INSIGHTS INTO THE STRUCTURE AND FUNCTION OF PRGW AND ITS CONSERVED CYSTEINES

A Dissertation
Submitted To
The Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

By
Jason Lewis Cutrera
December 2014

Examining Committee Members:

Bettina Buttaro, Ph.D., Advisory Chair, Microbiology and Immunology
Alexander Tsygankov, Ph.D., Committee Chair, Microbiology and Immunology
Cagla Tukel, Ph.D., Microbiology and Immunology
Marc Monestier, Ph.D., Microbiology and Immunology
Charles Grubmeyer, Ph.D., External Reader, Biochemistry

ABSTRACT

Insights into the structure and function of PrgW and its conserved cysteines

Jason Lewis Cutrera
Doctor of Philosophy in Microbiology and Immunology
Temple University, 2014
Doctoral Advisor: Bettina A. Buttaro

Enterococcus faecalis is a Gram-positive bacterial species that is typically a member of the human gastrointestinal tract microbiota. However, *E. faecalis* is also a nosocomial pathogen, which is involved in urinary tract infections, soft tissue infections and endocarditis. In recent times, the occurrence of antibiotic resistance has complicated the treatment of these infections. One of the major differences between commensal and pathogenic strains of *E. faecalis* is that pathogens contain multiple mobile elements such as plasmids, transposons and integrative conjugative elements (ICE). These elements allow for the acquisition and transfer of virulence factors and resistance genes. Conjugative plasmids are a class of plasmids present in *E. faecalis* whose transfer to host cells is induced by a small pheromone peptide, cCF10 (LVTLVFV). This peptide is initially encoded as a 22-amino acid precursor (pre-cCF10) from the signal sequence of the chromosomal *ccfA* gene and is then proteolytically cleaved by signal peptidase II and Eep. Once pCF10 has been transferred a host *E. faecalis* cell, it is exceptionally stable. A replicon clone is maintained in greater than 85% of host cells over 100 generations in the absence of selection, suggesting the stability of pCF10 is intrinsic to the replicon. Three unique features of the replication initiation protein PrgW may contribute to this

stability: (a) the interaction of PrgW with pre-cCF10, (b) disulfide bond formation at three conserved cysteines (C78, C275, and C307) in PrgW, and (c) processing of the nascent PrgW protein.

Replication initiation proteins associated with theta replicons, such as pCF10, are often self-contained units. To initiate plasmid replication, the replication initiation protein (PrgW) binds to direct repeats (*oriV*) in its own coding sequence (*prgW*). *In silico* analysis of PrgW suggests the existence of three distinct domains within the protein. The first 122 amino acids are homologous to a conserved domain present in related replication initiation proteins, which includes a Helix-Turn-Helix (HTH) DNA binding domain. This suggests that this domain of PrgW has a DNA-binding function and binds to the *oriV* site in *prgW*. The following 61 amino acids are not similar to any known sequence, and are encoded by the DNA sequence containing the direct repeats in the *oriV* site. This domain may or may not have a distinct function. The remaining sequence forms a domain that contains cysteines C275 and C307, and is also similar to no known structure. It is hypothesized that this domain is related to the stability of pCF10. C307 appears to be critical, as previous experiments indicate that its mutation alone affects plasmid stability.

Secondary structure analysis of this domain revealed the presence of multiple alpha-helices that contain distinct hydrophobic domains that possibly contribute to pre-cCF10 binding and PrgW tertiary structure. The positions of the conserved cysteines within these alpha-helices may stabilize a hydrophobic binding pocket that could potentially facilitate interaction with pre-cCF10.

PrgW has a predicted molecular weight of 38.6 KDa and can be detected in Western blots as a band with an apparent approximate molecular weight (mw) of 36,000. Previous data from our lab indicates that, when overexpressed in *E. faecalis*, four bands of PrgW are present with observed molecular weights of 40,000, 36,000, 24,000 and 18,000. Time course experiments revealed that the 40,000 mw form is converted to a 36,000 mw form independent. The 40,000 mw form is unstable (with a complete turnover in 30 minutes) while the 36,000 mw form has a half-life of greater than 4 hours. The 24,000 mw band does not have a DNA binding motif and is likely a turnover product. When the three conserved cysteines (and only cysteines) in PrgW are replaced with alanine, the 40,000 mw form is still processed to the 36,000 mw form. However, the cysteine to alanine mutants accumulate the 36,000 mw form.

ACKNOWLEDGEMENTS

Dedicated to my mother and father, and to my brothers for supporting me through this entire process. Also to all my friends who helped convince me to stick it through. Finally, to Olivia, Harry, Holly, John, Bertie, and Pascal, for keeping things from getting too boring.

TABLE OF CONTENTS

ABSTRACT ..... **iii**

ACKNOWLEDGEMENTS ..... **vi**

LIST OF FIGURES ..... **ix**

LIST OF TABLES ..... **x**

CHAPTER 1: INTRODUCTION **1**

- Enterococcus faecalis 1
- Antibiotic Resistance 2
  - β-lactam resistance 3
  - Aminoglycoside resistance 3
  - Glycopeptide resistance 4
  - Macrolide, lincosamide, and streptogramin A/B resistance 5
  - Tetracycline resistance 6
  - Antibiotic resistance transfer mechanisms 6
- Plasmids 7
- Plasmid Replication 7
  - Rolling-circle replication 7
  - Theta replication 8
- Enterococcal Plasmids 9
  - RCR plasmids 9
  - Inc18 plasmids 10
  - Conjugative plasmids 11
  - Conjugative transposons 12
  - Pheromone-inducible conjugative plasmids 13
- Virulence Factors 14
  - Cytolysin 14
  - Other secreted factors 15
  - Aggregation substance 16
  - Type-IV secretion systems 17
  - Biofilm formation 18
- pCF10 18
- PrgW 20
- Rep A_N proteins 23
- Processing of PrgW 24
- Goals 26

CHAPTER 2: MATERIALS AND METHODS **27**

- Bacterial strains and growth conditions 27
- DNA Manipulation 27
  - Plasmid extraction from E. coli 27
  - Agarose gel electrophoresis 30

DNA extraction from agarose gel 32
DNA primer design 32
Polymerase chain reaction 32
Restriction enzyme digestion 33
DNA ligation 33
Transformation of E. coli BL21(DE3) 35
Protein Manipulation 38
Preparation of whole-cell lysate 38
SDS-PAGE separation 39
Silver-staining procedure 40
Semi-dry electroblotting procedure (Western blotting) 42
Blocking and antibody incubation 43
Detection 44
Time-Course Assay 45
Conserved Domain Prediction 46
Secondary Structure Prediction 46
Protein Threading 46

CHAPTER 3: RESULTS 48
PrgW Structural Analysis 49
Prediction of the molecular weight and pI of PrgW 49
Comparison to PrgX X-ray crystal structure 49
Conserved domains 50
Secondary structure predictions 54
Helical-wheel projections show distinct hydrophobic regions on α-helices in the C-terminal domain of PrgW 55
Processing of PrgW 60
Previous studies 60
p40*, p36*, and p24* all appear at 37°C and are stable 60
p40* appears before p36* 64
p40* disappears after blocking of translation and p36* is a stable form of PrgW 66
PrgW cysteine mutants accumulate p36* 68

CHAPTER 4: DISCUSSION 71
Structure of PrgW 72
Conserved domain analysis of PrgW 72
PrgW secondary structure and binding pocket analysis 74
PrgW Processing 76
p40* and p36* 77
Role of the conserved cysteines 79
p24* 79
Conversion of p40* to p36* 80
Crystal Structure Analysis of Staphylococcal RepA 82
Comparison of PrgW and DnaD 84
Proposed PrgW Function in Replication 85
Conclusion 87

REFERENCES

APPENDIX

LIST OF FIGURES

Figure 1. Map of pheromone-inducible plasmid pCF10 21

Figure 2. Phyre2 threading data comparing PrgW with PrgX…………………………...51

Figure 3. Conserved domain database results for PrgW and RepA of pSK41 53

Figure 4. The positions of α-helices in the primary structure of PrgW 56

Figure 5. The position of major α-helices in the C-terminal region of PrgW................... 58

Figure 6. Possible arrangement of α-helices in the C-terminal domain of PrgW………..59

Figure 7. Time-course data of *E. faecalis* OG1RF(pPRGW) incubated at 37°C 61

Figure 8. Time-course data of *E. faecalis* OG1RF(pPRGW) incubated at 37°C and extended to 3.5 hours 63

Figure 9. Time-course data of *E. faecalis* OG1RF(pPRGW) incubated at 30°C 65

Figure 10. Time-course data of *E. faecalis* OG1RF(pPRGW) induced with low levels of nisin incubated at 37°C. 67

Figure 11. Time-course data of *E. faecalis* OG1RF(pPRGW) incubated with chloramphenicol at 37°C 69

Figure 12. Accumulation of PrgW p36* in *E. faecalis* OG1RF(pORI10) cysteine mutants in late exponential growth... 70

Figure 13. The conserved C-terminal sequence in PrgW……………………………......86

Figure 14. Hypotheses for PrgW function in replication 88

LIST OF TABLES

Table 1. Bacterial strains used in this study……………………………………...…...28

Table 2. Plasmids used in this study……………………………...…………………..31

Table 3. Primers used in this study…………………...…………………………………34

Table 4. PCR thermocycler program..36

Table 5. Restriction enzymes used in this study 37

Table 6. Acrylamide gel formulations…………………………………………………..41

CHAPTER 1

INTRODUCTION

Enterococcus faecalis

Enterococcus faecalis is a gram-positive bacterial species of the phylum Firmicutes. Enterococcus species are typically ovoid in shape, and found in chains when isolated from broth cultures. They are facultative anaerobes, capable of growing in both oxygen-rich and oxygen-deprived environments, and have an optimal growth temperature of 35°C, with an observed growth range of 10 to 45°C. Thus, enterococcus species are capable of colonizing a wide range of environments (Facklam, *et al.*, 2002, Hancock, 2006). In addition, enterococcus species have been shown to persist in nutrient-poor environments, resist desiccation and high salt concentrations, and adapt to detergent stress (Hancock, 2006).

Enterococcus species are typically present in the intestinal tracts of birds and mammals and is known to be a member of the normal flora in the human gastrointestinal tract and oral cavity (Gilmore, 2013). While not traditionally considered virulent, enterococcus species have emerged as one of the leading causes of nosocomial infections in recent times. *E. faecalis* has been implicated as a causative agent of urinary tract infections, soft tissue infections, and endocarditis, with many of these infections displaying a significant degree of antibiotic resistance (Jett, 1994, Murray, 1999). This includes vancomycin-resistant *E. faecalis*, a subset of vancomycin-resistant enterococci

(VRE). The conversion of commensal *E. faecalis* strains to antibiotic-resistant pathogenic strains is due to the presence of a wide range of mobile genetic elements.

In addition to causing resistant infections, these strains serve as important reservoirs for resistance genes, and are capable of spreading these resistance determinants to other organisms, such as *Staphylococcus aureus* (Lebreton, 2013). The prevalence of enterococcus strains in the environment and their ability to acquire and disseminate antibiotic resistance genes plays a significant role in public health (Gilmore, 2013). There is evidence to support the hypothesis that antibiotic resistance in enterococcus is selected for by agricultural practices. As the enterococcal strains colonizing animals become resistant, they spread this resistance to human enterococcal commensals. These commensals, in turn, spread resistance to other pathogenic organisms (Lebreton, 2013).

Antibiotic Resistance

Antibiotic resistance is a major concern in the public health community. Antibiotic-resistant infections continue to increase in number each year, and the incidence of vancomycin-resistant Gram-positive bacteria, including VRE, is becoming more common. While infections caused by resistant organisms pose an inherent threat, they represent a more indirect threat as well. Since many antibiotic-resistance determinants are located on mobile genetic elements, and are potentially capable of spreading from a less virulent organism, such as *Enterococcus faecalis*, to a much more virulent organism, such as *S. aureus*.

Enterococcal species, including *E. faecalis*, display antibiotic resistance due to both innate resistance mechanisms and mechanisms acquired from mobile elements, such as plasmids. In many cases, innate antibiotic resistance is a low-level phenomenon, while higher-level resistances are most often acquired by horizontal gene transfer.

β-lactam Resistance

Enterococcus species possess an innate low-level resistance to the β-lactam class of antibiotics (e.g. penicillin, ampicillin, cephalosporin, etc.) (Kak, 2002). While this resistance is weak in *E. faecalis*, roughly 90% of modern *E. faecium* infections are resistant to β-lactam antibiotics (Galloway-Peña, 2009). β-lactams interfere with the process of cell wall synthesis by binding to and inhibiting the function of penicillin-binding proteins (PBPs). The innate resistance of *E. faecium* arises due to PBP5, a class of PBPs that shows low-affinity binding towards β-lactam agents. While chromosomal in nature, the gene encoding PBP5 has been observed to transfer between cells in the presence of the independent vancomycin-resistance conjugative transposon *Tn5382* (Rice, 2005).

Aminoglycoside Resistance

Aminoglycosides function by inhibiting protein synthesis by binding to the 16S subunit of rRNA. Enterococcal species also possess an intrinsic low-level resistance to aminoglycosides (kanamycin, gentamycin, etc.), because they encode systems that limit the transport of these drugs across their cell membranes. In the past, this resistance was overcome by the use of β-lactams, vancomycin, and other cell-wall inhibitors, to increase penetration of the drug and increasing its concentration at the cell membrane. However, the recent increase of resistance towards cell-wall inhibiting agents has reduced the efficacy of this combination (Moellering, 1979). High-level resistances are acquired from extrachromosomal elements, such as transposons, and often consist of the production of an altered ribosomal S12 protein or an aminoglycoside nucleotidyltransferase (Arias, 2012).

Glycopeptide Resistance

The most important glycopeptide in the clinical setting is vancomycin, which is used as a "last resort" drug to treat gram-positive bacterial infections that have become resistant to other antibiotics., which is often the case for *E. faecalis*. Gram-negative organisms are innately immune to glycopeptides, as their outer membrane prevents the molecule from reaching its target on the cell wall. However, the emergence of vancomycin-resistant gram-positive organisms was first observed in 1986 (Uttley, 1988).

Like the β-lactam class of antibiotics, vancomycin functions by inhibiting cell wall synthesis. Unlike β-lactams, vancomycin does not bind enzymes involved in cell wall construction. Instead, it forms complexes with the D-alanine-D-alanine (D-Ala-D-Ala) terminus of the pentapeptide precursors, preventing further cell wall construction (Reynolds, 1989). There have been six vancomycin-resistance determinants identified in vancomycin-resistant enterococci (designated *vanA-E*, and *van G*) (Gholizadeh, 2000). Each of these systems possesses three components (Weaver, 2006). First, a two-component signal transduction system, VanRS, that detects the presence of vancomycin and induces the resistance operon. Second, an essential three-gene core system that encodes enzymes responsible for producing D-Lac or D-Ser precursors and eliminating D-Ala-D-Ala dipeptides (Weaver, 2006). The third component of these resistance operons are accessory genes (Weaver, 2006). The VanC, VanE, and VanG systems replace the D-Ala-D-Ala dipepetide precursors with D-Ala-D-Ser dipeptides, providing low-level vancomycin resistance. Each of these systems are chromosomally located and do not appear to spread between cells on mobile elements (Clark, 1998, Fines, 1999, McKessar, 2000, Kak, 2002). The VanA, VanB, and VanD systems replace their cell wall D-Ala-D-Ala dipeptides with D-Ala-D-Lac dimers. VanA is characterized by producing high-level resistance to vancomycin and other glycopeptide drugs, while VanB provides variable resistance to vancomycin only (Quintiliani, 1993). The decreased protection of the latter system is thought to be the result of defective sensing of the glycopeptide drug (Quintiliani, 1993). VanD produces a defective core enzyme, but is constitutively active, which produces a net low-level vancomycin resistance (Perichon, 1997). In addition, both VanA and VanB are frequently carried on mobile genetic elements. VanA is located on the Tn3-like transposon *Tn1546*, and is often found on conjugative plasmids (Leclercq, 1988). VanB has also been isolated from a number of mobile elements, particularly the transposon *Tn1547* (Quintiliani, 1993). VanD does not appear to be mobile (Perichon, 1997).

Macrolide, Lincosamide, and Streptogramin A/B Resistance

These classes of antibiotics (which include erythromycin and lincomycin) interfere with protein synthesis by preventing the binding of, or causing the early release of, the peptidyl-tRNA from the ribosome. Enterococci have acquired many methods of resistance towards these agents (the resistance is termed MLS resistence), including drug modification, target modification and efflux pumps (Arias 2012). The most common determinant providing resistance to these antibiotics is the *ermB* gene, which is found on the *Tn3*-like transposon *Tn917* (Portillo, 2000, Shaw, 1985). The *ermB* gene encodes a protein that methylates the 23S rRNA subunit, preventing the binding of these agents.

Tetracycline Resistance

Tetracycline inhibits protein synthesis by preventing the binding of aminoacyl-tRNA to the ribosome (Suarez, 1965). The most common tetracycline resistance determinant in *E. faecalis* is the *tetM* gene, which is found on the conjugative transposon *Tn916*, and *Tn916*-like genetic elements. The *tetM* gene confers tetracycline resistance to its host by encoding a protein that binds the ribosome and alters its conformation such that tetracycline cannot bind (Burdett, 1986). *tetM* is acquired by a host when the host accepts a *Tn916*-like element and the element inserts itself into the host chromosome (Su, 1992).

Antibiotic Resistance Transfer Mechanisms

As mentioned above, a significant number of antibiotic resistance genes that occur in *E. faecalis* are carried on mobile genetic elements, such as plasmids. This allows for the efficient transfer of these resistance genes between populations of that species (Dunny, 1995). The primary method of horizontal gene transfer in *E. faecalis* is through the transfer of plasmids and transposons via conjugation.

Plasmids

Plasmids are molecules of DNA that exist separately from the bacterial chromosome, and replicate and transfer independently within a suitable host species.

Plasmid Replication

Rolling-Circle Replication

The rolling-circle replication mechanism was first observed in bacteriophages, and the RCR plasmids use this mechanism to disseminate to many different bacterial species (Weaver, 2002). RCR plasmids are characterized by their small size, high copy number, and wide host range. The rolling-circle replication process proceeds as follows. First, a single-stranded nick is made in the leading strand at the double-stranded origin of replication (*dso*) by the plasmid-encoded replication initiation protein. The initiation protein also melts the DNA at the *dso* and presents the nick site as ssDNA. Next, host DNA helicase and DNA polymerase III (DNA Pol III) are recruited to the nick site, and replication is primed by the free 3' hydroxyl at the nick site. Replication proceeds, and once the first round is complete, the duplicated leading strand is released by the replication initiation protein. After release, host RNA polymerase recognizes a distinct origin (the single-stranded origin of replication, or *sso*) on the nascent strand and synthesizes an RNA primer at that site. The RNA primer is recognized by host DNA polymerase I (DNA Pol I) which begins synthesis, but DNA Pol III soon displaces DNA Pol I and completes the synthesis of the lagging strand. This process allows for the quick and efficient replication of the plasmid. In addition, the replication process requires few host proteins apart from DNA helicase, DNA Pol III, and DNA Pol I, so it is viable in a broad range of hosts. However, RCR plasmids must limit their replication, so they do not overwhelm their host. Replication control is most often achieved by limiting the synthesis of the replication initiation protein (Weaver, 2002).

