC AAA C
EAW168	TCA GCT (CC)<u>A TTG TTT</u>*GGT AGT GTT AAT GCG ATT ATA GGA TTG
EAW63	TTT ACC AGA CTC GAG* <u>TCA CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC</u>
pEW57	
EAW62	AGG AGA TAT ACC <u>ATG</u> *<u>ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA G</u>
EAW169	AAT CGC (TGC) <u>AAC ACT</u>* ACC AAA CAA TAC AGC TGA AGC
EAW170	AGT GTT (GCA) <u>GCG ATT</u>* ATA GGA TTG ACA TCT GAG AGA CC
EAW63	TTT ACC AGA CTC GAG* <u>TCA CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC</u>
pEW63	
EAW187	AGG AGA TAT ACC <u>ATG</u>* ATT ACT GTT GAT ATT ACA GTT AAT GAT GAA GGC
EAW188	GGT GGT GGT GCT CGA* G <u>CTT ATA ATT TAA TCT AAT ATT CTC ATT ATA TTC TTC TTC AAT AG</u>
pEW68	
EAW45	CGA TGC ATG **CCA TGG** *<u>TAT ACA GGA GGT GCA AAG TAT GTT TGC TAT TAT TG</u>
EAW196	<u>CGG AAT GCA CCT CAC TTA TAA TTT AAT CTA ATA TTC TCA TTA TAT TC</u>
EAW197	GTG AGG TGC ATT CCG * <u>ATC GAA TCC CTT CGG AGC G</u>
EAW198	<u>CCG CGG TAA TAA ACT ATC GAA GGA AC</u>
EAW199	AGT TTA TTA CCG CGG * <u>GCA GTA GCT GAA TAA TTT TGT CTA GTT AAC ACC</u>
EAW48	GGC GAT ATC **GGA TCC** *<u>TTT ACC ACC GTC ACC GCC</u>
pEW72	
EAW217	GTT GAT GAG G<u>GG ATC</u>*C ATG TTA AAA TTA AAC TTA CAA TTC TTC GCA TCT A
EAW218	GCT AAG CTT GGT CGA* C <u>TTA TTC AGC TAC TGC ATA TAC AGA AAC TTG TTT TTT G</u>
pEW73	
EAW223	GTT GAT GAG G<u>GG ATC</u>* C ATG GCA TCT AAA AAA GGG G
EAW218	GCT AAG CTT GGT CGA* C <u>TTA TTC AGC TAC TGC ATA TAC AGA AAC TTG TTT TTT G</u>
pEW74	
EAW224	GTT GAT GAG G<u>GGA TC</u>*C ATG TTA AAA TTA AAC TTA CAA GCA GC

EAW218	GCT AAG CTT GGT CGA* C <u>TTA TTC AGC TAC TGC ATA TAC AGA AAC TTG TTT TTT G</u>
pEW75	
EAW227**	TAA GTC GAC CAA GCT*T AGG AGG TGG ATC ATG TTA AAA TTA AAC TTA CAA TTC TTC TAA A*AGC TTA GCT AGC TAG
EAW228**	CTA GCT AGC TAA GCT*T TTA GAA GAA TTG TAA GTT TAA TTT TAA CAT GAT CCA CCT CCT A*AGC TTG GTC GAC TTA
pEW76	
EAW88	GAA AGG AGG T**GG ATC C** *ATG <u>GCA TCT AAA AAA GGG GTA AGT TCT ACA AAA AAC G</u>
EAW229	GCA GGA GCT C**GA ATT** *C TTA ATG ATG ATG ATG ATG ATG <u>TTC AGC TAC TGC ATA TAC AGA AAC TTG TTT TTT G</u>

Table 3. Relative abundance of ribosomal proteins in the aberrant 50S MS1 peaks obtained upon overexpression of PrpC34A.

 Ribosomal 50S wild type and aberrant subunits were analyzed by mass spectrometry. Tryptic digests of each sample were separated by liquid chromatography and subjected to primary mass spectrometry. The area under the peak for each peptide in the MS1 spectrum (MS1 area) represents the amount of peptide detected in the sample. The MS1 areas for each peptide detected for the indicated ribosomal proteins are tabulated above. The ratio of each peptide from each ribosomal protein to the amount of each peptide from L3 or L4 from the same sample was calculated (i.e. peptides from wt 50S were compared to wild type L3 and L4; peptides from aberrant 50S were compared to aberrant L3 and L4) and then the ratio of aberrant/wt for each peptide was calculated and averaged. The n value indicates the number of individual comparisons (i.e. for L21 there were two unique peptides, each of which was compared to two unique peptides for L3 to give an n of 4). For L24, L32, L33 and L36, where only a single unique tryptic peptide was detected in both samples, we report the range rather than the standard deviation, as the n is only two. In a wild-type ribosome, the proteins should be stoichiometric - a complete 50S subunit should contain one copy of each of the proteins we examined. This was observed for L3 and L4, which are thought to be incorporated early in ribosome assembly – their ratios and most of the other ribosomal protein ratios were the same in the wild type and aberrant 50S peaks. There was a striking three fold or greater decrease in the relative amounts of L18, L20, L16, L27 and L25. Not detected: L9, L28, L34, L35.

Ribosomal protein	Representative Peptides	Wild type 50S peptide peak areas	Aberrant 50S peptide peak areas	Fold ribosomal protein peptide ratio change (Aberrant/ WT)				
				Compared to L3 ((aberrant x/ aberrant L3)/(wt x/ wt L3))	Std dev	Compared to L4 ((aberrant x/ aberrant L4)/(wt x/ wt L4))	Std dev	N
L13	LSSEVASILR QFFMANESNIER	250560344 107986136	746860941 502062870	1.408	0.50	1.315	0.45	4
L22	EAYANEGPTLK VLMSALANAEHNYDMNTDELVVK	72855340 209342247	290447116 696624960	1.350	0.36	1.261	0.32	4
L11	TQDQAGLIIPVEISVYEDR MQDLNAADEEAAMR	109995577 115698823	397904473 376827343	1.270	0.32	1.185	0.28	4
L10	AGIEGLDEFLTGPTAIATSSEDAVAAAK GLTVAEVTDLR	140041746 118973788	445187080 393102301	1.197	0.29	1.117	0.26	4
L4	NVLSTLEQPK AQENGLTVVDAFNFEAPK	336156228 287000000	814015487 1038000000	1.115	0.38	1.000	0.00	4
L3	IGMTQVFGENGELIPVTVVEAK KIGMTQVFGENGELIPVTVVEAK	780091714 173438694	1745552815 594947630	1.000	0.00	0.977	0.32	4
L6	VLELVGVGYR EQVGALASNIR	270806222 308875027	555945068 980926050	0.970	0.34	0.901	0.31	4
L1	SVAVTTTMGPGVK VSFTEQLIENFNTLQDVLAK	134801260 57921310	454820838 82420208	0.886	0.48	0.827	0.44	4
L17	NVEILNEDETTQTALQK DLATSLIISER	235500304 307344293	563020926 615607093	0.811	0.21	0.757	0.19	4
L30	HLVTVEEK TNSSVVVEDNPAIR	50170721 131939444	63484853 404968006	0.800	0.44	0.747	0.40	4
L7/L12	ANHEQIIEAIK EMSVLELNDLVK	206033961 413887198	479887793 726647782	0.754	0.22	0.704	0.20	4
L5	FNTEVTENLMK VLDNAVEELELITGQKPLVTK	146432962 506753382	285458875 978962971	0.716	0.17	0.670	0.15	4
L21	VGAPTVEGATVTATVNK MFAIIETGGK	814361978 415922960	1727599000 718684233	0.711	0.19	0.663	0.17	4
L19	KISSGVGVER IQVFEGVVIK	74180725 240347206	132478634 491698641	0.707	0.18	0.660	0.16	4
L23	MAVEEIFNVK VASVNIMNYKPK	256220634 114218375	454271325 234774687	0.706	0.18	0.670	0.16	4
L31	LDISSDSHPFYTGR TSSEMMEWEDGKEYPVIR	106385716 110680881	240923104 156818947	0.680	0.25	0.634	0.22	4
L15	EYAIVNLDQLNKFEDGTEVTPALLVESGVVK SGGGVRPGFEGGQLPLFR	129352920 43451981	254646048 74248737	0.679	0.17	0.634	0.15	4
L2	VGNALPLQNIPVGTVVHNIELKPGK ATIQVGNLQHELVNVGK	351239287 269212060	439771384 533457363	0.597	0.22	0.557	0.20	4
L14	TANIGDVIVCTVK IVSLAPEVL	185468056 12834771	350581861 10969082	0.507	0.26	0.473	0.24	4
L29	FQLATGQLEETAR DLTTSEIEEQIK	330713575 256226592	482457414 326708468	0.505	0.13	0.471	0.11	4
L18	GVTLAQASSK DSDIATTATK	50396014 92701187	45511112 81780390	0.330	0.08	0.308	0.07	4
L20	AFAQLV MLSEI	224658307 209751898	180951452 84673810	0.223	0.10	0.208	0.09	4
L16	ILFEVAGVSEEVAR +2 ILFEVAGVSEEVAR +1 GAVEGWIAVVKPGR	65049732 6470425 113833683	37124254 3119370 43675322	0.177	0.05	0.165	0.05	6
L27	GGDDTLFAK IYPGENVGR	38518456 87239977	20382700 16775068	0.133	0.08	0.124	0.07	4
L25	TVEVPVQLVGEAVGAK VPAVVYGYGTK	1185560675 189806833	629321059 29347836	0.126	0.09	0.118	0.08	4