Theta Replication

The pheromone-inducible conjugative plasmids observed in *E. faecalis* are an example of plasmids that undergo theta-replication. This method of replication is named for the distinct structures (shaped like the Greek letter theta, θ) seen in electron micrographs of the process. The origin of replication in theta-replicating plasmids (*oriV*) is defined as the minimal *cis* element required for replication initiation, and often contains regions of direct or inverted repeat sequences spaced at approximately 11 bp apart. Theta-replication in plasmids is characterized by the need for a plasmid-encoded replication initiation protein (del Solar, 1998). This replication initiation protein contains a helix-turn-helix (HTH) domain that recognizes these repeats and binds to the DNA at *oriV*. After binding the *oriV*, the replication initiation protein melts a small section of DNA, and host DnaA binds to the site. DnaA causes additional melting at the origin and stabilizes ssDNA. The host DNA replication machinery (DNA helicase, DNA Pol III, RNA Primase, etc.) is then recruited to the site, and the plasmid is replicated in a fashion similar to the bacterial chromosome (del Solar, 1998). Unlike RCR plasmids, theta-replicating plasmids generally have a low copy number (del Solar, 1998).

Enterococcal Plasmids

There are three main classes of plasmids that exist in enterococcus species: rolling-circle replicating plasmids (RCR plasmids), Inc18 plasmids, and pheromone-inducible conjugative plasmids (also known as pheromone-responsive plasmids) (Weaver, 2002). Each of these classes is distinguished by their replication mechanism. While the RCR plasmids and Inc18 plasmids have been observed to replicate in other bacterial species, the pheromone-responsive plasmids appear to only replicate in enterococcus species. However, plasmids that contain replicons related to those found in pheromone-responsive plasmids (RepA_N replicons) have been observed in a wide range of Gram-positive species (Weaver, 2009).

RCR Plasmids

Rolling-circle-replicating plasmids are common in enterococci. However, due to their small size, these plasmids do not carry the machinery necessary to mobilize them. Instead, they most likely take advantage of the presence of the conjugation machinery introduced on a separate plasmid and are co-transferred during transfer of the conjugative plasmids (Weaver, 2002). In addition, they do not commonly carry antibiotic resistance determinants. However, the *E. faecalis* plasmid pAMα1 possesses a tetracycline resistance gene that is amplified by recombination events (Clewell, 1975).

Inc18 Plasmids

The incompatibility group 18 (Inc18) plasmids are a class of approximately 15 plasmids that can be found in Firmicutes, including enterococci. These plasmids are approximately 25 to 30 kb in size, maintain a low copy number in their host, and are classified by their type of post-segregational killing systems (Weaver, 2002, Palmer, 2010). These plasmids are characterized by a unique replication initiation mechanism (Le Chatelier, 2001). The most studied example of an Inc18 plasmid is pAMβ1 in *E. faecalis*. In this plasmid, a plasmid-encoded protein, RepE binds the origin of replication (*ori*) and melts a section approximately 20 nucleotides long. Host RNA polymerase is then recruited to the replication site and displaces RepE. The RNA polymerase then synthesizes an RNA transcript. The transcript is then cleaved into a short RNA primer. DNA Pol I is recruited and extends replication at the primer. Replication then proceeds unidirectionally. (Le Chatelier, 2001).

Inc18 plasmids are capable of disseminating to a wide range of gram-positive hosts, and are carriers of antibiotic resistance genes. The plasmid pAMβ1 encodes MLS resistance, while another Inc18 plasmid, pIP501, encodes MLS and chloramphenicol resistance (Thompson, 2003). These plasmids have been shown to transfer resistance to other bacteria, including gram-positive species and also *E. coli* (Kurenbach, 2003). Many of these elements lack a type-IV secretion system, and it is thought that they instead take over the transfer machinery encoded on another mobilizable element and are co-transferred with conjugative plasmids (Palmer, 2010).

Conjugative Plasmids

Conjugative plasmids are a subclass of theta-replicating plasmids that have the ability to initiate and conduct their own transfer from cell to cell via conjugation (i.e. they are self-conjugative). Jensen, *et al.* proposed a classification scheme for conjugative plasmids in gram-positive bacteria by dividing the plasmids into families based on homologous features of their respective replicons (Jensen, *et al.*, 2010). As a result of this scheme, conjugative plasmids are sorted into 19 Rep families, or into a unique category if the replicon does not resemble that of an established family.

A second classification scheme was proposed by Palmer, *et al.* In this scheme, conjugative plasmids that exist in enterococcal species were grouped by their host range, and divided into three categories. The member of each category were correlated with its corresponding Rep family (Palmer, *et al.*, 2010). The first category is the pheromone responsive plasmids, a family of low copy number plasmids that replicate and transfer in response to a pheromone molecule produced by a host cell. The Rep families most commonly represented in this category are Rep8 and Rep9 among enterococcal plasmids (Palmer, *et al.*, 2010) and Rep15 and Rep19 among staphylococcal plasmids (Jensen, *et al.*, 2010). These plasmids have a very narrow host range and each such plasmid is almost always stable only within its host species. The pheromone responsive plasmids are also the most well-studied family of conjugative plasmids in *E. faecalis*. The second category contains any conjugative plasmid known to have a broad host range, and the most commonly represented Rep family is Rep1, which includes members of the Inc18 family of plasmids described above (Jensen, *et al.*, 2010, Palmer, *et al.*, 2010). The third category of conjugative plasmids includes any plasmid in which the host range has not been determined. The most prominent members of this category include the gram-positive plasmids pHTβ and pMG1, both originally recovered from *E. faecium* (Palmer, *et al.* 2010). The Rep family classification of these plasmids is not yet determined.

Conjugative Transposons

Transposons are small mobile genetic elements that are capable of excising themselves from their host element and inserting into a foreign element. Therefore, copies of a single transposon can be found within many different plasmids and the chromosomes of a wide range of organisms.

Transposons that occur in *E. faecalis* can be grouped into three general categories: Tn3-like transposons, composite transposons, and conjugative transposons (Weaver, 2002). The *Tn3*-like transposons are a large group of transposons first described in gram-negative bacteria. The most well-studied member of this family in Enterococcus species is *Tn917*, which encodes the *ermB* macrolide/lincosamide/streptogramin resistance gene (Perkins, 1984).

Composite transposons are characterized by terminal insertion sequences (IS elements) that flank their genes. These transposons are implicated in the spread of antibiotic resistance determinants, including the vancomycin-resistance genes *vanA* and *vanB* (Heaton, 1995, Quintiliani, 1996).

Conjugative transposons are one class of transposons that were first characterized in *E. faecalis* but have since been observed in a wide range of Gram-positive organisms (Buu-Hoï, 1980, Bertram 1991). These transposons are characterized by their ability to form a circular intermediate after excision, and to mediate their own transfer between cells through conjugation. The best-studied conjugate transposon is *Tn916* isolated from *E. faecalis*. *Tn916* is 18 kb in size and has a modular structure (Flannagan, 1994). Its excision from and reintegration into host genomes is mediated by a λ-phage-like Xis and Integrase, respectively, both of which are located on the left end of the transposon. The right-hand section of the transposon contains all the genes required for conjugation, as well as the origin of transfer (*oriT*) (Flannagan, 1994). Within this section of the transposon, *Tn916* possesses several genes that appear similar to a type-IV secretion system (Alvarez-Martinez, et al., 2009). Situated between these two sections is a *tetM* tetracycline resistance determinant (Flannagan, 1994).

Pheromone-Inducible Conjugative Plasmids

Pheromone-inducible conjugative plasmids are a class of plasmids commonly found in *E. faecalis*, but also occur in other *Enterococcus* species, though less commonly (Palmer, *et al.*, 2010). These plasmids facilitate their own transfer between individual cells in response to a small peptide that is produced by plasmid-free recipient *E. faecalis* cells. Plasmids of this class are highly stable within a population of cells, and are capable of disseminating a wide range of genetic elements throughout a population (Weaver, 2009, Palmer, 2010). Factors that have been isolated from these plasmids include ultraviolet resistance and a wide range of antibiotic resistances, including resistance operons for tetracycline, chloramphenicol, gentamycin, macrolides/lincosamides/streptogramin antibiotics, and vancomycin. Virulence factors, such as cytolysin and aggregation substance have also been isolated from these plasmids. In addition, they can also carry transposons and conjugative transposons. Furthermore, recent evidence indicates that these plasmids can mediate exchange of chromosomal material between strains of enterococci, perhaps by some type of high frequency recombination events, such as those described for the F-plasmids of *Escherichia coli* (Manson, 2010).

Pheromone-responsive plasmids are unique to *Enterococcus*, and do not replicate outside the genus (Clewell, 2002). However, it was discovered that these plasmids respond to pheromone signals produced by other gram-positive genera as well. The chromosomally-encoded *S. aureus* lipoprotein CamE is processed into a form that possesses pheromone activity towards the plasmid pAM373 (Flannigan, 2002). Additionally, the *S. aureus* plasmid pSK41 encodes the TraH lipoprotein that exhibits cAD1 activity upon processing (Firth, 1994).

Virulence Factors

The pheromone responsive plasmids are known to harbor a number of virulence factors. The most important factors that are carried on these plasmids include cytolysin, aggregation substance, type-IV secretion systems, and genes that promote biofilm formation.

Cytolysin

Cytolysin is a secreted pore-forming toxin that is distantly related to streptolysin S and bacteriocins and exhibits both hemolytic and bactericidal activity (Gilmore, 2002). The operon that encodes cytolysin can be found within the large pheromone-responsive plasmids or in the *E. faecalis* chromosome on a pathogenicity island (Shankar, 2002). Cytolysin production occurs in approximately one third of *E. faecalis* strains (Arias, 2012). The operon consists of six genes (*cylL*$_S$, *cylL*$_L$, *cylM*, *cylB*, *cylA*, and *cylI*), each of which is essential to the production of mature toxin. The genes *cylL*$_S$ and *cylL*$_L$ respectively encode the small and large peptides that comprise the inactive lytic component of the toxin. After translation, the lytic component is modified by *cylM*. These modifications are consistent with those seen in lantibiotic bacteriocins and are characterized by the dehydration of serine and threonine residues into lanthionine and β-methyl lanthionine, and the formation of a thio-ester bridge between the resulting double-bond and the sulfhydryl group as a cysteine (Coburn and Gilmore, 2003). The lytic components are then transported across the membrane by CylB, and then activated by the serine protease CylA. The gene product of *cylI* confers immunity to cytolysin, so that cells that produce the toxin are not lysed themselves.

The role of cytolysin in pathogenesis is well-studied. Intraperitoneal lethal challenge studies in mice indicated that *E. faecalis* strains expressing functional cytolysin were more than 10 times more toxic than those strains that did not (Ike, *et al.* 1984). In addition, the presence of cytolysin was shown to promote the presence of *E. faecalis* in the bloodstream (Huycke, *et al.* 1998).

Other Secreted Factors

In addition to cytolysin, *E. faecalis* secretes two other factors that contribute to pathogenicity. One is a zinc metalloproteinase known as gelatinase (GelE). The other factor is a serine proteinase, SprE. These proteins promote the pathogenicity of *E. faecalis* by causing degradation of host tissue, modulation of the host immune response, and the activation of autolysin, an enzyme that contributes to biofilm formation (Thomas, 2009). Both the *gelE* and *sprE* genes are encoded on the same transcript and are regulated by the Fsr quorum-sensing system (Pinkston, 2011). Several studies indicate that *E. faecalis* strains that do not secrete functioning GelE and SprE are significantly less virulent (Arias, 2012).

Aggregation Substance

Aggregation substance is a protein that is localized to the cell surface of *E. faecalis* that mediates the adhesion of the bacterial cell to other bacterial and eukaryotic cells. Operons that encode aggregation substance are often found on pheromone-responsive plasmids that also possess antibiotic resistance, such as pAD1, pPD1, and pCF10 (Ehrenfeld, 1986, Arias, 2012, Clewell, 2002). A very different aggregation substance was found on pAM373 (Clewell, 2002). Aggregation substance is produced during conjugative transfer of pheromone-responsive plasmids after induction by pheromone produced by recipient cells and in the human body when the conjugation system is induced by human plasma. This allows for the clumping of *E. faecalis* during conjugation, which promotes the efficient conjugative transfer of genetic material in liquid culture. Aggregation substance contains two Arg-Gly-Asp (RGD) motifs, similar to those found in eukaryotic fibronectin, which promote binding to host integrins (Galli, 1989, Kao, 1991). In addition to facilitating the transfer of genetic material, aggregation substance has been observed to affect pathogenesis by increasing the binding of *E. faecalis* to eukaryotic cells (Olmstead, 1994). An example of this is the *E. faecalis* endocarditis model in which aggregation substance allows the formation of large bacterial aggregates on heart valves (Schlievert, 1998).

Type-IV Secretion Systems

Pheromone-inducible conjugative plasmids facilitate their transfer to neighboring cells through a plasmid-encoded type-IV secretion system (T4SS). These secretion systems are found in both Gram-positive and Gram-negative bacteria, often carried by plasmids and other mobile genetic elements. The T4SS systems are characterized by their unique ability to secrete both DNA and proteins from the host cell directly into a neighboring cell. The T4SS of the *Agrobacterium tumefaciens* plasmid pTi is the most well-studied, but all T4SS components appear to be evolutionarily related (Fronzes, 2009).

All T4SS systems follow a similar pattern of assembly. The first stage in this process is the assembly of the translocation pore. The substrates required to assemble the pore are recruited by a membrane-anchored cytoplasmic ATPase known as the Type-IV coupling protein, or T4CP (VirD4 in pTi, PcfC in the pheromone-inducible plasmid pCF10, etc.) (Fronzes, 2009, Alvarez-Martinez, 2009). The coupling protein interacts with a membrane-bound ATPase (VirB4 in pTi, PrgJ in pCF10) and with the structural pore proteins (Vir6-10 in pTi, putatively PrgD, F, H, I, K, L and PcfH in pCF10) to form a channel that penetrates both membranes (Fronzes, 2009, Alvarez-Martinez, 2009). Pilum-associated proteins (VirB2 and VirB5 in pTi, the pilum-associated proteins of pCF10 are not known) also assemble on the outer membrane at the pore location (Fronzes, 2009). Next, a third ATPase (the cytosolic VirB11 in pTi) interacts with the translocation pore, inducing pilum assembly or secretion (Fronzes, 2009, Alvarez-Martinez, 2009). There is no known counterpart to VirB11 in pCF10, and it is thought that PrgJ assumes this role (Li, *et al.*, 2012).

Biofilm Formation

Biofilms are masses of bacterial colonies that are supported by a matrix. These structures can be formed by many different species of bacteria, including *E. faecalis*. In enterococcal biofilms, the matrix is secreted by the host cells and consists of polysaccharides, protein, and extracellular DNA (eDNA). Biofilms may contain many localized heterogeneous sub-populations of cells, each one having different virulence and antibiotic resistance traits. Horizontal gene transfer is prevalent within the sub-populations, however, due to the close proximity of the cells to one another. This is a concern for the clinical health community, as biofilms promote the spread of virulence factors and antibiotic resistance. Biofilms are can also grow on both biotic and abiotic surfaces, including catheters, implanted devices, and hospital surfaces. Biofilms are also very resilient, and are difficult to remove once established.

The formation of biofilms by *E. faecalis* is heavily influenced by the surface protein Esp. The gene that encodes this protein is present on a pathogenicity island (Shankar, 2002) and the protein shares structural similarity to a *S. aureus*-encoded protein that is associated with biofilm formation (Toledo-Arana, 2001). While Esp appears to promote the formation of biofilms among *E. faecalis*, it does not appear to be necessary for the process, as strains that lack a functional Esp protein can still form biofilms (Toledo-Arana, 2001).

pCF10

The pheromone-responsive plasmid pCF10 is 67,673 bp in size and carries the transposon *Tn925*, which in turn carries the *tetM* tetracycline resistance gene. The transfer of pCF10 occurs only between cells of *E. faecalis*, and is induced by the processed 7-amino acid signal sequence cCF10 (LVTLVFV). This pheromone peptide is derived from a precursor, pre-cCF10, which is encoded as a 22-amino acid sequence. Pre-cCF10 itself is derived from the signal sequence of the chromosomal *ccfA* gene, a putative lipoprotein. First, the *ccfA* gene product is proteolytically cleaved into a pro-lipoprotein by signal peptidase II (SpaseII). The gene product is further processed into mature cCF10 by the Eep endoprotease and possibly another carboxyl peptidase (Antiporta, 2002, Buttaro, 2000).

Since *ccfA* and all the proteases required to process mature cCF10 are chromosomal, cells that do not contain the pCF10 plasmid produce the pheromone peptide. cCF10 interacts with the pCF10-encoded PrgZ receptor on a donor cell. PrgZ facilitates the import of cCF10 into the donor cell. The cCF10 then interacts with the PrgX transcriptional repressor, changing its conformation such that it no longer binds DNA and repression is lost. The release of PrgX allows for the transcription of many plasmid-encoded genes, including Aggregation Substance (PrgB) and the Type IV secretion system that facilitates plasmid transfer between cells (see above for the proteins involved). Cells that already contain pCF10 still produce mature cCF10, but do not induce self-conjugation because of the secreted inhibitor iCF10 and the insertion of PrgY into their membranes, which reduces the amount of cCF10 released from the membrane (Dunny, 2007).

One of the notable characteristics about pCF10 is its extremely high stability. The plasmid can persist in its host reservoir with 98% stability for 100 generations without antibiotic selection (Meloni, submitted). This plasmid also displays a very limited host range and will only replicate in *E. faecalis* (it can be transferred to other strains during conjugation, but is not maintained) (Meloni, submitted). Previous studies have indicated that the host range of pMSP6050, a plasmid containing the entire functional replicon of pCF10, can be extended to *Lactococcus lactis* when the latter is engineered to produce pre-cCF10 (Meloni, submitted).

The transposon *Tn925* is a conjugative transposon normally found on pCF10. It is approximately 18 kb in size and is highly similar to *Tn916* (Christie, 1987, Showsh, 1996). It carries a *Tn916*-like *tetM* tetracycline resistance determinant, and has been shown to mediate transfer of the *tetM* determinant between cells of *E. faecalis* (Torres, 1991). Previously performed mixed biofilm transfer experiments indicate that *tetM* was able to transfer from *E. faecalis* to *S. aureus* at a rate of 10^{-8} (Massie-Schuh, thesis). It is not currently known whether or not *Tn925* alone mediates the transfer of *tetM*, or whether pCF10 mediates the transfer. In any case, *Tn925* is thought to integrate into the *S. aureus* chromosome after transfer (Massie-Schuh, thesis).

PrgW

The pCF10 replication initiation protein, PrgW is a member of the RepA_N superfamily of replication proteins, which belong to plasmids that are found in low-GC Gram-positive organisms (Weaver, 2009). These proteins typically bind to direct repeated sequences on the plasmid and then initiate replication of the plasmid at the origin of replication (Weaver, 2009). Previous immunoprecipitation and electromobility shift experiments have shown that PrgW binds a set of direct repeats located within the nucleotide sequence for PrgW itself (Meloni and Mazeffa, submitted).

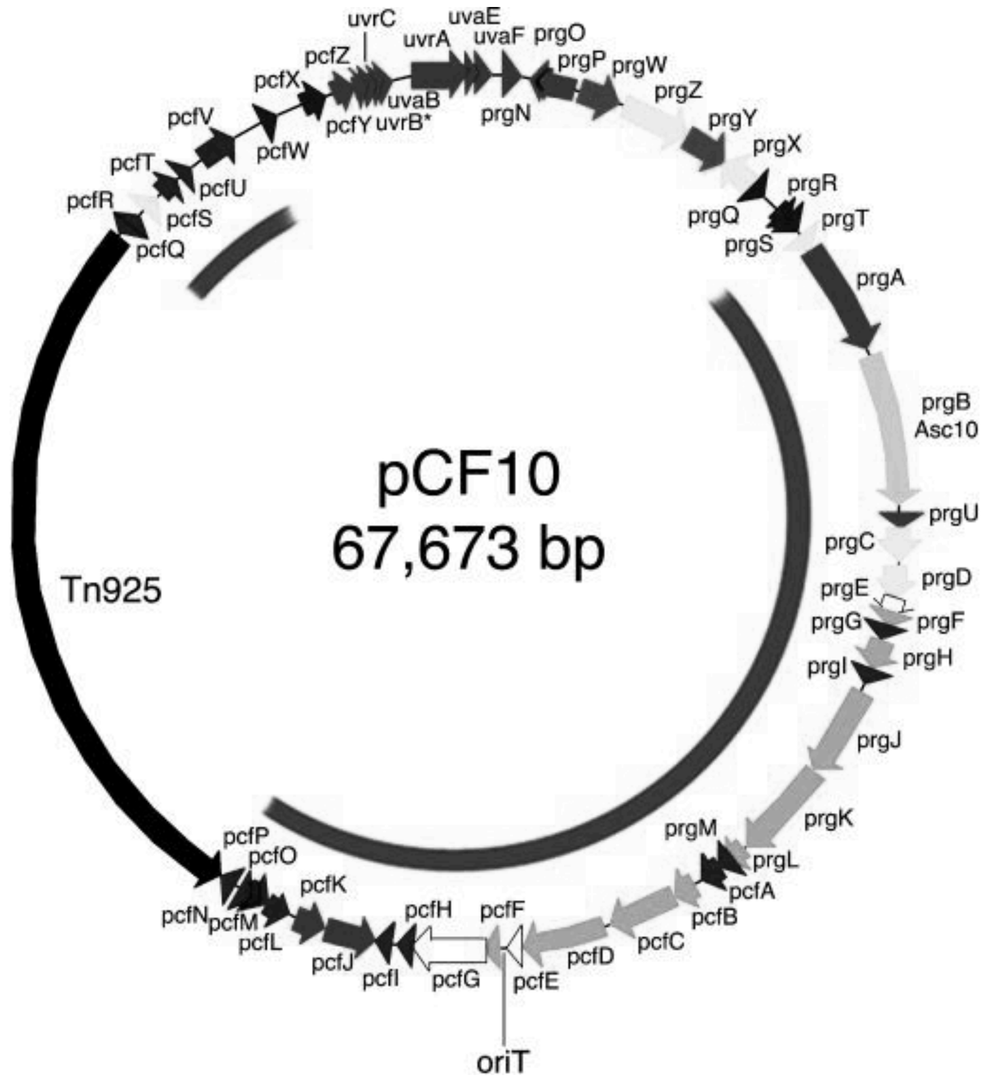

Figure 1. Map of *E. faecalis* pheromone-inducible plasmid pCF10 (Hirt, *et al.*, 2005).

Several experiments have also been performed detailing the host range of pCF10, which is related to the proper function of PrgW. Experiments using *Enterococcus hirae* have demonstrated that mature cCF10 is not necessary to determine the host range of PrgW. *E. hirae* is a species closely related to *E. faecalis* that does not produce mature cCF10, but instead produces a form of pre-cCF10 that contains mutations in the sequence corresponding to cCF10 (Meloni and Mazeffa, submitted). Despite this, when pCF10 is transferred into *E. hirae*, the plasmid is maintained (Meloni and Mazeffa, submitted). These experiments suggest that while cCF10 induces transfer of pCF10, it is pre-cCF10 that is important in establishing host range. Column chromatography experiments demonstrated that pre-cCF10 and mature cCF10 bind purified PrgW with different affinities (Buttaro, unpublished observations).

Previous experiments have shown that three cysteine residues in PrgW (C78, C275, and C307) are important for the stability and host range of pCF10. When the plasmid vector pMSP6050 was engineered with a *prgW* gene in which each of the cysteine residues has been replaced with alanine (the resulting plasmid is called pMSP6-78/275/307) (Utter, submitted), it was found that it still functioned as a replication protein, but displayed decreased stability in *E. faecalis*. It was also discovered that plasmid containing triple-mutant PrgW displayed a wider host range than the wild-type plasmid. When pMSP6-78/275/307 was electroporated into *L. lactis*, a low number of transformants were recovered. These transformants were confirmed to be *L. lactis* by 16S rRNA analysis and the presence of triple-mutant PrgW was confirmed by restriction analysis.

PrgW may have a secondary function as a transcriptional regulator of the *prgZY* operon. The gene *prgZ* encodes the cCF10 receptor gene (*prgZ*) and *prgY* encodes the protein that prevents self-induction of conjugation by donor-produced cCF10. Strains in which the 5' end of *prgW* was deleted but the *prgZ* promoter was left intact and in-frame failed to express the gene products of the *prgZY* operon (Ruhfel, *et al.*, 1993). In addition, the presence of PrgW induces transcription of *prgZY* in *E. faecalis* .Finally, direct and specific binding of PrgW to the *prgZY* promoter can be demonstrated (Meloni, unpublished observations).

Rep A_N proteins

PrgW is a member of the RepA_N family of proteins. Comparative analyses of the primary structures of several RepA_N proteins indicate that they fall into two general classes: those that possess three conserved cysteines and those that had either only one of these cysteines or none of them. The former group of proteins was limited to *E. faecalis* strains, while the latter group was found in other enterococcal and staphylococcal species (Utter, submitted). The RepA_N proteins belonging to the *Staphylococcus aureus* plasmid pSK41 and a putative plasmid found in *E. faecalis* strain TX0104 (hereafter called pTX0104) are 98% similar, the only difference being the deletion of the last repeat in the direct repeat region of the gene (Utter, submitted).

When the primary structure of PrgW from pCF10 is analyzed, three distinct regions can be recognized. Comparison of the primary structures of the replication initiation proteins of pSK41 and pTX0104 yields the following. The third of the protein towards the N-terminal is highly conserved between pCF10 and pTX0104/pSK41, sharing 61% identical amino acids and a 69% nucleotide identity. This is followed by the direct repeat region: the sequence of the direct repeats of pTX0104 and pSK41 share 100% identity to each other, except that pSK41 is lacking the 5th repeat. The pSK41/pTX0104 direct repeats are 55% similar to pCF10. Binding of the replication initiation proteins to the direct repeats have been documented by EMSA and DNase I footprinting analysis for pSK41 and EMSA for pCF10 (Kwong, 2004, Meloni, Submitted). The last portion of the gene following the direct repeats still shares 54% nucleotide identity, but only 16% amino acid identity and lacks C278 and C307. The disparity between the nucleotide and amino acid similarities of the two proteins allows us to hypothesize that the transfer of the pSK41 ancestor to *S. aureus* is an effect of the loss of C278 and C307 due to changes in the reading frame of the C-terminal portion of the protein. As the N-terminal regions of these proteins display strong homology, it can be further hypothesized that they emerged from a common ancestor protein, and the replication protein of pSK41 only recently spread to *S. aureus*.

Processing of PrgW

Previous experiments in our lab indicate that PrgW is processed, and that processing is required for the function of the protein (Massie-Schuh, thesis). Functional PrgW has a predicted molecular weight of 38,635. Analysis of multiple blots with multiple polyclonal antibodies suggest that PrgW is processed into several fragments: designated p40*, p36*, p24*, and p18*, based on the approximate apparent molecular weight of the fragment (40,000, 36,000, 24,000 and 18,000, respectively). The appearance of these bands in any given strain of *E. faecalis* is dependent on the expression level of PrgW in that clone (Massie-Schuh, thesis). In strains that possess the native plasmid pCF10, only the p36* and p24* bands are visible. Strains that contain the minimal replicon, pORI10, where expression of *prgW* is linked to an *erm* promoter, protein concentrations are increased, but p36* and p24* are still the only fragments visible on a Western blot. In strains that possess the nisin-inducible overexpression vector pPRGW, the p40* band is visible after nisin induction, along with the p36*,p24* and the p18* bands (Massie-Schuh, thesis). Full length PrgW, when expressed in *E. coli*, has an apparent approximate molecular weight of 40,000. This, suggests that PrgW is transcribed as the p40* form and then processed to the p36* form, which is then further processed into the p24* and p18* fragments.

Experimental evidence also indicates that the pCF10 replicon activity is related to the processing of PrgW. The sequence identity of each of the protein fragments is not known at this time. In addition, the nature of the processing is not known, and could include any or all of PrgW modification, cleavage, or folding into specific conformations. As these experiments demonstrate that both p40* and p36* are detected by antibodies to the N-terminal and also antibodies to the C-terminal of PrgW, the early processing step is likely a modification of PrgW folding or the result of a small fragment that is deleted from either end so that the antibody epitope is retained. Since p24* does not react with the anti-N-terminal antibody, it is likely for cleavage to occur during the step that forms that band.

Goals

Based on the possible processing of PrgW and its function as both a replication initiation protein and transcriptional regulator, the first goal these experiments are to elucidate information about the higher-order structure of the pCF10 replication initiation protein, PrgW. The primary structure of PrgW is analyzed using the NCBI Conserved Domain Database, as well as the Jpred Secondary Structure prediction server. The information yielded by these two online servers will then be projected onto the three-domain model of PrgW. In this way, the function of these domains may be predicted.

The second goal of these experiments is to extend the current knowledge of PrgW processing. Time course experiments are performed to obtain information about the order of PrgW fragment processing. The stability of the different fragments is analyzed by protein synthesis inhibition experiments. The dependence of the processing on the conserved cysteines is also determined.

CHAPTER 2

MATERIALS AND METHODS

Bacterial strains and growth conditions

Strains used in this study are *Enterococcus faecalis* OG1RF, *E. faecalis* OG1RF pORI10, *E. faecalis* OG1RF pORI10-78/275/307 (and the single- and double-mutants), *E. faecalis* OG1RF pPRGW, and *E. faecalis* OG1RF pPRGW-78/275/307. DNA was transformed into *E. coli* DH5α, and protein expressed in *E. coli* BL21 and BL21(DE3). All Enterococcus strains were grown in Todd-Hewitt (TH) broth (see Appendix) at 37°C, statically. All *E. coli* strains were grown in Lysogeny Broth (LB) or Terrific Broth (TB) (see Appendix) at 37°C, aerated by shaking at 250 rpm. *Enterococcus* and *E. coli* strains containing the plasmids pORI10, pPRGW, or derivatives of either of these were grown in the presence of 50 μg/ml erythromycin (Sigma) for selection.

DNA Manipulation

Plasmid Extraction From E. coli

E. coli was inoculated in 25 ml of LB in Becton-Dickinson Falcon™ 50 ml conical centrifuge tubes. Cultures were incubated overnight for 12 to 18 hours at 37°C while shaking at 250 rpm. The overnight cultures were then divided into two 12.5 ml portions in Becton-Dickinson Falcon™ 15 ml conical centrifuge tubes and pelleted in a Fisher Scientific™ Centrific 225 centrifuge at 6000 rpm for 10 minutes at 4°C.

Table 1. Bacterial strains used in this study

Strain	Description	Reference
E. coli		
DH5α	Cloning host: [F^- *endA1 hsdR17* (r_k^- m_k^+) *supE44 thi-1* λ^- *recA1 gyrA96 relA1 Δ*(*lacZYA-argF*) *U169 Φ80dlacZΔM15*]	Gibco BLR
BL21	Protein expression host: [F^- *ompT hsdS$_B$* (r_B^- m_B^-) *gal dcm ton*A]	Novagen
BL21(DE3)	Protein expression host: [*fhuA2 [lon] ompT gal (λ DE3) [dcm]* Δ*hsdS*] *(λ DE3) = λ sBamHIo* Δ*EcoRI-B int::(lacI::PlacUV5::T7 gene1) i21* Δ*nin5*	NEB
E. faecalis		
OG1RF	Plasmid-free laboratory reference strain: [Gel^+, Spr^+, Rif^R, Fus^R]	Dunny, *et al.*, 1978

The supernatant was then discarded. Pelleted cells were then either processed immediately, or stored for up to 3 days at -20°C. Plasmid extraction was performed using the Promega Wizard® Plus Midipreps DNA Purification System kit, following a modified procedure based on the manufacturer's instructions.

The cell pellet was first resuspended in 3 ml of the provided resuspension solution. Then 3 ml of the provided lysis solution was added and the tube inverted at least five times to mix. In order to prevent DNA damage, 3 ml of the provided neutralization solution was added within two minutes and the tube mixed by inversion at least five times. The lysis mixture was then separated in a Fisher Scientific™ Centrific 225 centrifuge at 6000 rpm for 10 minutes at 4°C. After centrifugation, the supernatant was collected by straining through cheesecloth to filter out aggregated cell debris.

The provided binding resin is resuspended and 3 ml is then added to the collected supernatant and mixed by swirling gently. A provided midicolumn is then connected to a vacuum manifold, and the resin mixture is transferred to the column. The vacuum is applied until all the liquid passed through the column and then stopped to prevent over-drying of the resin. The resin is then washed twice with 15 ml of the provided wash solution and once more with 3 ml of Pharmco 100% ethanol by applying vacuum to draw the solution through the column.

The column is then removed from the manifold and the lower portion of the column is cut away from the upper portion. The lower portion of the column is placed in an Eppendorf microcentrifuge tube and centrifuged in an Eppendorf 5415C tabletop centrifuge at 14,000 rpm for 5 minutes. The column is then removed from the tube and the tube is discarded. The column is placed in a new microcentrifuge tube and 125 μl of

Baxter sterile water preheated to 70°C is added to the column. The column is then centrifuged at 14,000 rpm for 1 minute. The column is then discarded and the flow-through is centrifuged again at 14,000 rpm for 5 minutes to separate the fine resin particles. The plasmid-containing supernatant is then transferred to a clean microcentrifuge tube and was stored at -20°C. Plasmid isolation was confirmed by visualization with agarose gel electrophoresis and ethidium bromide staining.

Agarose Gel Electrophoresis

The presence and size of DNA molecules was visualized by agarose gel electrophoresis. The electrophoresis apparatus was assembled by placing a gel tray into a FisherBiotech Mini-Horizontal (FB-SB-710) electrophoresis cell unit. An appropriate comb was then placed into the gel tray. Agarose powder (Fisher BioReagents®) was dissolved in 1x TAE buffer (see Appendix) at a concentration depending on the size of the DNA molecule to be examined. A gel of 0.8-1% agarose was used for visualizing plasmids, while a 1.5% agarose gel was used to visualize PCR products. The agarose was dissolved in 1x TAE buffer by placing the mixture in a microwave and heating for 1 minute.

The samples were prepared by adding 2-4 µl of 6x loading dye (see Appendix) to 10-20 µl of sample. In some cases, the sample was diluted 1:10 and 2-4 µl of dye added to 10-20 µl of the dilution. The entire dye-sample mixture was then added to a well. In order to gauge the size of the DNA, 10 µl of Fisher Biotech 1 kb Full Scale DNA ladder was added to the first well. The electrodes on the electrophoresis cell were connected to the appropriate terminals of a FisherBiotech FB105 power source.

Table 2. Plasmids used in this study

Plasmid	Description	Reference
pCF10	Pheromone-responsive to cCF10: Tn*925* encodes *tetM* (Tet^R); contains PrgW	Dunny, *et al.*, 1981
pVA891	*E. coli* plasmid: Erm^R	Macrina, *et al.*, 1983
pORI10	Derivative of pVA891. Contains *prgW* which functions as the pCF10 origin of replication: Erm^R	Meloni, *et al.*, submitted
pMSP6050	Derivative of pVA891. Contains pCF10 minimal stability replicon consisting of *uvrA, uvaE, uvaF, prgN, prgO, prgP, prgW, prgZ*, and truncated *prgY*: Erm^R	Hedberg, *et al.*, 1996
pET28b	*E. coli* expression vector. Encodes an insert with His_6 tag: Kan^R	Novagen
pET28b-*prgW*	Derivative of pET28b. Encodes *prgW* with N-terminal His_6 tag	Dunny Lab
pET28b-N122	Derivative of pET28b. Encodes the first 122 amino acids of PrgW with N-terminal His_6 tag	This study
pMSP3535	*E. coli* and *E. faecalis* shuttle vector. Nisin inducible promoter: Erm^R	Bryan, *et al.*, 2000
pORI-78	Derivative of pORI10. *prgW* contains C78A substitution: Erm^R	Utter, thesis
pORI-275	Derivative of pORI10. *prgW* contains C275A substitution: Erm^R	Utter, thesis
pORI-307	Derivative of pORI10. *prgW* contains C307A substitution: Erm^R	Utter, thesis
pORI-78/275	Derivative of pORI10. *prgW* contains C78A and C275A substitutions: Erm^R	Utter, thesis
pORI-78/307	Derivative of pORI10. *prgW* contains C78A and C307A substitutions: Erm^R	Utter, thesis
pORI-275/307	Derivative of pORI10. *prgW* contains C275A and C307A substitutions: Erm^R	Utter, thesis
pORI-78/275/307	Derivative of pORI10. *prgW* contains C78A, C275A, and C307A substitutions: Erm^R	Utter, thesis
pPRGW	Derivative of pMSP3535. *prgW* controlled by the nisin inducible promoter: Erm^R	Meloni, *et al.*, submitted
pPRGW-78/275/307	Derivative of pPRGW. *prgW* contains C78A, C275A, and C307A substitutions: Erm^R	Utter, thesis

The gel was then run at 70 to 100 volts for approximately 45 minutes to 2 hours, depending on the speed of the dye front. Following electrophoresis, the gel was placed into a 1x TAE bath containing 0.5 μg/ml ethidium bromide for 20 minutes in order to stain the DNA. The gel was then visualized using the BioRad Molecular Imager® Gel-Doc™ XR ultraviolet imaging system, as per the manufacturer's instructions.

DNA Extraction From the Agarose Gel

In order to minimize damage done to the DNA, the gel was exposed to UV light for as little time as possible. The band to be extracted was excised using a clean razor blade or scalpel, and the removed block placed into a Eppendorf™ 1.5 ml clear microcentrifuge tube. The gel fragment was then weighed and the DNA extracted using the Qiagen QIAEX®II Gel Extraction Kit following the Qiagen protocol.

DNA Primer Design

Primers were designed using the IDT Technologies™ website. Lyophilized primer stocks were reconstituted in sterile water to a final concentration of 100 μM. Primer working dilutions (1:10 dilution in sterile water) were made for use in PCR reactions. The primers used in this study are listed below.

Polymerase Chain Reaction

In order to amplify a piece of target DNA, a polymerase chain reaction (PCR) was carried out. The PCR mixture was prepared on ice and consisted of 0.25 μl of Promega

GoTaq Polymerase, 10 μl of Promega 5x GoTaq colorless buffer, 2.5 μl of 10 mM dNTP mixture, 2.5 μl of the forward primer, and 2.5 μl of the reverse primer. The DNA template to be amplified was added until the reaction mixture was a total of 50 μl. Between 25 and 50 μl of mineral oil was added on top of the reaction mixture. The PCR reaction was carried out in a MJ Research, Inc. PTC-100 thermocycler, according to the parameters below. After polymerase chain reaction, the samples were purified using the Qiagen QIAquick® PCR Purfication Kit, following the manufacturer's instructions.

Restriction Enzyme Digestion

Restriction enzyme digestion was performed on purified plasmids using enzymes and buffers prepared by New England Biolabs. The digestion conditions depend on the enzyme used, and are detailed in the table below. The digested plasmids were then separated from the removed insert by agarose gel electrophoresis and excised as described above.