L24	VVVEGVNIMK	117209777	215355913	0.678 +/- 0.20	n/a	0.633 +/- 0.18	n/a	2
L32	ISVPGMTECPNCGEYK	54564520	138423473	0.936 +/- 0.28	n/a	0.874 +/- 0.24	n/a	2
L33	VNVTLACTECGDR	40363666	35505371	0.325 +/- 0.10	n/a	0.303 +/- 0.08	n/a	2
L36	VMVICENPK	128621813	30990312	0.089 +/- 0.03	n/a	0.083 +/- 0.02	n/a	2

Protocols

Allelic Exchange:

1) Select *S. aureus* carrying pMAD derivative on erythromycin (5ug/mL) and Xgal (200 ug/mL). Colonies should be blue.

2) Grow overnight at 42-44C in erythromycin

3) Subculture in erythromycin at 42-44C overnight

4) Subculture in erythromycin at 42-44C overnight

5) Serially dilute culture and plate on erythromycin Xgal. Grow overnight at 42-44C. Colonies should be blue. These represent the cointegrate - the plasmid should be forced to integrate via homologous recombination under these conditions.

6) Pick blue colonies and innoculate into media without erythromycin at 30C. This allows the plasmid to excise via homologous recombination, hopefully leaving the desired mutation in the bacterial chromosome.

7) Subculture in media without erythromycin at 30C overnight

8) Subculture in media without erythromycin at 42-44C overnight. This will begin to select for bacteria that no longer contain the plasmid.

9) Repeat step 8

10) Serially dilute culture and plate on TSA Xgal. Grow at 30C for 24-48 hours. Blue product from Xgal cleavage takes a long time to develop in this system, it's best to wait so that blue will be easily discernable from white colonies.

11) Pick white colonies and patch them on: TSA erythromycin Xgal then TSA Xgal. Grow overnight at 30C. Candidates will be white and die on erythromycin. These must be checked by PCR and sequencing to ensure that the allelic exchange has occurred correctly.