DNA Ligation

The plasmid vector and cloned insert were ligated using T4 DNA-ligase (Promega). A molar ratio of 3:1 insert to vector was used, and approximately 150 ng of vector DNA was used. The amount of insert used in the reaction was calculated as follows:

Insert mass (ng) = ((150 ng vector × insert size (kb))/vector size (kb)) × 3

The ligation mixture was then prepared by mixing together 100 ng of vector, the

Table 3. Primers used in this study

Name	Description	Sequence	T_m (°C)
jlc006	Forward primer for the amplification of the N-terminal domain of PrgW.	5'-GCATATGATAATCGTAAACTACCAGTGAG-3'	55.3
jlc007	Reverse primer for the amplification of the N-terminal domain of PrgW.	5'-GGAATTCCTAAGGTTTTCCGAGATATAGC-3'	54.9

calculated amount of insert, 1 µl of Promega 10x T4 Ligation Buffer, and 0.5 units of Promega T4 DNA Ligase. Baxter sterile water was added until the total volume of the mixture was 10 µl. The reaction mixture was then allowed to incubate at room temperature for 3 hours. The ligation was confirmed by agarose gel electrophoresis, and the plasmid extracted from the gel as above.

Transformation of E. coli *BL21(DE3)*

A provided tube of *E. coli* BL21(DE3) was placed in ice and allowed to thaw for 10 minutes. Transformation was performed as per the manufacturer's protocol, located at https://www.neb.com/protocols/1/01/01/transformation-protocol-c2527.

Table 4. PCR thermocycler program

Step	Temperature (°C)	Time
1	95	2 min
2	95	30 sec
3	50	30 sec
4	72	1 min/kb
5	Repeat steps 2-4 35x	
6	72	10 min
7	4	Indefinitely

Table 5. Restriction enzymes used in this study

Enzyme	Site	Manufacturer	Catalog #	Buffer	Temperature	Heat Shock
EcoRI	G/AATTC	NE Biolabs	R0101S	Buffer 4	37°C	65°C
NdeI	CA/TATG	NE Biolabs	R0111S	Buffer 4	37°C	65°C

Protein Manipulation

Preparation of Whole-Cell Lysate

Enterococcus faecalis cultures were inoculated in 50 ml of Todd-Hewitt (TH) broth, with appropriate antibiotic, in Becton-Dickinson Falcon™ 50 ml conical centrifuge tubes. Cultures were incubated overnight for 15-18 hours at 37°C. Cultures were then diluted 1:10 in fresh TH broth containing the appropriate antibiotic and the optical density was measured using a Spectronics Genesys 5 spectrophotometer. The culture was incubated to mid-exponential phase (an OD_{600} of approximately 0.5). During incubation, bead-beater tubes were prepared by adding 1.5 g of Research Products International Zirconia-Alumina beads to 2 ml Sarstedt Micro Tubes. Sampling was performed by removing cell culture from the incubator and collecting it evenly into two 50 mL Falcon™ tubes. The amount of cell culture removed was adjusted depending on the optical density, using the following formula:

$$\textbf{Amount Cultured} = (1/OD_{600}) * 2 * \textbf{Desired Concentration}$$

The tubes were pelleted in an Eppendorf 5430 tabletop centrifuge at 5000 rpm at room temperature for 5 minutes. Afterwards the tubes were collected, and the supernatant was discarded. The cells were then re-suspended in 2 ml lysis buffer (see Appendix) that hand been chilled in ice. The re-suspended cells were added to the prepared bead-beater tubes, and the tubes were placed inside a Biospec Priducts mini-bead-beater. They were then beat at maximum speed for 1 minute and cooled on ice for 30 seconds afterwards. This beating procedure was repeated twice more, and then each supernatant was carefully collected and placed in a Eppendorf™ 1.5 ml clear microcentrifuge tube. The tubes were then either used for SDS-PAGE immediately, or stored at -20°C for up to 3 days.

SDS-PAGE Separation

A 10x stock solution of SDS tank buffer (see Appendix) was diluted 1:10 in DI water to prepare a 1x working solution. The 1x tank buffer was gently mixed for 1 to 3 hours in order to remove air bubbles. Two clean 18 cm x 16 cm glass plates were washed with detergent and tap water, rinsed with DI water, and then wiped with a Kimwipe™ (Kimberly-Clark) soaked with 70% ethanol (Fisher) to remove protein impurities. Two 0.75 mm spacers were then wiped with 70% ethanol and placed on either end of one of the glass plates. The second glass plate was placed on top of the first one and carefully aligned. Two side clamps were secured to the sides of the plate apparatus and the whole assembly was secured to a casting stand.

A 12% acrylamide resolving gel (see below) was mixed and poured into the plate apparatus. At least 1 mL of DI water was added to the top of the resolving gel until it set, at which point the water was removed. A 5% acrylamide stacking gel (see below) was then mixed and poured on top of the resolving gel, and a 12- or 15-well 0.75 mm comb was placed inside. The stacking gel was allowed to set. The comb was then removed, and the apparatus was placed inside the electrophoresis chamber and the tank buffer was gently poured on top of it and into the chamber.

Samples were prepared by collecting 20 μl of whole cell lysate in a 1.5 ml microcentrifuge tube and mixing it with 4 μl 5x loading dye (see Appendix) and 1μl beta-mercaptoethanol (Fisher Biotech). The sample tubes were then clipped shut and placed in boiling DI water for 3-5 minutes. The samples were then loaded into the SDS-PAGE gel, along with GE Amersham Full-Range Rainbow Molecular Weight Markers. The 5x loading dye was diluted 5:1 with DI water and the resulting 1x dye was then added to any empty wells in order to reduce spreading.

The electrodes of the chamber were then connected to a power source (either GE Amersham EPS601 or Hoefer Scientific PS1500), and current was applied. Overnight runs were set to run at 70 volts (400 mA, 10W) for 15-18 hours. Faster runs were set to run at 80 volts (400 mA, 10W) until the dye front reached the top of the resolving gel, and then the voltage was increased to 180 volts. After the dye front ran out of the bottom of the gel, the power was stopped, and the apparatus was removed from the chamber. The glass plates were then separated and the gel removed.

Silver-Staining Procedure

The acrylamide gel was removed and placed in a clean glass dish. Silver staining was performed with a Pierce SilverSNAP Silver Staining Kit. During all steps in this process, the gel was not touched, except with clean gloves or glass or plastic implements. In addition, the dish was placed on a rocker and rocked gently during all steps of this procedure. Reagents were discarded either using vacuum, or were poured out while securing a corner of the gel with a gloved hand. The manufacturer's protocol was followed. The gel was first washed twice for 5 minutes each with DI water. It was then fixed with a 30:10:60 mixture of 100% ethanol (Pharmco), glacial acetic acid (Fisher Chemicals), and DI water. Fixing was performed by exposing the gel to the mixture for 15 minutes, then discarding the mixture, and exposing it again for at least 15 minutes.

Table 6: Acrylamide gel formulations

Acrylamide gel (%)	40% acrylamide, 0.8% bis-acrylamide (ml)	4x Resolving or Stacking buffer[a] (ml)	DI H_2O (ml)	10% Ammonium Persulfate (µl)	TEMED (µl)
Stacking gel					
Stacking	**1.124**	**2.5**	**6.16**	**100**	**10**
Resolving gel					
10	3.75	3.75	7.5	50	10
12	**4.5**	**3.75**	**6.75**	**50**	**10**
15	5.63	3.75	5.62	50	10
18	6.75	3.75	4.5	50	10

a: see Appendix

This second fixing step could be extended overnight. The gel was then washed twice for 5 minutes each in 10% ethanol, and then twice more for 5 minutes each in DI water, discarding the washing reagent each time. A sensitizer solution of 50 µl Sensitizer mixed with 25 ml of DI water was prepared. In addition, a staining solution of 0.5 ml Enhancer mixed with 25 ml stain was prepared at this time. The sensitizer was added to the dish and the gel was allowed to incubate for 1 minute. The sensitizer solution was then discarded and the gel was washed in DI water twice, for 1 minute each wash. The water was discarded and the staining solution was added. The gel was incubated in the staining solution for at least 30 minutes, though this could be extended to overnight without affecting the stain performance. During this time, the developing solution was prepared by mixing 0.5 ml Enhancer with 25 ml of Developer. The stain was then discarded and the gel washed twice with DI water for exactly 20 seconds each. The developing solution was then added, and the gel developed until bands could be seen. The developer solution was then discarded, and the reaction stopped by washing the gel with 5% acetic acid. The wash was discarded, and the gel was then incubated in 5% acetic acid until the results were recorded.

Semi-Dry Electroblotting Procedure (Western Blotting)

A Fisher Biotech Semi-Dry Blotting Unit (FB-SDB-2020) was cleaned by wiping the plates with DI water. Three sheets of Whatman™ 3MM chromatography paper were cut to a size larger than the gel and soaked in Towbin buffer (see Appendix). The chromatography paper was then removed from the buffer and placed on the bottom plate of the apparatus, being careful to avoid air bubbles. The acrylamide gel was removed from the electrophoresis chamber and placed in a clean glass dish filled with Towbin buffer. The gel was allowed to incubate in the buffer for up to 5 minutes before being removed and placed carefully onto the chromatography paper. 3-5 ml of Towbin buffer was added on top of the gel. A sheet of GE Amersham Hybond™-ECL nitrocellulose membrane was cut to the size of the gel and added to the Towbin buffer briefly to soak. The membrane was then carefully placed on top of the gel, avoiding the formation of air bubbles. Three more sheets of chromatography paper were soaked in Towbin buffer and placed on top of the membrane. Excess air and buffer were gently squeezed out by rolling, and paper towels were used to remove buffer from around the paper. The lid of the apparatus was then secured firmly, and the electrodes of the apparatus was connected to a Hoefer Scientific PS1500 power source. Current was applied for 1 hour at 10 volts (400 mA, 10 W).

Blocking and Antibody Incubation

During the blotting procedure, 5% blocking buffer was prepared by adding 5 g of evaporated milk to 100 ml of 1x TBS-Tween (see Appendix). The nitrocellulose membrane was removed from the blotting apparatus and placed in the blocking buffer. The membrane was incubated in the blocking buffer for at least 1 hour, though the blocking could be extended to overnight without affecting performance, if stored at 4°C. The primary antibody was prepared by adding either 1:2000 (for the anti-N-terminal antibody) or 1:4000 (for the anti-C-terminal antibody) of primary antibody to 50 ml of 1x TBS-Tween. The blocking buffer was discarded and the membrane was washed twice with 1x TBS-Tween. The membrane was then incubated in the primary antibody for 1 hour. In the case of a weak-binding antibody, the primary antibody incubation was extended to overnight at 4°C. During this incubation, the secondary antibody was prepared by adding 1.67 µl of Li-Cor® Odyssey® Goat anti-Rabbit fluorescent antibody to 50 ml of 1x TBS-Tween with 3% milk. Both the secondary antibody and the prepared solution were protected from light. The primary antibody solution was discarded, and the gel was again washed twice with 1x TBS-Tween, and then washed three more times with 1x TBS-Tween for 5 minutes each. The gel was then incubated in the secondary antibody solution (protected from light) for 1 hour. The secondary antibody solution was then discarded and the gel washed twice with 1x TBS-Tween, and three times with 1x TBS-Tween for 5 minutes each. The gel was allowed to remain in the TBS-Tween solution until detection.

Detection

Detection occurred using a LiCor® Odyssey® Infrared Scanner. The nitrocellulose membrane is placed in the lower left-hand corner scanner face, making sure that no air bubbles are present. The size of the membrane is noted. The cover is then closed, and the Odyssey v3.0 computer program is run. A new file is created for the scan, and the scanning dialog box is opened. The scanning preset is set to "Membrane." The scanning area is set to the size of the membrane to be scanned, and the membrane is scanned on channel 700 with an intensity of 4.5. Scanning is then begun. If the resulting scan is too bright or dim, the scan is repeated adjusting the intensity as necessary.

Time-Course Assay

The time course assay is performed as follows. *Enterococcus faecalis* OG1RF is inoculated into 50 ml of TH broth (with erythromycin added to a concentration of 50 µl/ml) in a 50 ml Falcon™ tube and incubated in a water bath at 37°C overnight for 12-18 hours. At least 30 minutes prior to sampling, 450 ml of TH broth (erm 50) is placed in the water bath and allowed to warm to 37°C. The overnight culture is then added to the warmed media, and the bottle is kept in the water bath for incubation. The culture in incubated until the OD_{600} = approximately 0.5. Optical density measurements are performed using a Spectronics Genesys 5 spectrophotometer. The first (uninduced) sample is collected at this time. Each sample is 20x concentration, and sampling is done as follows. First, 2 ml of culture is collected into a spectrophotometer analysis cell, and the OD_{600} is recorded. The OD measurement is normalized to 1.0 and multiplied by 40 ml. The volume of culture indicated by this calculation is divided evenly into two 50 ml Falcon™ tubes. The tubes were spun down in a Eppendorf 5430 tabletop centrifuge at room temperature at 5000 rpm for 5 minutes. The contents of both tubes were combined and re-suspended in 2 ml of Lysis Buffer (see Appendix). Sample processing is continued as per "Preparation of whole-cell lysate," above. After the collection of the uninduced sample, nisin (Sigma-Aldrich) was added to the culture to a concentration of 25 ng/ml. The second sample is collected 15 minutes after nisin induction and process in the same way as the uninduced sample. Sample collection is continued every 15 minutes, until 2 hours after induction. The lysate of each sample is placed in a freezer overnight, and SDS-PAGE performed the next day, and Western Blotting after that.

Conserved Domain Prediction

Conserved domain prediction was performed using the NCBI Conserved domain prediction server located at http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi. Primary sequence data was entered in FASTA format, and the default search settings were used. The PrgW primary structure was obtained from the NCBI protein database (located at http://www.ncbi.nlm.nih.gov/protein/), and is designated YP_195765.1.

Secondary Structure Prediction

Secondary structure prediction was performed using the JPred3 secondary structure prediction served located at http://www.compbio.dundee.ac.uk/www-jpred/. The PrgW primary structure was input as a raw amino acid sequence, which was obtained from the NCBI Protein database, as above. The default analysis settings were used.

Helical wheel projections were created by inputting the primary structure of a predicted α-helix into the RZ Lab helical wheel projection script (http://rzlab.ucr.edu/scripts/wheel/wheel.cgi).

Protein Threading

Protein threading data was obtained using Phyre2 (Protein Homology/Analogy Recognition Engine v. 2.0) located at http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index. The PrgW primary structure data was obtained from the NCBI protein database, as described above. The PrgX protein database structure data was obtained in .ent file format from the Protein Data Bank in Europe (located at http://www.ebi.ac.uk/pdbe/). Comparison of the PrgW primary structure to the PrgX template was performed using the one-to-one threading method, using the default analysis settings.

CHAPTER 3

RESULTS

PrgW is a protein transcribed by a gene carried on the *E. faecalis* plasmid pCF10. It is capable of binding both single- and double-stranded DNA, and is known to have two roles. First, it acts as a replication initiation protein. In this capacity, it binds to the *oriV* site located in the *prgW* gene. It is thought to melt the DNA in this area, allowing host replication machinery to begin to replication of pCF10. However, PrgW also serves as a transcriptional regulator. In this role, it binds to the promoter of the *prgZY* operon, where it is thought to transcriptionally upregulate that operon (Meloni, submitted).

Previous studies have indicated that the binding of pre-cCF10 influences the replication initiation function of PrgW (Meloni, submitted). These experiments demonstrated that the presence of pre-cCF10 allows PrgW function in organisms that normally cannot support pCF10 replication. Another set of experiments demonstrated that a similar effect was observed in PrgW mutants where the three conserved cysteines – C78, C275, and C307 – were mutated into alanines (Utter, thesis). In order to better understand the influences of pre-cCF10 binding and the conserved cysteines on PrgW function, it is important to determine and understand the structure of PrgW.

PrgW Structural Analysis

Prediction of the Molecular Weight and pI of PrgW

In order to relate the structure and function of PrgW, it is necessary to understand how the protein is expected to behave under different conditions. Towards that end, the primary structure of PrgW is input into the ExPASy pI/Mw computation tool (http://web.expasy.org/compute_pi/). This tool calculates the theoretical pI and molecular weight of a protein, based on its amino acid composition (the pI is calculated from the individual pK values of the amino acids, and the molecular weight is calculated from the average isotopic masses of the amino acids). The results of this prediction indicate that PrgW has a theoretical pI of 6.87 and molecular weight of 38,766.40 Da.

Comparison to PrgX X-Ray Crystal Structure

PrgW is a member of the RepA_N family of replication initiation proteins. BLASTP analysis of the PrgW structure indicates that the structure of the protein is similar only to other proteins of the RepA_N family. Further analysis by the Phyre2 homology identification algorithm indicates that this similarity extends only to 77 amino acids in the N-terminal domain of PrgW, while the rest of the protein does not show similarity to any other cataloged structure.

PrgX is a protein also encoded by pCF10 that acts as a transcriptional regulator, inhibiting many of the conjugative transfer genes on the plasmid. As a transcriptional regulator, it possesses the ability to bind DNA. The pheromone signal cCF10 binds to

PrgX: the peptide changes the conformation of PrgX, inhibiting its DNA binding and allowing transcription to occur. PrgX was extensively studied, and its X-ray crystal structure is known (Shi, *et al.*, 2005). Analysis of the cCF10 binding pocket indicates that it is formed by a series of six α-helices. Despite the fact that cCF10 is a very hydrophobic peptide, the inside surface of the binding cleft does not show high hydrophobicity. Instead, five hydrophobic amino acid residues (I82, F86, I200, I283, and L282) are responsible for the interaction between PrgX and cCF10 (Shi, *et al.*, 2005).

Since PrgW also binds both DNA and the cCF10 peptide, comparison of the crystal structure of PrgX to PrgW might provide insight into the higher-order structure of PrgW. However, when the primary structure of PrgW is compared to the X-ray crystal structure of PrgX using Phyre2 one-to-one threading analysis, little information is revealed. When compared to both the A and B chains of PrgX, the predicted PrgW structure shares only 11% identity with the known PrgX structure, with 96.5% confidence. In addition, the only amino acid residues that were aligned between the two structures were amino acids 68-86 in PrgW and amino acids 18-35 in PrgX. Notably, this alignment is located in the proposed DNA-binding domain of each respective protein.

Conserved Domains

PrgW binds to both single- and double-stranded DNA in the *oriV* of *prgW* in order to initiate pCF10 replication (Meloni and Mazeffa, submitted). It also binds to dsDNA in the intergenic region between *prgW* and *prgZY* and increases transcription of *prgZY* (an operon encoding the pheromone receptor PrgZ and the protein PrgY, which prevents self-induction of conjugation) (Ruhfel, *et al.*, 1993, Meloni, unpublished

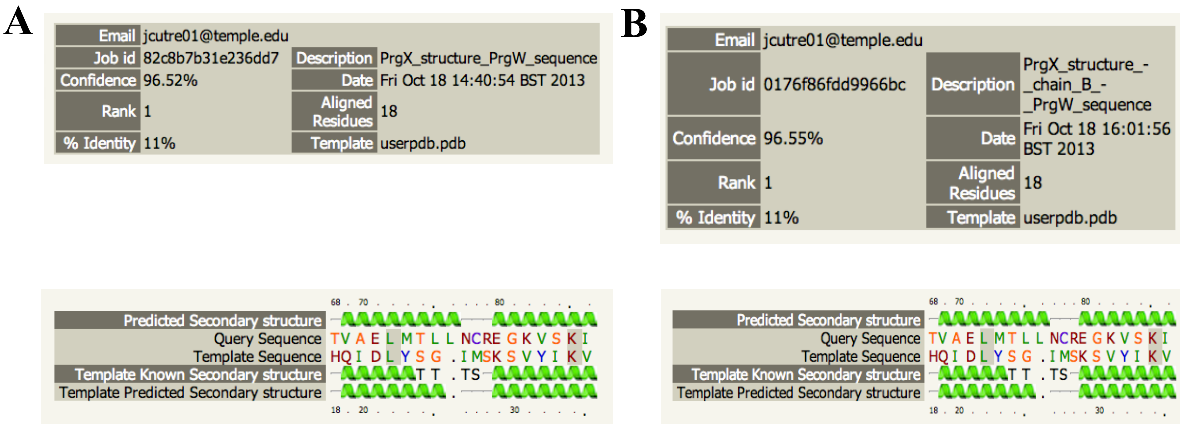

Figure 2. Phyre2 threading data comparing PrgW with the PrgX A-chain (A) and B-chain (B). Alignment data is shown on top, while the schematic of the predicted structure similarity is shown on bottom.

observations). This suggests that the C-terminal end of PrgW does not play a necessary role in DNA binding.