Isolation of gDNA from *S. aureus* using DNAzol (Life Technologies Frederick, MD)

1) Centrifuge 250ul of overnight culture to harvest cells. Discard supernatant

2) Resuspend cells in 100ul TE and 2ul of 5mg/mL lysostaphin

3) Incubate 30 minutes at 37C in a water bath.

4) Add 500ul DNAzol, mix gently. Incubate 5 minutes at 65C.

5) Transfer mixture to a Qiagen (Venlo, Limburg) miniprep spin column and spin 10k for 1 minute to bind DNA

6) Wash with 750ul PE (ethanol wash buffer from Qiagen miniprep kit)

7) Spin the column dry to discard excess ethanol

8) Repeat steps 6 and 7

9) Place column in a clean microfuge tube for gDNA collection. Add 50ul prewarmed (65C) TE or Elution buffer to column and let stand for 5 minutes.

10) Elute gDNA by centrifugation

11) Store gDNA at -20C

(Ref: African Journal of Biotechnology 2003 2(8):251-253)

Vita

Erin Ashley Wall was born on April 29th, 1987 in Texas City, Texas and spent most of her life before graduate school in Houston, Texas. She graduated from Mirabeau B. Lamar High School in 2005 and received a full scholarship to the University of Houston Honors College, which she attended for two years before transferring to Texas Agricultural and Mechanical University. She received her Bachelor of Science degree in Animal Science from Texas A&M in 2010.

During her time at Texas A&M, Erin pursued her abiding interests in environmental microbiology and virology. She worked for the USDA at the Agricultural Research Service South Plains Area Research Center under Dr. Robin Anderson, gaining lab experience and an appreciation for the role that microbiota play in the lives of ruminants. Here she first witnessed a wastewater isolate of Vancomycin-Resistant *Enterococcus*. This work and a stint in the laboratory of Dr. Susan Payne, her virology professor at the TAMU College of Veterinary Medicine, caused her to become interested in mobile genetic elements and antimicrobial resistance. Naturally, this incited her to join the laboratory of Dr. Gail E. Christie at Virginia Commonwealth University in Richmond, VA for her PhD. She hopes to work on developing solutions to antimicrobial resistance using structure-based drug design and genetics.

Publications:

Wall EA, Caufield JH, Lyons CE, Manning KA, Dokland T, Christie GE. Specific N-terminal cleavage of ribosomal protein L27 in staphylococcus aureus and related bacteria. Molecular Microbiology. 2014 Nov 11.

Damle PK, **Wall EA**, Spilman MS, Dearborn AD, Ram G, Novick RP, Dokland T, Christie GE. The roles of SaPI1 proteins gp7 (CpmA) and gp6 (CpmB) in capsid size determination and helper phage interference. Virology. 2012 Oct 25;432(2):277-82. doi: 10.1016/j.virol.2012.05.026

Spilman MS, Damle PK, Dearborn AD, Rodenburg CM, Chang JR, **Wall EA**, Christie GE, Dokland T. Assembly of bacteriophage 80α capsids in a *Staphylococcus aureus* expression system. Virology. 2012 Dec 20;434(2):242-50. doi: 10.1016/j.virol.2012.08.031

Awards:

American Heart Association Predoctoral Grant 13PRE16800002 to Erin A. Wall
Evaluation of Two Essential Genes in Staphylococcus aureus: A Potential Avenue for Drug Design   July 2013-2015

Alpha Epsilon Lambda Nomination Award   February 2015

Phi Kappa Phi School of Medicine Nomination Award   March 2014

Cold Spring Harbor Tuition Stipend Protein Purification and Characterization Course
                                                        April 2013

Virginia Branch ASM Travel Grant   November 2014

Best Poster Presentation *Characterization of Mutations Affecting SaPI1 Capsid Size Determination.* Phage Virus Assembly meeting   October 2011

Phi Kappa Phi Honor Society   March 2011