Previous studies indicate that the N-terminal domains of RepA_N proteins contain the DNA binding motif. Kwong, *et al.* performed electrophoretic mobility shift assays (EMSAs) on the Rep protein of the *S. aureus* plasmid pSK41, and demonstrated that the N-terminal segment of that protein is sufficient to bind DNA (Kwong *et al.*, 2004). In addition, a similar study carried out by Francia, *et al.* demonstrated that the N-terminal domain of the RepA protein of the *E. faecalis* pAD1 plasmid was sufficient to specifically bind the *oriV* site (Francia, *et al.*, 2004).

The National Center for Biotechnology Information maintains the Conserved Sequence Database (CDD). The CDD is a collection of known protein primary structures that possess conserved domains – structural regions of high sequence and functional similarity. This database can be used as a resource to aid in the determination of protein function. The interface system accepts a primary amino acid structure as input, and analyzes the sequence using Reverse Position-Specific BLAST (RPS-BLAST), which is a variation of the more commonly used PSI-BLAST (Marchler-Bauer, *et al.*, 2010). It then compares the results to known sequences that are already in the database. Results are returned as annotated regions in the sequence, which are predicted to conform to a conserved domain. The results are ranked in order of: 1) specific hits that have an alignment score higher than a specific threshold, 2) superfamily annotations where the sequence conforms with a particular superfamily present in the database, and 3) annotation by multi-domain models (Marchler-Bauer, *et al.*, 2010).

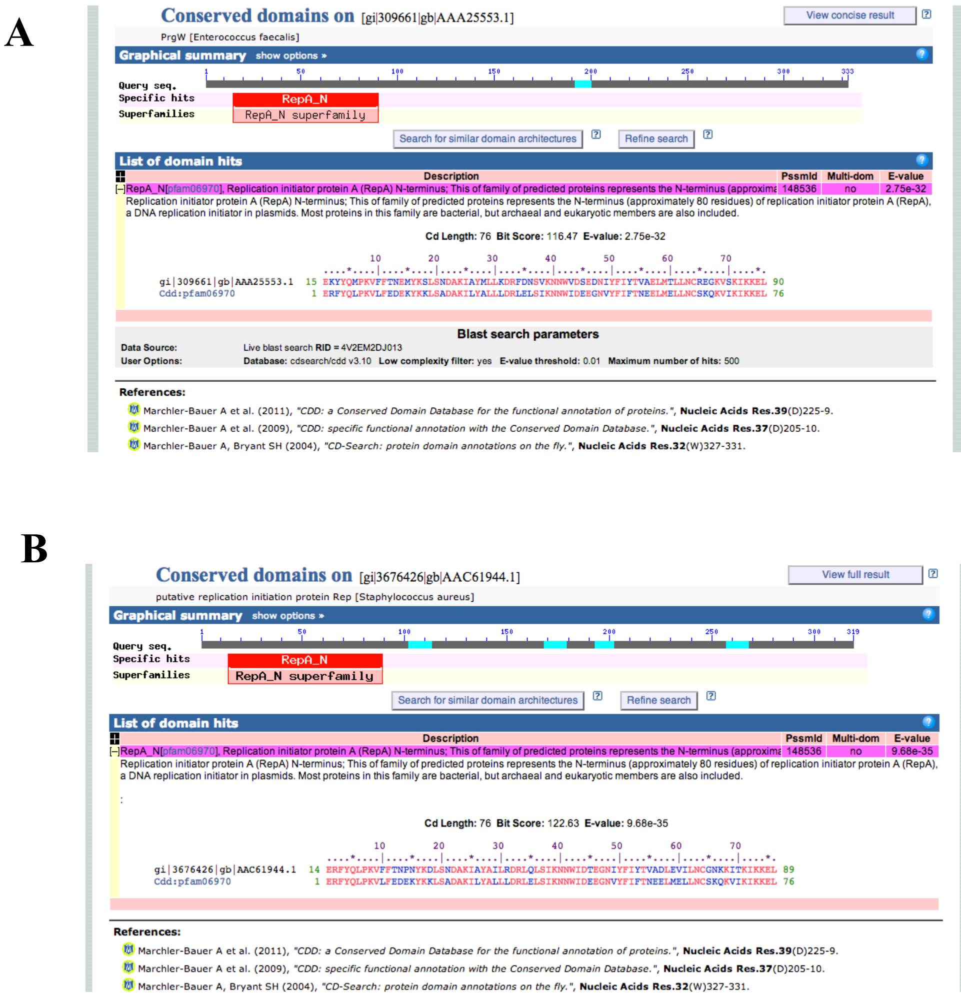

Figure 3. Conserved domain database results for (A) PrgW and (B) RepA of pSK41. Both proteins possess a RepA_N superfamily domain near the N-terminus of the protein. No other conserved domains were found in either protein.

Analysis of the primary structure of PrgW with the NCBI conserved sequence database indicates the presence of a conserved RepA_N superfamily domain within the N-terminal region of PrgW. The RepA_N superfamily is a class of domains that possess a helix-turn-helix (HTH) motif and are most often associated with DNA binding. In PrgW, this conserved domain exists between amino acids 15 and 90. No other conserved domains were identified in PrgW (Figure 3A). The sequence encoding the first 122 amino acids of the N-terminal domain of PrgW was cloned into the expression vector pET28b with the intention of purifying the N-terminal domain for testing in gel-shift assays. The clone was successfully constructed and confirmed by sequencing, however, the protein could not be detected in or recovered by Ni-chromatography under reducing or native conditions from *E. coli* BL21.

This result was compared to the same database search using the pSK41 RepA protein. Like in PrgW, RepA possesses a conserved domain consistent with the RepA_N superfamily between amino acids 14 and 89. Like in PrgW, no other conserved domains were identified in RepA (Figure 3B).

Secondary Structure Predictions

In order to analyze the secondary structure of PrgW, the primary sequence of the protein was entered into the Jpred secondary structure prediction server. While the primary structure of a protein can be determined relatively easily with modern techniques, the gap between the knowledge of primary and secondary protein structure continues to increase (Cole, C., *et al.* 2008). The Jpred 3 Secondary Structure Prediction Server is a internet-based server dedicated to predicting the secondary structure of a protein, based on a primary structure input (either as an amino acid sequence, or a multiple-sequence alignment). The server utilizes an algorithm called Jnet that analyzes the primary structure data and makes a secondary prediction using PSI-BLAST, the Position-Specific Scoring Matrix (PSSM), and the Hidden Markov Model (HMM). The accuracy of the prediction is rated at 81.5%, with the accuracy of solvent accessibility rated at 88.9%, 82.4%, and 77.8%, for >0%, >5%, and >25% relative solvent accessibility thresholds, respectively (Cole, C., *et al.* 2008). Output is displayed as the positions of likely α-helix, β-sheet, and coiled coil structures set on the primary structure.

Jpred predicted 11 α-helices in PrgW, localized to both the N- and C-terminal domains of the protein (Figure 4). The presence of five of the putative helices in the N-terminal domain is consistent with the hypothesis that this region of PrgW is a DNA-binding domain. The remaining six helices are localized to the C-terminal domain.

Helical-Wheel Projections Show Distinct Hydrophobic Regions on α-Helices in the C-Terminal Domain of PrgW

In order to gain insight on the structure and organization of the α-helices in the C-terminal domain of PrgW, helical-wheel projections of the longest three helices in the domain are created (Figure 5). These helices are termed Helix 3, 4, and 5, as they are the third, fourth, and fifth respective helix in the C-terminal domain. Examination of the projection of each of these helices reveals the presence of distinct hydrophobic sections present on the exterior. These hydrophobic regions take the form of a coherent band running along a single section of the helix exterior, and can potentially allow each helix to interact with a hydrophobic region within the protein or a foreign hydrophobic

MENFQFYKSKEIYREKYYQMPKVFFTNEMYKSLSNDAKIAYMLLKDRFDNSVKNNWVDSEDNIYFIY

TVAELMTLLNCREGKVSKIKKELEAVNLLKQKKGRVKKMNGKIETTPNLLYLGKPNVTSEDVFKINE

EQNNVNPVIAKIANTEKVNDINENPVIAKIANTEKVNDINENPVIAKIADNLFYTSYLDTLDTLDTK

KEALQQQLLLDQFSEVQENTFLSKDSLKFIAAFSDTIKEAHEMVGTIIRAKTKVEKEYNVVLIGEDY

QEEIDKCLRRVMHKIKTDSTVKSPKGLFYKSFYNLFVECALEKENVKNRDNSKGSEITLHNWLEK

Figure 4. The positions of α-helices (gray boxes) in the primary structure of PrgW, as predicted by the Jpred secondary structure prediction server. All other amino acids are predicted to conform to a coiled-coil structure.

molecule. One possible arrangement of these helices is shown in Figure 6. It should also be noted that the hydrophobic regions on helices 4 and 5 also contain conserved cysteines 275 and 307, respectively.

A

MENFQFYKSKEIYREKYYQMPKVFFTNEMYKSLSNDAKIAYMLLKDRFDNSVKNNWVDS

EDNIYFIYTVAELMTLLNCREGKVSKIKKELEAVNLLKQKKGRVKKMNGKIETTPNLLY

LGKPNVTSEDVFKINEEQNNVNPVIAKIANTEKVNDINENPVIAKIANTEKVNDINENP

VIAKIADNLFYTSYLDTLDTLDTKKEALQQQLLLDQFSEVQENTFLSKDSLKFIAAFSD

3 4 5

T**IKEAHEMVGTIIRAKTKVEKE**YNVVLIGE**DYQEEIDKCLRRVMHKIK**TDS**TVKSPKGL**

FYKSFYNLFVECALEKENVKNRDNSKGSEITLHNWLEK

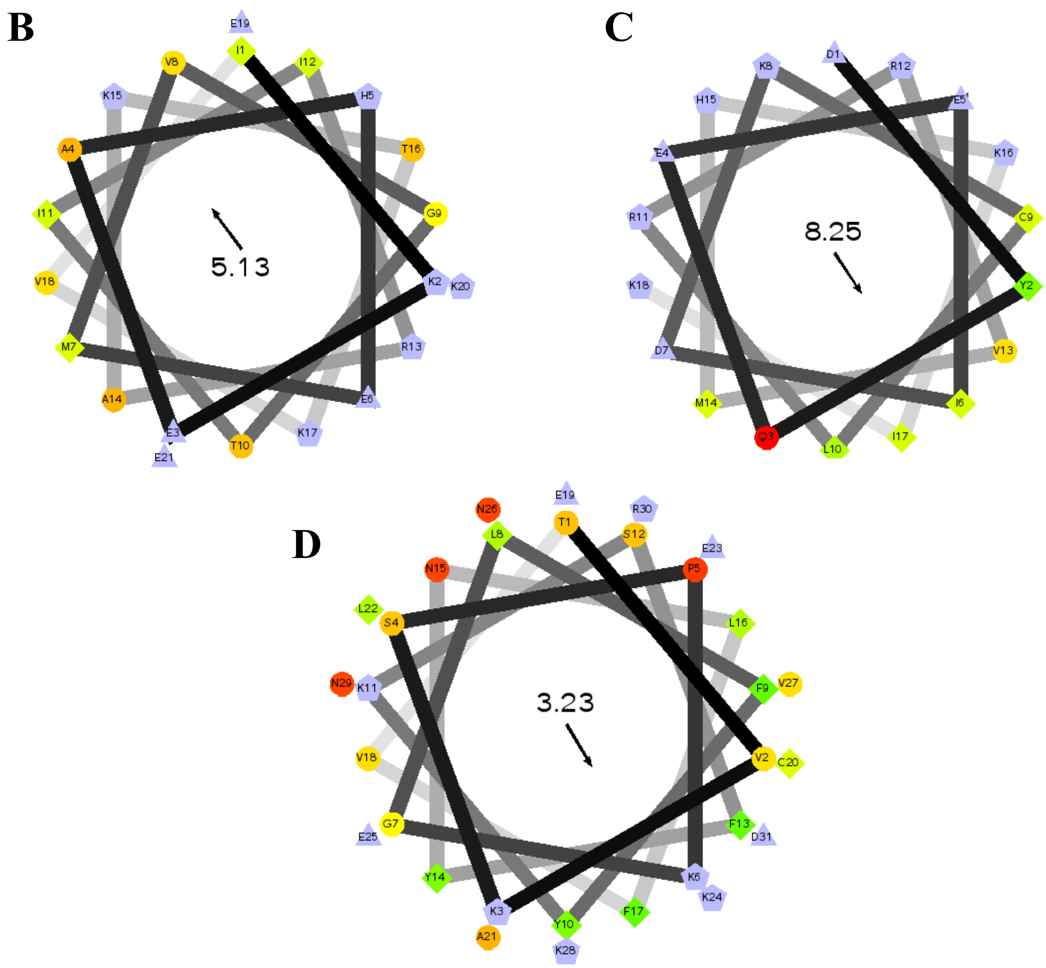

Figure 5. The position and helical-wheel projections of major α-helices in the C-terminal region of PrgW. (A) Positions of helices 3, 4, and 5 in the PrgW primary structure. (B) Helical-wheel projections of helix 3, (C) helix 4, and (D) helix 5. In B, C, and D, the amino acids are labeled by shape and color. Circles indicate polar residues, diamonds indicate hydrophobic residues, petagons indicate basic residues, and triangles indicate acidic residues. Color indicates the degree of hydrophobicity, with green being the most hydrophobic and red being the most polar. Charged residues are depicted as light blue.

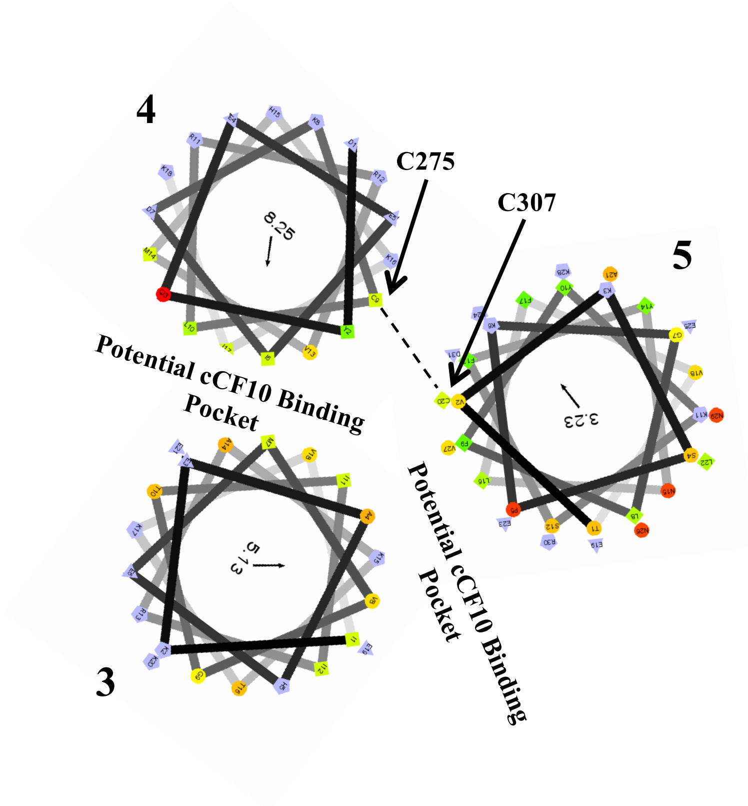

Figure 6. Possible arrangement of α-helices 3, 4, and 5 in the C-terminal domain of PrgW. Note the putative disulfide bond and presence of the cCF10 binding pocket.

Processing of PrgW

Previous Studies

Data from previous experiments in our lab suggested that PrgW might be processed. When PrgW expressed from *E. Faecalis* carrying pCF10 is analyzed by Western blotting, two bands are observed. One band migrates at an apparent molecular weight of approximately 36,000 kDa (p36*), and the other runs at approximately 24,000 kDa (p24*) (Massie-Schuh, thesis). These proteins are difficult to detect and require at least 100 μg/lane of whole cell lysate to be detected by Western blotting. However, when PrgW is expressed from an inducible promoter, two hours after protein induction four different forms were observed: the p36* and p24* bands are present, as well as one running at approximately 40,000 kDa (p40*) and another running at approximately 18,000 kDa (p18*) (Massie-Schuh, thesis). These forms are much more abundant and easily detected in approximately 20 μg of whole cell lysate. To determine if these observed forms were due to processing or are the result of an overexpression artifact, further experiments were done. Unless stated otherwise, all experiments were performed with a cell culture that was incubated to an OD_{660} of 0.5.

p40, p36*, and p24* All Appear at 37°C and Are Stable*

In order to explain the appearance of the different processed forms of PrgW in different strains, the evolution of each band is examined over time. The p40* band of PrgW appears when the protein is over-expressed, but not when it is under a native

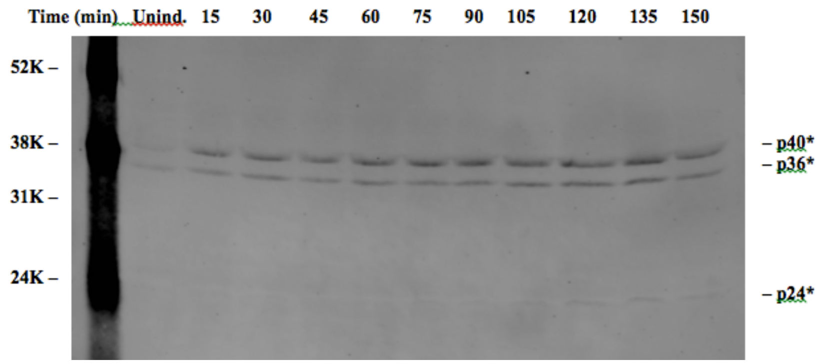

Figure 7. Time-course data of *E. faecalis* OG1RF(pPRGW) induced with nisin and incubated at 37°C. Representative Western blot of three experiments.

promoter. This suggests one of two likely possibilities: that p40*may accumulate in whole cell lysates from the overexpression vector because of a bottleneck in the conversion of p40* to p36*, or that p40* is very unstable and, due to the low amounts of PrgW present in pCF10-containing cells, it turns over too rapidly to be detectable.

Time-course experiments were run on *E. faecalis* OG1RF(pPRGW), a strain that overexpresses PrgW under control of a nisin-inducible promoter. Overnight cultures were diluted 1:100 and induced with nisin when they reached an OD_{660} of 0.05. Whole cell lysates were prepared at various time points post induction. Lysates of all strains were adjusted to an OD_{660} = 1.0 and then concentrated 20X, yielding a protein concentration of 1.0-1.1 mg/ml. The lysates were then loaded on 12% reducing denaturing polyacrylamide gels at approximately 20 μg WCL protein/lane. Western blotting was performed with a polyclonal C-terminal antibody against amino acid residues 313-333 and quantified with a labeled secondary antibody detected by a LI-COR Odyssey scanner. This antibody does not detect p18*, which does not contain the N or C terminus of the protein (Massie-Schuh, thesis).

The first experiment was performed as a two-hour time course, incubating the culture at 37°C, in order to observe the processing pattern over time. Anti-PrgW_C antibodies were used in the Western blot, and recognized the p40*, p36*, and p24* fragments of PrgW (Figure 7). It is observed that the p40* and p36* bands appear early, and the p24* band does not begin to appear until 120 minutes. The presence of p40* and p36* bands in the uninduced sample are because pPRGW possess an leaky promoter and will express a low level of protein even when uninduced.

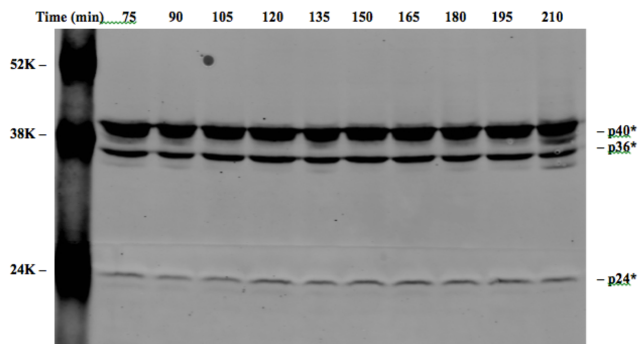

Figure 8. Time-course data of *E. faecalis* OG1RF(pPRGW) induced with nisin, incubated at 37°C, and extended to 3.5 hours. Representative Western blot of two experiments.

In order to determine the dynamics of the PrgW p24* fragment, an extended time-course experiment was performed on *E. faecalis* OG1RF(pPRGW) at 37°C. The samples were collected every 15 minutes, starting at 75 minutes and continuing until 210 minutes. As before, Anti-PrgW_C antibodies detected the p40*, p36*, and p24* fragments (Figure 11A). It is observed that the concentration of the p24* does not change over a longer period of time.

p40 Appears Before p36**

In order to gain a more detailed picture of the early evolution of the bands, it is necessary to examine the early evolution of p40* and p36*. Towards that end, two different approaches were attempted. First, the above experiment was repeated, except that *E. faecalis* OG1RF(pPRGW) was incubated at 30°C in order to examine PrgW processing at a slower rate. As before, the Anti-PrgW_C antibodies recognize both the p40* and p36* fragments and the relative amounts of these fragments are measured (Figure 8). It is observed that the absolute amount of the p40* fragment decreases over time after induction, and the amount of p36* increases. In addition, the p24* fragment is not detectable by the end of the course.

The second approach examines the generation of PrgW after induction with 1/10 the normal concentration of nisin at 37°C. The results are depicted in Figure 9. It is observed that the absolute amount of the p40* fragment increases over time. The p36* fragment appears starting at 45 minutes after culture, and also increases over time. As in the previous experiment, the p24* fragment is not detected by the end of the time-course. Taken together, these experiments indicate that p40* appears before p36*.

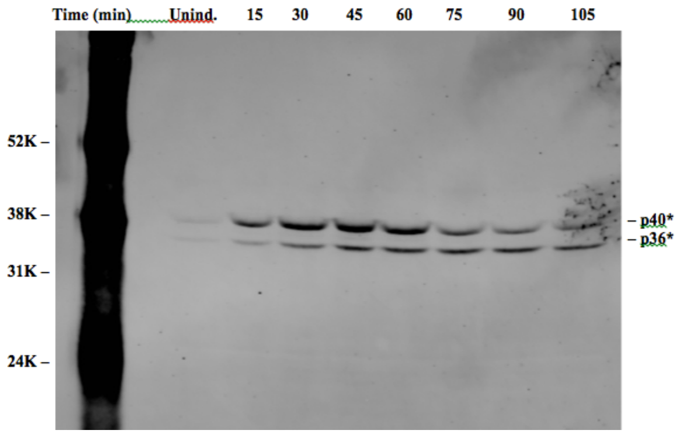

Figure 9. Time-course data of *E. faecalis* OG1RF(pPRGW) induced with nisin and incubated at 30°C. Representative Western blot of data of two experiments.

p40 Disappears After Blocking of Translation and p36* is a Stable Form of PrgW*

In order to better understand why the p40* band is absent when PrgW is expressed under a native promoter, it is necessary to determine whether or not p40* is stable when transcription is halted. Time-course experiments were performed at 37°C while halting protein expression with the antibiotic chloramphenicol. Anti-PrgW_C antibodies detected the p40* and p36* fragments of PrgW and the relative amounts of the fragments were measured (Figure 10). The chloramphenicol was added to the culture 15 minutes after induction. Like in the previous experiments, p40* is strongly induced by 15 minutes after addition of nisin. However, after the addition of chloramphenicol, it is observed that the amount of p40* sharply decreases within 30 minutes. The p40* fragment then decreases and its presence is never re-established. In contrast to this, the presence of p36* is fully established by 30 minutes after induction. Unlike the p40* fragment, the presence of p36* is not decreased by the cessation of transcription. Instead, the amount of P36* remains constant throughout the experiment. As before, the p24* fragment is not detected these experiments.

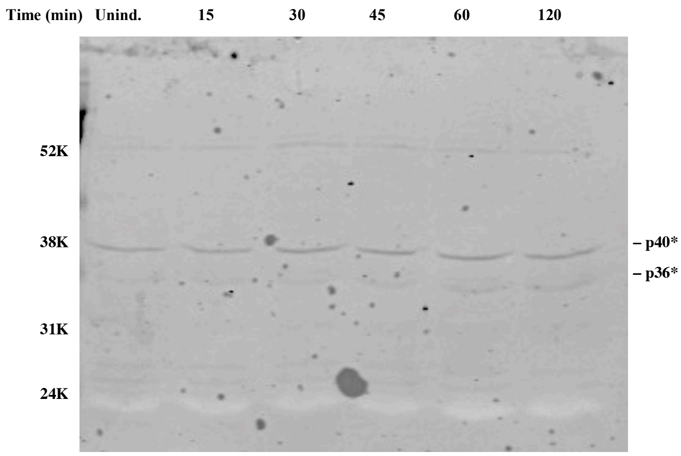

Figure 10. Time-course data of *E. faecalis* OG1RF(pPRGW) induced with low levels of nisin (1/10 normal inducing concentration) and incubated at 37°C. Representative Western blot of two experiments.

*PrgW Cysteine Mutants Accumulate p36**

In order to examine the role of disulfide bonding on the processing of PrgW, experiments were preformed to compare the relative amounts of PrgW fragments in *E. faecalis* OG1RF containing wild-type pORI10 and the various single, double and triple cysteine mutants by Western blotting of late-exponential phase cultures taken at an OD_{660} of approximately 0.8. When examined with the anti-PrgW_C antibody, the apparent amount of the p36* fragment relative to the p40* fragment changes significantly between each mutant strain (Figure 12). Each of the mutants shows a net accumulation of p36* relative to p40*. The C275A and C275A/C307A mutants show the highest retention of p36*, with the remaining mutants show lesser amounts of apparent p36* retention.

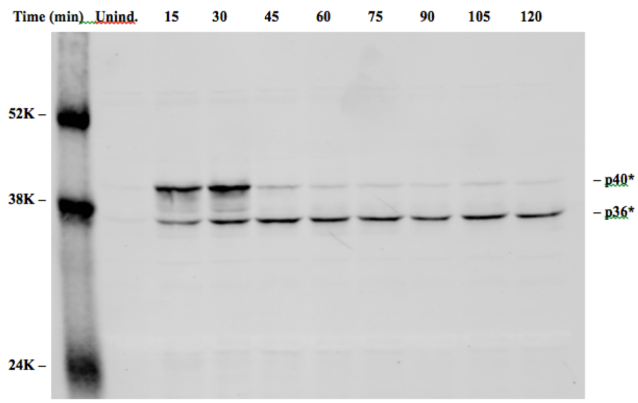

Figure 11. Time-course data of *E. faecalis* OG1RF(pPRGW) induced with nisin and incubated at 37°C. Protein expression was stopped with chloramphenicol added 15 minutes after induction. Representative Western blot of three experiments.

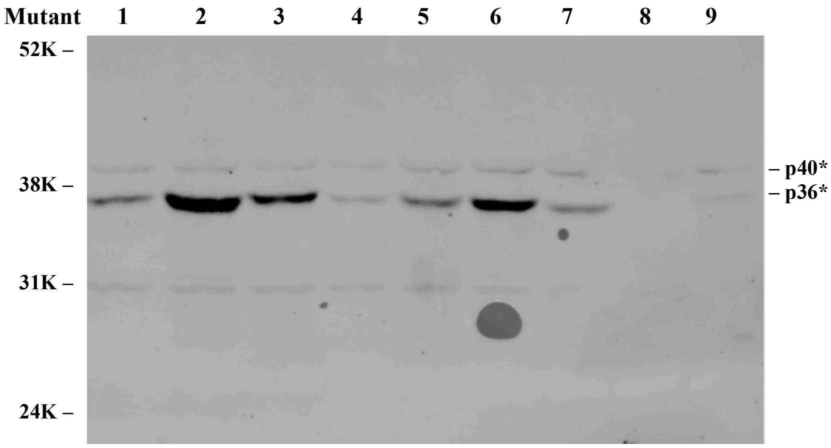

Figure 12. Accumulation of PrgW p36* in *E. faecalis* OG1RF(pORI10) cysteine mutants in late exponential growth ($OD_{660\ nm}$ 0.8). Mutants are as follows: 1 is C78A, 2 is C275A, 3 is C307A, 4 is C78A/C275A, 5 is C78A/C307A, 6 is C275A/C307A, 7 is the triple mutant, 8 contains no plasmid, and 9 is the pORI10 wild-type control. Representative Western blot of cysteine mutants detected with anti-PrgW_C. Data is representative of three experiments.

CHAPTER 4

DISCUSSION

PrgW serves as both the replication initiation protein for pCF10 and a transcriptional regulator of *prgZY* transcription. To gain a better understanding of the possible structure-function relationships of PrgW, *in silico* analysis tools are used to predict the structure of this protein and protein studies were done to gain further insights into its processing. These studies lead to the following conclusions: (1) PrgW possesses a conserved domain near the N-terminal consistent with the DNA-binding domains of other RepA_N proteins. However, no other conserved domains were located on the protein, suggesting that all DNA-binding activity is restricted to the N-terminal region. (2) A putative pre-cCF10 binding pocket exists in PrgW. This pocket is likely hydrophobic in nature and is quite distinct from the binding pocket known to exist in the cCF10 binding protein PrgX. (3) PrgW is detected in two forms when overexpressed – p40* and p36*. The p40* form is unstable and converted to p36*. Since this form is not detected in *E. faecalis* expressing PrgW from pCF10, it is likely to be an accumulation of the nascent form due to a bottleneck in processing of p36*. Since the p24* form does not have a DNA binding motif, it is hypothesized to be a degradation product. (4) The three conserved cysteines (C78, C275 and C307) are not important for the formation of p36*, but their mutation leads to p36* accumulation.

Structure of PrgW

Conserved Domain Analysis of PrgW

Using the NCBI Conserved Domain Database, it is possible to infer the function of a particular domain of a protein by locating a highly similar region of another protein with a known function. The NCBI Conserved Domain Database program revealed that amino acids 15 to 90 on the N-terminal region of PrgW belonged to the RepA_N domain superfamily. Similar analysis of the RepA protein of the *S. aureus* plasmid pSK41 also indicated the presence of a RepA_N superfamily domain from amino acids 14 to 89, also near the N-terminal of that protein. The N-terminal region of RepA is known to bind DNA, specifically the *repA* gene (Kwong, 2004). Thus, it is possible to infer that the N-terminal region of PrgW has a similar DNA-binding function, because of the similarity of the domains.

While this analysis indicates that PrgW possesses one DNA-binding domain, PrgW itself is hypothesized to have three different binding activities. Previous data indicates that it can bind to double-stranded DNA – both that of its own gene during replication and that in the *prgZY* promoter. Additionally, it is predicted to bind single-stranded DNA within its own gene during replication. The fact that this N-terminal region is the only one predicted to bind DNA suggests that it may be responsible for all three activities (Figure 14). This can be tested in future studies by introducing a variant PrgW with the N-terminal domain deleted and observing the replication activity of the gene, as well as EMSA studies to observe the binding. A clone of amino acids 1-122 was constructed in *E. coli* and confirmed by PCR. However, in *E. coli* BL21, no protein could be isolated or detected after induction of protein expression with IPTG. This suggests this particular fragment is unstable in BL21. Specific binding activities can be tested by introducing point mutations into the N-terminal region and observing replication activity, *prgZY* promoter activity, and EMSA results. These data would indicate whether or not the C-terminal domain controls if the protein acts as a transcriptional regulator or replication initiation protein, as well as contributing to protein activity in other ways such as protein multimerization for coating and melting DNA at the origin of replication.

Additionally, p24* does not bind to anti-peptide polyclonal antibodies directed against amino acids 1-20 of PrgW but still reacts with antibodies to the C-terminal 20 amino acids. Taken together, these data suggest that the p24* fragment probably no longer contains the HTH motif. N-terminal sequencing of the fragment could be used to confirm this hypothesis.

The database analysis revealed that no other known conserved domains appear in PrgW. Although negative data, this finding is significant, as it lends support to the hypothesis that pCF10 replicated by a theta-replication mechanism. While the N-terminal domain is a DNA-binding domain, no nickase or other conserved domain associated with rolling-circle replication is present. Analysis of the nucleotide sequence of *prgW* indicates the presence of a series of inverted direct repeats converging on a poly-A tract. The replication proteins of other RepA_N plasmids, such as pAD1 RepA and pSK41 RepA show a similar pattern within their gene sequence. Since previous experiments have demonstrated that PrgW exhibits both double- and single-stranded DNA binding activity specific to the pCF10 *oriV*, it is possible that PrgW binds DNA at the *oriV* and then melts it. This would allow DnaA and other host replication proteins access to the replication site, in a manner similar to RepE in Inc18 plasmids.

Analyses of other replication initiation proteins from theta-replicating plasmids indicate the presence of consistent domains responsible for the recruitment of host DnaA to the replication site (del Solar, 1998). For example, the *Pseudomonas savastanoi* plasmid replication initiation protein pPS10-RepA contains a leucine zipper motif responsible for interaction with DnaA. However, while these domains might be conserved within phylogenetic groups, many such interaction domains exist, and none are conserved across all replication initiation proteins. Thus, it is possible that PrgW utilizes a novel mechanism for interacting with host DnaA, despite the lack of conserved protein interaction domains, or even that such interaction is not required for PrgW to initiate replication. The specific interaction between PrgW and host replication proteins must be elucidated to test this hypothesis.

PrgW Secondary Structure and Binding Pocket Analysis

Analysis of PrgW using the Jpred 3 secondary structure prediction server predicted the presence of 6 alpha-helices in the C-terminal domain. Projection of the three longest alpha-helices present with the Helical Wheel Prediction Program shows that each of these helices has one side on the exterior that is distinctly hydrophobic. Furthermore, it is possible for these helices to be aligned in such a manner as to create a putative interaction site for the hydrophobic pre-pheromone pre-cCF10 (MKKYKRLLLMAGLVTLVFVLSA). It can be hypothesized that a disulfide bond between conserved cysteines 275 and 307 could stabilize an interaction between helices 4 and 5, presenting a distinct hydrophobic face. This face interacts with the hydrophobic region of helix 3, holding it into place. This creates a pocket lined with hydrophobic amino acid side-chains, in which a small hydrophobic peptide, such as pre-cCF10 could enter and bind.

PrgX, the negative regulator of conjugation, binds both mature cCF10 and the inhibitor iCF10. Unlike the hydrophobic binding pocket suggested in PrgW, PrgX binds these molecules in a more hydrophilic cleft, with the interaction being mediated by five hydrophobic amino acid residues. As expected, comparison of the primary structure of PrgW to the known crystal structure PrgX found little similarity between the two structures. No similarity was found in the putative cCF10-binding regions of either of these proteins. Only 18 amino acids aligned between either of these proteins, and in both cases, this alignment occurred within the putative DNA-binding domains of both (Shi *et al.*, 2005). In addition, attempts to thread the PrgW primary structure onto a PrgX template have failed. Thus, the crystal structure for PrgW must be obtained by X-ray crystallography. Attempts to do this are not yet successful, due to poor yield of PrgW upon induction.

The difference between the cCF10 binding pocket domains in PrgW and PrgX are consistent with a difference in the biological functions between pre-cCF10 and mature cCF10. PrgW is the repressor of conjugation. Its activity is controlled by two peptides; iCF10 and cCF10. The peptide iCF10 is produced by pCF10-containing cells and prevents the self-induction of conjugation. iCF10 binds to PrgX and induces tetramer formation, allowing PrgX to act as a repressor. When a donor cell is present, the concentration of cCF10 pheromone is increased. cCF10 displaces iCF10, causing the

PrgX tetramers to dissociate, relieving the repression and allowing conjugation to occur. The binding of pre-cCF10 to a distinct pocket in PrgW would separate the conjugation regulation function from plasmid replication. This idea is supported by the fact that iCF10 inhibits plasmid conjugation, but does not cure the plasmid. If replication were solely influenced by the interaction of PrgW and mature cCF10, then iCF10 would likely inhibit the replication process. However, mapping of the binding site of PrgW with photo-affinity crosslinking and a study of the binding efficacy of pre-cCF10, cCF10 and iCF10 to PrgW must be performed to support this idea. Furthermore, the binding of pre-cCF10 to PrgX must be analyzed to examine the differences in binding.

PrgW Processing

The time-course experiments add more insight into the processing of PrgW. Data from previous experiments indicate that when PrgW is detected by Western blotting of whole cell lysates of *E. faecalis* containing pCF10, two bands react with the antisera; p36* and p24*. However, two additional bands; p40* and p18* forms appear when PrgW is overexpressed in *E. faecalis*. The p40* and p36* forms react with antibodies directed to the N-terminus (aa 1-20) and the C-terminus of PrgW (aa 313-333). The p24* form binds only the anti-C-terminal antibody, while the p18* form binds neither antibody but binds only to a mixture of polyclonal antibodies recognizing PrgW peptides 64-79 and 249-270. The p18* band is often accompanied by smaller degradation products. In addition, the p36* band appears in any system where the replicon is active. Induction time courses were performed to determine the order of the appearance of the bands and protein synthesis inhibition studies were done using the overexpression system to examine the stability of the products.

p40 and p36**

The p40* band and the p36* band appeared within 15 minutes after induction of protein expression. If low culture growth temperatures (30^{o}C) or low concentrations of inducer were used, p40* appeared slightly sooner than p36*. Full-length PrgW purified from *E. coli* runs at an apparent approximate molecular weight of 40,000 after the His-tag has been removed by thrombin cleavage (Buttaro, unpublished observations). This is consistent with the idea that p40* is the nascent protein. In stability assays, the p36* band was very stable and did not significantly decrease in two hours after the inhibition of protein synthesis with chloramphenicol. The p40* band was unstable and degraded within 15 minutes after inhibition of protein synthesis. The p36* band did not increase, suggesting p40* was turned over and not converted to p36*.

One phenomenon to note is that the lysate of an uninduced culture of *E. faecalis*(pPRGW) also displayed the p40* and p36* bands, albeit weakly. This is an effect of the promoter: it is "leaky" and will express protein even without induction. As the p40* band is undetectable with the native promoter even when loading 100 µg of cell lysate, this means that pPRGW overexpresses the protein at least 5-fold. This continuous overexpression of PrgW is most likely the reason we still see the p40* band in uninduced samples.

There are two discrepancies that are visible when PrgW is analyzed with SDS-PAGE. First, neither band appears at PrgW's predicted molecular weight of 38 kDa.

SDS-PAGE works by denaturing proteins and treating them with SDS, an anionic detergent. The SDS binds to the protein, holding it in a denatured form and imparting a negative charge to the molecule. However, there is evidence that SDS does not bind all proteins equally well, particularly those with many hydrophobic residues (Rath, *et al.*, 2009). Furthermore, protein folding is also thought to affect SDS binding, determining which parts of the protein the detergent can access (Rath, *et al.*, 2009). Since PrgW has a number of hydrophobic residues, such as those that potentially form the pre-cCF10 binding pocket, it can be argued that the difference between the predicted molecular weight of PrgW and the observed band might be due to incomplete SDS binding (solvation).

Secondly, p40* and p36* run at different apparent molecular weights from each other. The reason for this is currently unknown. Since PrgW has the potential to form three different possible intramolecular disulfide bonds (C78-C275, C78-C307, and C275-C307), it is quite possible that the differentation is due to incomplete or temporary reduction of the protein. The reducing agent β-mercaptoethanol is argued to be unstable under many different gel conditions, and thus, could have allowed PrgW to oxidize, thus forming disulfide bonds and assuming a folded conformation. This would have led to the appearance of two bands, based on how the disulfide bonds reformed. Since both p40* and p36* appear under non-reducing conditions, this does not indicate that the bands are purely an artifact of SDS-PAGE. In order to fully understand the effects of the reduction conditions, the bands must be studied using different redox reagents, such as dithiothretol (DTT).

Role of the Conserved Cysteines

The conserved cysteines in PrgW (C78, C275, and C307) were not required for conversion of p40* to p36*. However, they appeared to affect the ratio of p40* to p36* (Fig. 12). The ratio of p40* to p36* shifted toward 50:50 with a trend toward decreased levels of p40* in cells expressing PrgW cysteine mutants. This increase in p36* was more readily observed in late exponential cultures of cells expressing PrgW from a high-copy replicon clone (pORI10), where there was a marked accumulation of p36* in whole cell lysates during exponential growth as compared to pORI10 expressing wild-type PrgW (Fig. 12). These data suggest that while the cysteines are not necessary for the conversion of p40* to p36*, they change the ratio in favor of p40*. This increase in the relative amount of p36* in cysteine mutants could be due to (a) increased conversion of p40* to p36* (b) to increased stability of p36* or (c) reduced formation of intermolecular disulfide bonds leaving more protein in the soluble fraction (see Crystal Structure Analysis of Staphylococcal RepA). Because the amount of p40* is probably the result of a balance between conversion to the stable p36* and degradation of p40*, more experimentation is needed to distinguish between these models. Since the C275 and C307 are both within the putative binding pocket for pre-cCF10, pre-cCF10 may also have a role in the turnover of p36*.

*p24**

Since p24* probably does not contain the HTH motif for DNA binding, it is likely to be a turnover product of p36*. While the data suggest that p36* is inherently more stable than p40*, it is formally possible that chloramphenicol interferes with the proteolysis system needed to turn over p36*, but not the system needed to degrade p40* (Sussman, 1969, Rozkov, 2001). This could be tested by checking for the appearance of p24* without the use of chloramphenicol in *E. faecalis* strains that contain mutant ATP-dependent proteases. The p24* form is present in both wild-type PrgW and cysteine-to-alanine substitution PrgW (Nina Alfaro, unpublished observations), suggesting that its formation is not dependent on the cysteines.

Conversion of p40 to p36**

Numerous attempts to purify p36* from *E. faecalis* overexpressing PrgW have not been successful. Four hypotheses could be formed for the conversion of p40* to p36*: (1) alternative folding, (2) proteolytic processing, (3) protein modification, and (4) interaction with pre-cCF10.

Alternative Folding. After translation, proteins are folded into their correct conformation. This process is often performed by chaperone proteins, such as the DnaJ/DnaK complex and the GroES/GroEL complex. Many of these chaperone systems are conserved across different genera and homologs to both DnaJ/DnaK and GroES/GroEL exist in *E. faecalis* (Teng, 2001). This can serve as an explanation of the presence of p40* in overexpression systems. If p40* is normally converted to p36* by a chaperone system, and such a system works at a constant rate, increasing the amount of nascent p40* by overexpression will leave a significant amount of unprocessed p40* in the cell.

Proteolytic Processing. The conversion of p40* to p36* is due to a small cleavage event at the C-terminal end of the protein. There is a conversed motif at the C-terminal end of all members of the RepA_N conjugative plasmid family: HNW[YL]E[NYK][QK]. The sequence in PrgW is HNWLEK and has a predicted molecular weight of 826 Da. Since this motif is highly charged, it could significantly affect the folding of PrgW, and thus, its mobility. Since the antisera used to detect PrgW is a polyclonal sera directed against the last 20 amino acids, this would leave enough of the epitope intact to allow the antibodies to bind. C-terminal sequencing of the purified fragment could be used to address this possibility. In the absence of protein purification, a genetic mutation eliminating the motif could also be used to determine the effect of the motif on PrgW migration.

Post-translational modification. Post-translational modification in bacteria is less common than in eukaryotic cells, and generally increases the apparent molecular weight of the protein. Despite this, post-translational processing is possible, and the conversion of p40* to p36* is due to a modification leading to folding differences that allow for faster migration. MS/MS analysis of fragmented PrgW could be used to detect these modifications.

Interaction with pre-cCF10. A portion of the p40* fragment is turned over into the p36* fragment, while the excess is degraded. As the p40* fragment is not produced to a very high degree under the native PrgW promoter, the band is not observed in *E. faecalis*(pCF10) strains. It is also possible that the nascent PrgW interacts with a limited amount of pre-cCF10 present in the host, which serves to stabilize the protein. The portion of the nascent protein that is not stabilized assumes a different conformation which is visualized as the p40* band. This conformation is inactive and is quickly degraded by host proteases. The apparent accumulation of p36* by PrgW mutants in this case could be the result of altered pre-cCF10 binding efficiency due to the cysteine mutations. Alternatively, as in the first model, the apparent accumulation might be the result of reduced turnover of p36* into p24*, which is supported by the reduced amount of p24*.

Crystal Structure Analysis of Staphylococcal RepA

Very recently (mid-summer 2014), Schumacher, *et al.* published an analysis of the RepA proteins belonging to pSK41 and pTZ2162, two RepA_N plasmids of *S. aureus*. As part of this analysis, the crystal structures of these proteins were determined. The analysis of both Staphylococcal proteins led to the same general conclusions, so I will not describe the minute differences between the two RepA proteins studied.

Structural analysis of the N-terminal domain (i.e. the first ~120 amino acid residues) of RepA showed that it had a helix-turn-helix motif, highly homologous to that of other RepA_N proteins. Furthermore, it was determined that the N-terminal domain bore strong similarity to the DNA-binding domain of DnaD. DnaD is one of the members of the replication initiation complex, responsible for recruiting helicase to the origin of replication. Specifically, the N-terminal domain of DnaD promotes the assembly of multiple DnaD domains into a scaffold. The scaffold recognizes DnaA at the origin of replication and binds the DNA at that location. It then binds the proteins DnaG and DnaI/DnaB to complete the replication initiation complex and recruit helicase. However, unlike DnaD, the RepA N-terminal domain does not promote scaffold assembly, instead binding directly to DNA at the origin of replication.

As is the case for other RepA_N proteins, the specific role of the C-terminal domain (i.e. the last ~120 amino acid residues) in RepA is unknown. The crystal structure of this domain indicates that it consists of a particular arrangement of five distinct α-helices. Furthermore, the analysis concludes that this domain showed strong structural homology to the C-terminal domain of the putative DnaD protein of *E. faecalis* V583. Yeast two-hybrid studies were carried out and it was demonstrated that the RepA-DNA complex interacted with host DnaG. This data led them to conclude that the C-terminal domain of RepA might have a role in the formation of the helicase recruitment complex.

Of particular interest are the many parallels between this analysis of staphylococcal RepA and the findings about PrgW presented in this work. First, the primary structure of PrgW is consistent with the three-domain model of RepA_N proteins, as described previously in detail. The N-terminal domain of PrgW is quite homologous to that of pSK41, thus suggesting similarity to staphylococcal DnaD as well. These findings also suggest a role for the C-terminal domain of PrgW, in the recruitment and assembly of the host helicase recruitment complex.

The pheromone-responsive plasmid pCF10 has two unique features not present in Staphylococcal RepA_N plasmids: conserved cysteines and the ability to bind pre-cCF10, both of which enhance replication initiation function of PrgW. We have recently shown that conversion of PrgW to p36* is necessary to form intermolecular disulfide bonds, which are readily apparent in fresh preparations of PrgW from growing cells (Francisco-Calva, unpublished observations). In addition, growth under conditions of oxidative stress (hydrogen peroxide) increases the copy number of pCF10 from an average of 3 to an average of 6 (Buttaro, unpublished observations), while mutation of the conserved cysteines reduces copy number to an average of 1 (Utter, thesis). Taken together, these data suggest a possible second role of the conserved cysteines – the stabilization of intermolecular interactions at the origin, allowing for more efficient recruitment of helicase for replication. This efficiency of helicase recruitment can explain the differences in plasmid copy number.

Further analysis must be carried out in order to test this hypothesis. First of all, in order to confirm the validity of the model, an analysis of the primary structure similarity between *E. faecalis* DnaD and PrgW should be performed. Secondly, the crystal structure of PrgW should be determined and compared to that of DnaD. Finally, gel-shifting and yeast two-hybrid experiments should be performed in order to determine the interaction profile of the PrgW C-terminal domain.

Comparison of PrgW and DnaD

In order to compare the structure of PrgW with that of *E. faecalis* DnaD, preliminary simulations were performed. The Phyre2 threading algorithm was used to attempt to thread the primary structure of PrgW onto the crystal structure of DnaD, but no similarity was found. In addition, attempts to thread PrgW to Staphylococcal RepA also found no similarity between the two programs. Clustal Omega analysis (http://www.ebi.ac.uk/Tools/msa/clustalo/) was performed to directly compare the primary structures of PrgW and DnaD (Figure 13A). The analysis showed little identity between the two structures, except for a 6-amino acid sequence (TLHNWL), which matched exactly. A consensus alignment between many different *E. faecalis* RepA_N proteins that was previously generated in our lab shows that this 6-amino acid sequence is highly conserved (Figure 13B). It is thought that this sequence might allow PrgW to mimic DnaD, and thus promote replication.

Proposed PrgW function in replication

Based on the data presented in this work, it is possible to form a hypothesis about the function of PrgW (Figure 14). PrgW is first expressed as a nascent peptide (p40*). p40* has no biological function and is degraded quickly after transcription. Some of the p40* interacts with pre-cCF10, which protects it from degradation and causes it to change conformation into the p36* form. By default, the p36* assumes an "open" conformation, in which pre-cCF10 recruits it to the membrane, where it potentially interacts with membrane-associated oxidoreductases which break any disulfide bond formed between C275 and C307. Pre-cCF10 orients the α-helices so that the cysteines are facing outside, where they interact and bonded to other PrgW molecues. PrgW multimers then bind *oriV* and initiate replication by mimicking DnaD. Some amount of PrgW instead loses its interaction with pre-cCF10 and, by some unknown method, forms an intramolecular disulfide bond between C275 and C307. This forces the pre-cCF10 binding pocket closed and this PrgW acts as a transcriptional enhancer for *prgZY* transcription.

A

```
CLUSTAL O(1.2.1) multiple sequence alignment

gi|58616129|ref|YP_195765.1|          MENFQFYKSKEIYREKYYQMPKVFFTNEMYKSLSNDAKIAYMLLKDRFDNSVKNNWVDSE
gi|384512895|ref|YP_005707988.1|      ML------------SIREYLEAGE------------TTISNLL-FDHYQ---KIGL--KD
                                      *             .:* :  :            :.*: :*  *:::    *     .:

gi|58616129|ref|YP_195765.1|          DNIYFIYTVAELMTLLNCREGKVSKIKKELEAVNLLKQKKGRVKKMNGKIETTPNLLYLG
gi|384512895|ref|YP_005707988.1|      DEFLFLLQ------LF--RSQNAGDLFPD-------------LMAIAETMGKTPDTIYKL
                                      *:: *:        *:  *. :...:  :             :  :  .: .**: :*

gi|58616129|ref|YP_195765.1|          KPNVTSEDVFKINEEQNNVNPVIAKIANTEKVNDINENPVIAKIANTEKVNDINENPVIA
gi|384512895|ref|YP_005707988.1|      LNQLVSRGFIRIETQQNQKGQMMD---------TYDLLPVFQKIQLFLQ--TSKEKQVVA
                                        ::.*. .::*: :**:   ::           :  **: **    :    :*: *:*

gi|58616129|ref|YP_195765.1|          KIADN--LFYTSYLDTLDTLDTKKEALQQQLLLDQFSEVQENTFLSKDSLKFIAAFSDTI
gi|384512895|ref|YP_005707988.1|      NHEDEIKQLYQGFEKEF------------------------GRPLSPIELEMIGQWLNTD
                                      :  *:   :* .: . :                        **  .*::*. : :*

gi|58616129|ref|YP_195765.1|          KEAHEMVGTIIRAKTKVEKEYNVVLIGEDYQEEIDKCLRRVMHKIKTDSTVKSPKGLFYK
gi|384512895|ref|YP_005707988.1|      HYSPELIRLALRE-AVLNQAYSLKYI-----DRILLAWER--KNITTKEQVAADQK----
                                      : : *::   :*  : ::: *.:  *     :.*  . .*  ::*.*.. * : :

gi|58616129|ref|YP_195765.1|          SFYNLFVECALEKENVKNRDNSKGSEITLHNWLEK----
gi|384512895|ref|YP_005707988.1|      KRKDSMIQNEIE---QQGQTQESLPKVTLHNWLNPEDSE
                                      .  : :::  :*     : : :..  ::******:
```

B

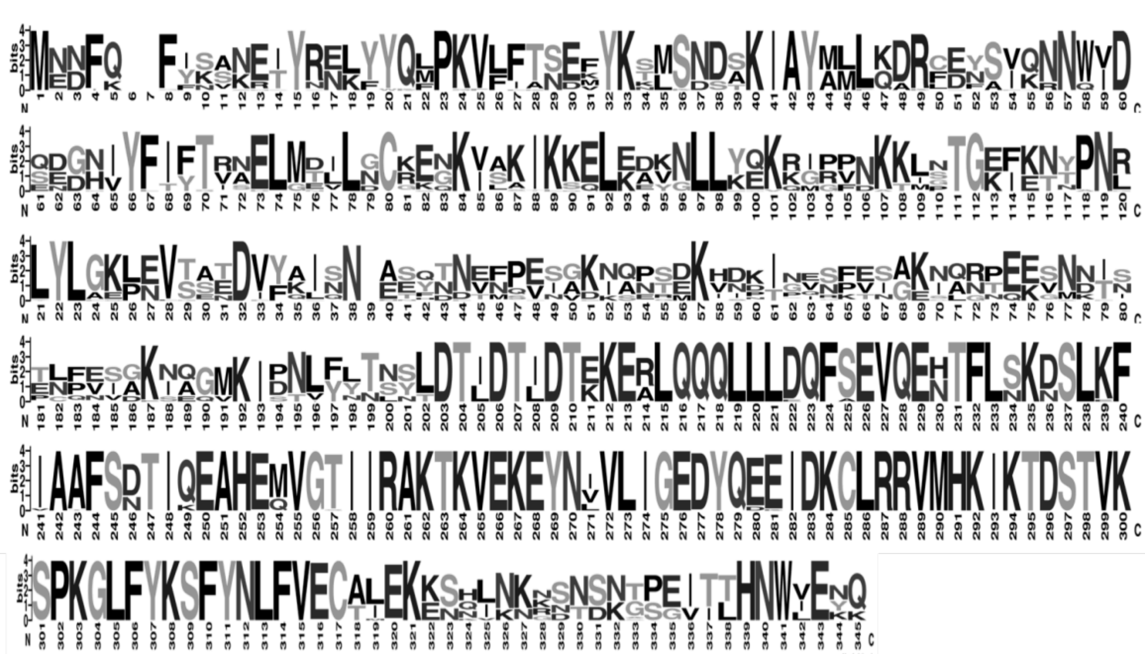

Figure 13. The conserved C-terminal sequence in PrgW. (A) Primary structure comparison of PrgW (top) and *E. faecalis* DnaD (bottom). Six amino acids near the C-terminus (TLHNWL) share common identity. (B) Consensus sequence alignment of RepA_N proteins in *E. faecalis*. Note the high degree of conservation between amino acids 337 and 343.

Conclusion

Taken together, these data add further insight into the structure and function of the unique replication initiation protein PrgW. Although many possibilities exist, these data can be used to formulate the following testable hypothesis for the function of the cysteines and pre-cCF10 in pCF10 biology (Fig. 14). Understanding the interaction of pre-cCF10 with PrgW may allow us to design molecules that will disrupt PrgW structure rather than enhance its replication activity. Since PrgW is essential for pCF10 replication, this would allow curing of this family of important mediators of antibiotic resistance and limiting its spread. Gaining a better understanding and the role of the conserved cysteines in plasmid replication will also give us insights into a unique mechanism of adjusting plasmid copy number in response to redox state of the bacterial cells.

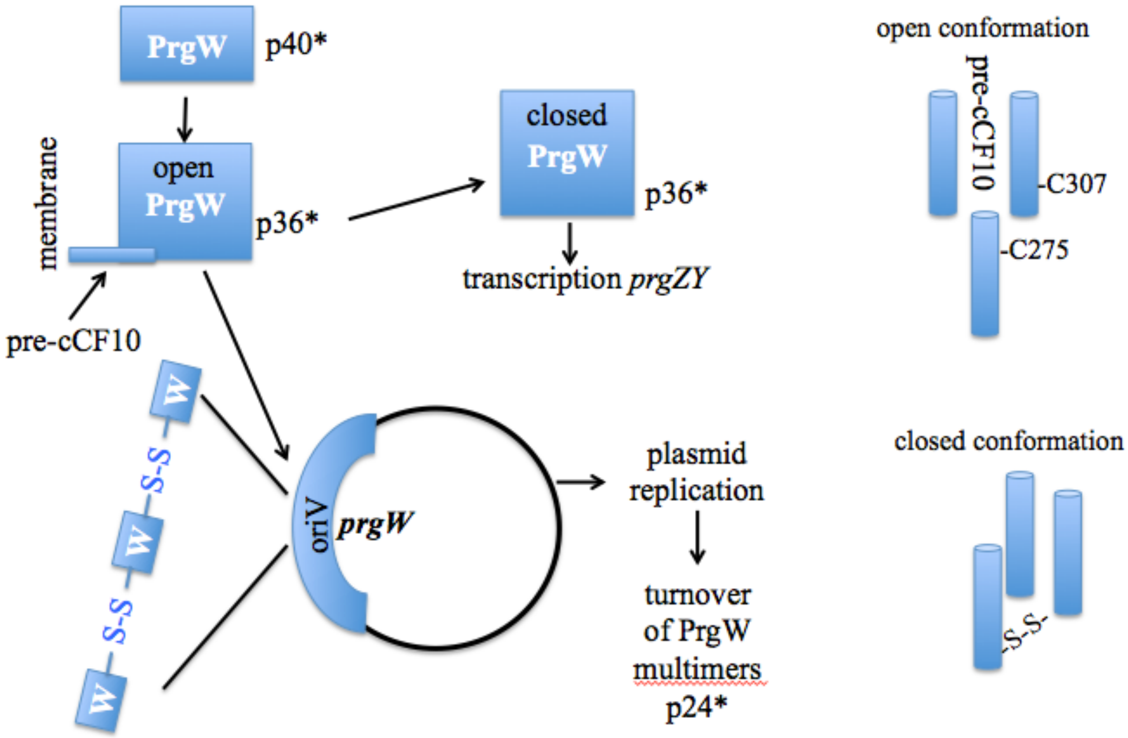

Figure 14. Hypotheses for PrgW function in replication. Pre-cCF10 is present in the cell membrane in limiting concentrations. Pre-cCF10 binds to p40*, promoting its conversion to p36* and recruiting it to the membrane. p36* can then bind to the origin of replication in PrgW. If there are cysteines present, intermolecular disulfide bonds help promote the binding of multiple PrgW proteins to the series of direct repeats in the origin of replication. Once sufficient copies of PrgW are present, the helicase complex assembles and plasmid replication is initiated. p36* is degraded after use as a replication initiation protein. p36* that does not interact with pre-cCF10 forms an intermolecular disulfide bond, promoting *prgZY* transcriptional activity.

REFERENCES

Alvarez-Martinez, C.E., and Christie, P.J. (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. R.* **73**: 775-808.

Antiporta, M.H., Dunny, G.M. (2002) ccfA, the genetic determinant for the cCF10 peptide pheromone in *Enterococcus faecalis* OG1RF. *J. Bacteriol.* **184**: 1155–1162.

Arias, C.A. and Murray, B.E. (2012) The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat. Rev. Microbiol.* **10**: 266-278.

Arthur, M., and Courvalin, P. (1993) Genetics and mechanisms of glycopeptide resistance in enterococci. *Antimicrob. Agents Chemother.* **37**: 1563-1571.

Arthura, M., Reynolds, P., and Courvalin, P. (1996) Glycopeptide resistance in enterococci. *Trends Microbiol*. **4**: 401-407.

Bertram, J., Strätz, M., and Dürre, P. (1991) Natural transfer of conjugative transposon Tn916 between gram-positive and gram-negative bacteria. *J. Bacteriol.* **173**: 443-448.

Burdett, V. (1986) Streptococcal tetracycline resistance mediated at the level of protein synthesis. *J Bacteriol.* **165**: 564-569.

Buttaro, B.A., Antiporta, M.H., Dunny, G.M. (2000) Cell-associated pheromone peptide (cCF10) production and pheromone inhibition in *Enterococcus faecalis*. *J. Bacteriol.* **182**: 4926–4933.

Buu-Hoï, A., and Horodniceanu, T. (1980) Conjugative transfer of multiple antibiotic resistance markers in *Streptococcus pneumoniae*. *J Bacteriol.* **143**: 313-320.

Clark, N.C., Teixeira, L.M., Facklam, R.R., and Tenover, F.C. (1998) Detection and differentiation of *vanC-1*,*vanC-2*, and *vanC-3* glycopeptide resistance genes in enterococci. *J Clin Microbiol.* **36**: 2294-2297.

Christie, P.J., Korman, R.Z., Zahler, S.A., Adsit, J.C., Dunny, G.M. (1987) Two conjugation systems associated with *Streptococcus faecalis* plasmid pCF10: identification of a conjugative transposon that transfers between *S. faecalis* and *Bacillus subtilis*. *J Bacteriol.* **169**:2529-2536.

Clewell, D.B. (1990) Movable genetic elements and antibiotic resistance in enterococci. *Eur J Clin Microbiol Infect Dis.* **9**: 90-102.

Clewell, D.B. and Dunny, G.M. (2002) Conjugation and genetic exchange in enterococci, in The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance. Edited by M.S. Gilmore. ASM Press: Washington, D.C.

Clewell, D.B., Yagi, Y., and Bauer, B. (1975) Plasmid-determined tetracycline resistance in *Streptococcus faecalis*: evidence for gene amplification during growth in presence of tetracycline. *Proc Natl Acad Sci U S A*. **72**: 1720-1724.

Coburn, P.S., and Gilmore, M.S. (2003) The *Enterococcus faecalis* cytolysin: a novel toxin active against eukaryotic and prokaryotic cells. *Cell Microbiol.* 2003 **5**: 661-669.

Cole, C., Barber, J.D., and Barton G.J. (2008) The JPred3 secondary structure prediction server. *Nucleic Acids Res.* **36**: W197-W201.

Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M., and Barton, G.J. (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics* **14**: 892-893.

del Solar, G., Giraldo, R., Ruiz-Echevarría, M.J., Espinosa, M., and Díaz-Orejas, R. (1998) Replication and Control of Circular Bacterial Plasmids. *Microbiol. Mol. Biol. Rev.* **62**: 434-464.

Dunny, G. (2007) The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. *Philos. T. R. Soc. B* **362**: 1185-1193.

Dunny, G.M., Leonard, B.A. (1997) Cell-cell communication in gram-positive bacteria. *Annu Rev Microbiol*. **51**: 527-564.

Dunny, G.M., Leonard, B.A., and Hedberg, P.J. (1995) Pheromone-inducible conjugation in *Enterococcus faecalis*: interbacterial and host-parasite chemical communication. *J Bacteriol.* **177**: 871-876.

Ehrenfeld, E.E., Kessler, R.E., and Clewell, D.B. (1986) Identification of pheromone-induced surface proteins in *Streptococcus faecalis* and evidence of a role for lipoteichoic acid in formation of mating aggregates. *J Bacteriol.* **168**: 6-12.

Facklam, R.R., Carvalho, M.G.S., Teixeira, L.M. (2002) History, taxonomy, biochemical characteristics, and antibiotic susceptibility testing of enterococci, in The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance. Edited by M.S. Gilmore. ASM Press: Washington, D.C.

Fines, M., Perichon, B., Reynolds, P., Sahm, D.F., and Courvalin, P. (1999) VanE, a new type of acquired glycopeptide resistance in *Enterococcus faecalis* BM4405. *Antimicrob Agents Chemother.* **43**: 2161-2164.

Flannagan, S.E., Zitzow, L.A., Su, Y.A., and Clewell, D.B. (1994) Nucleotide sequence of the 18-kb conjugative transposon Tn916 from *Enterococcus faecalis*. *Plasmid.* **32**: 350-354.

Fronzes, R., Christie, P.J., and Waksman, G. (2009) The structural biology of type IV secretion systems. *Nat. Rev. Microbiol.* **7**: 703-714.

Galloway-Peña, J.R., Nallapareddy, S.R., Arias, C.A., Eliopoulos, G.M., and Murray, B.E. (2009) Analysis of clonality and antibiotic resistance among early clinical isolates of *Enterococcus faecium* in the United States. *J Infect Dis.* **200**: 1566-1573.

Gholizadeh, Y., and Courvalin, P. (2000) Acquired and intrinsic glycopeptide resistance in enterococci. *Int J Antimicrob. Ag.* **16**: 11-17.

Gilmore, M.S., Coburn, P.S., Nallapareddy, S.R., and Murray, B.E. (2002) Enterococcal virulence, in The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance. Edited by M.S. Gilmore. ASM Press: Washington, D.C.

Gilmore, M.S., Lebreton, F., and van Schaik, W. (2013) Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. *Curr. Opin. Microbiol.* **16**: 10-16.

Grohmann, E., Muth, G., and Espinoza, M. (2003) Conjugative plasmid transfer in gram-positive bacteria. *Microbiol. Mol. Biol. R.* **67**: 277-301.

Hancock, L.E. and Gilmore, M.S. (2006) Pathogenicity of enterococci, in Gram-Positive Pathogens, Second Edition. Edited by Fischetti, V.A., Novick, R.P., Ferretti, J.J., Portnoy, D.A., and Rood, J.I. ASM Press: Washington, D.C.

Heaton, M.P., Discotto, L.F., Pucci, M.J., and Handwerger, S. (1996) Mobilization of vancomycin resistance by transposon-mediated fusion of a VanA plasmid with an *Enterococcus faecium* sex pheromone-response plasmid. *Gene.* **171**: 9-17.

Hedberg, P.J., Leonard, B.A.B., Ruhfel, R.E., Dunny, G.M. (1996) Identification and characterization of the genes of *Enterococcus faecalis* plasmid pCF10 involved in replication and in negative control of pheromone-inducible conjugation. *Plasmid.* **35**: 46-57.

Hidron, A.I., Edwards, J.R., Patel, J., Horan, T.C., Sievert, D.M., Pollock, D.A., and Fridkin, S.K. (2008) Antimicrobial-resistant pathogens associated with healthcare-associated infections: annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006–2007. *Infect. Cont. Hosp. Ep.* **29**: 996-1011.

Hirt, H., Manias, D.A., Bryan, E.M., Klein, J.R., Marklund, J.K., Staddon, J.H., Paustian, M.L., Kapur, V., and Dunny, G.M. (2005) Characterization of the pheromone response of the *Enterococcus faecalis* conjugative plasmid pCF10: complete sequence and comparative analysis of the transcriptional and phenotypic responses of pCF10-containing cells to pheromone induction. *J. Bacteriol.* **187**: 1044-1054.

Huycke, M.M. (2002) Physiology of enterococci, in <u>The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance</u>. Edited by M.S. Gilmore. ASM Press: Washington, D.C.

Jensen, L.B., Garcia-Migura, L., Valenzuela, A.J.S., Løhr, M., Hasman, H., Aarestrup, F.M. (2010) A classification system for plasmids from enterococci and other gram-positive bacteria. *J. Microbiol. Meth.* **80**: 25-43.

Jett, B.D., Huycke, M.M. and Gilmore, M.S. (1994) Virulence of enterococci. *Clin. Microbiol. Rev.* **7**: 462–478.

Kak, V. and Chow, J.W. (2002) Acquired antibiotic resistances in enterococci, in <u>The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance</u>. Edited by M.S. Gilmore. ASM Press: Washington, D.C.

Kao, S.M., Olmsted, S.B., Viksnins, A.S., Gallo, J.C., and Dunny, G.M. (1991) Molecular and genetic analysis of a region of plasmid pCF10 containing positive control genes and structural genes encoding surface proteins involved in pheromone-inducible conjugation in *Enterococcus faecalis*. *J Bacteriol.* **173**: 7650-7664.

Kurenbach, B., Bohn, C., Prabhu, J., Abudukerim, M., Szewzyk, U., and Grohmann, E. (2003) Intergeneric transfer of the *Enterococcus faecalis* plasmid pIP501 to *Escherichia coli* and *Streptomyces lividans* and sequence analysis of its *tra* region. *Plasmid.* **50**: 86-93.

Kwong, S.M., Skurray, R.A., and Firth, N. (2004) *Staphylococcus aureus* multiresistance plasmid pSK41: analysis of the replication region, initiator protein binding and antisense RNA regulation. *Mol. Microbiol.* **51**: 497-509.

Le Chatelier, E., Janniere, L., Ehrlich, S.D., and Canceill, D. (2001) The RepE initiator is a double-stranded and single-stranded DNA-binding protein that forms an atypical open complex at the onset of replication of plasmid pAMβ1 from gram-positive bacteria. *J. Biol. Chem.* **276**: 10234-10246.

Lebreton, F., van Schaik, W., McGuire, A.M., Godfrey, P., Griggs, A., Mazumdar, V., Corander, J., Cheng, L., Saif, S., Young, S., Zeng, Q., Wortman, J., Birren, B., Willems, R.J., Earl, A.M., Gilmore, M.S.. (2013) Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *MBio.* **4**: e00534-13.

Lebreton, F., van Schaik, W., McGuire, A.M., Godfrey, P., Griggs, A., Mazumdar, V., Corander, J., Cheng, L., Saif, S., Young, S., Zeng, Q., Wortman, J., Birren, B., Willems R.J.L., Earl, A.M., Gilmore, M.S. (2013) Emergence of Epidemic Multidrug-Resistant *Enterococcus faecium* from Animal and Commensal Strains *MBio* **4**: 1-10.

Leclercq, R., Derlot, E., Duval, J., and Courvalin, P. (1988) Plasmid-mediated resistance to vancomycin and teicoplanin in *Enterococcus faecium*. *N Engl J Med.* **319**: 157-161.

Lins, R.X., Andrade, A.O., Junior, R.H., Wilson, M.J., Lewis, M.A.O., Williams, D.W., and Fidel, R.A.S. (2013) Antimicrobial resistance and virulence traits of *Enterococcus faecalis* from primary endodontic infections. *J. Dent.* **41**: 779-786.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S.H. (2011) CDD: a conserved domain database for the functional annotation of proteins. *Nucleic Acids Res.* **39**: D225-D229.

Massie-Schuh, E. (2103). The processing of replication initiation protein PrgW in *Enterococcus faecalis* is necessary for activity and stable maintenance of pCF10. Masters Thesis. Temple University, USA.

McKessar, S.J., Berry, A.M., Bell, J.M., Turnidge, J.D., and Paton, J.C. (2000) Genetic characterization of *vanG*, a novel vancomycin resistance locus of *Enterococcus faecalis*. *Antimicrob Agents Chemother.* **44**: 3224-3228.

Meloni, M., (2005). The replication of the pheromone-dependent plasmid pCF10 of *Enterococcus faecalis* and the function of its replication protein PrgW. Ph.D. Thesis. Temple University, USA.

Meloni, M., Mazeffa, V.J., Buttaro, B.A. The precursor of the peptide pheromone cCF10 effects maintenance of the conjugative plasmid pCF10. *J. Bacteriol.* Submitted.

Moellering, R.C., Korzeniowski, O.M., Sande, M.A., and Wennersten C.B. (1979) Species-specific resistance to antimocrobial synergism in *Streptococcus faecium* and *Streptococcus faecalis*. *J Infect Dis.* **140**: 203-208.

Murray, B.E., Weinstock, G.M. (1999) Enterococci: new aspects of an old organism. *Proc. Assoc. Am. Physicians* **111**: 328–334.

Palmer, K.L., Kos, V.N., and Gilmore, M.S. (2010) Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Curr. Opin. Microbiol.* **13**: 632-639.

Perichon, B., Reynolds, P., and Courvalin, P. (1997) VanD-type glycopeptide-resistant *Enterococcus faecium* BM4339. *Antimicrob Agents Chemother.* **41**: 2016-2018.

Perkins, J.B. and Youngman, P.J. (1984) A physical and functional analysis of Tn917, a *Streptococcus* transposon in the Tn3 family that functions in *Bacillus*. *Plasmid.* **12**: 119-138.

Pinkston, K.L., Gao,P., Diaz-Garcia, D., Sillanpää, J., Nallapareddy, S.R., Murray, B.E., and Harvey, B.R. (2011) The Fsr quorum-sensing system of *Enterococcus faecalis* modulates surface display of the collagen-binding MSCRAMM ace through regulation of *gelE*. *J Bacteriol.* **193**: 4317-4325.

Portillo, A., Ruiz-Larrea, F., Zarazaga, M., Alonso, A., Martinez, J.L., and Torres, C. (2000) Macrolide resistance genes in *Enterococcus* spp. *Antimicrob Agents Chemother.* **44**: 967-971.

Quintiliani, R., Evers, S., and Courvalin, P. (1993) The *vanB* gene confers various levels of self-transferable resistance to vancomycin in enterococci. *J Infect Dis.* **167**: 1220-1223.

Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., and Deber, C.M. (2009) Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc Nat Acad Sci.* **106**: 1760-1765.

Reynolds, P.E. (1989) Structure, biochemistry and mechanism of action of glycopeptide antibiotics. *Eur J Clin Microbiol Infect Dis.* **8**: 943-950.

Rozkov, A. (2001) Control of Proteolysis of Recombinant Proteins in *Escherichia coli*. Department of Biotechnology, Royal Institute of Technology, Stockholm, Sweden. ISBN 91-7283-160-X

Ruhfel, R.E., Manias, D.A., Dunny, G.M. (1993) Cloning and characterization of a region of the *Enterococcus faecalis* conjugative plasmid, pCF10, encoding a sex-pheromone binding function. *J. Bacteriol*. **175**: 5253-5259.

Schlievert, P.M., Gahr, P.J., Assimacopoulos, A.P., Dinges, M.M., Stoehr, J.A., Harmala, J.W., Hirt, H., and Dunny, G.M. (1998) Aggregation and binding substances enhance pathogenicity in rabbit models of *Enterococcus faecalis* endocarditis. *Infect. Immun.* **66**: 218-223.

Schumacher, M.A., Tonthat, N.K., Kwong, S.M., babu Chinnam, N., Liu, M.A., Skurray, R.A., and Firth, N. (2014) Mechanism of staphylococcal multiresistance plasmid replication origin assembly by the RepA protein. *Proc. Nat. Acad. Sci.* **111**: 9121-9126.

Shi, K., Brown, C.K., Gu, Z.Y., Kozlowicz, B.K., Dunny, G.M., Ohlendorf, D.H., Earhart, C.A. (2005) Structure of peptide sex pheromone receptor PrgX and PrgX/Pheromone complexes and regulation of conjugation in *Enterococcus faecalis*. *Proc. Nat. Acad. Sci*. **102**: 18596-18601.

Showsh, S.A., and Andrews, R.E. (1996) Functional comparison of conjugative transposons Tn*916* and Tn*925*. *Plasmid* **35**: 164-173.

Su, Y.A., He, P., and Clewell, D.B. (1992) Characterization of the *tet(M)* determinant of Tn916: evidence for regulation by transcription attenuation. *Antimicrob Agents Chemother.* **36**: 769-778.

Sussman, A.J., and Gilvarg, C. (1969) Protein Turnover in Amino Acid-starved Strains of *Escherichia coli* K-12 Differing in Their Ribonucleic Acid Control. *J. Biol. Chem.* **244**: 6304-6308.

Tannock, G.W. and Cook, G. (2002) Enterococci as members of the intestinal microflora of humans, in The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance. Edited by M.S. Gilmore. ASM Press: Washington, D.C.

Teng, L.J., Hsueh, P.R., Wang, Y.H., Lin, H.M., Luh, K.T., and Ho, S.W. (2001) Determination of *Enterococcus faecalis groESL* Full-Length Sequence and Application for Species Identification. *J. Clin. Microbiol.* **39**: 3326-3331.

Thompson, J.K., and Collins, M.A. (2003) Completed sequence of plasmid pIP501 and origin of spontaneous deletion derivatives. *Plasmid.* **50**: 28-35.

Toledo-Arana, A., Valle, J., Solano, C., Arrizubieta, M.J., Cucarella, C., Lamata, M., Amorena, B., Leiva, J., Penades, J.R., and Lasa, I. (2001) The enterococcal surface protein, Esp, is involved in *Enterococcus faecalis* biofilm formation. *Appl. Environ. Microbiol.* **67**: 4538-4545.

Utter, B.D. (2012). Pheromone-interacting replication protein controls Enterococcal conjugative plasmid host range and stability through disulfide bonds. Ph.D. Thesis. Temple University, USA.

Utter, B.D., Buttaro, B.A. Pheromone-interacting replication protein controls Enterococcal conjugative plasmid pCF10 host-range and stability through disulfide bonds. *Proc. Nat. Acad. Sci.* Submitted.

Weaver, K.E. (2006) *Enterococcal genetics*, in Gram-Positive Pathogens, Second Edition. Edited by Fischetti, V.A., Novick, R.P., Ferretti, J.J., Portnoy, D.A., and Rood, J.I. ASM Press: Washington, D.C.

Weaver, K.E., Kwong, S.M., Firth, N., Francia, M.V. (2009) The RepA_N replicons of Gram-positive bacteria: a family of broadly distributed but narrow host range plasmids. *Plasmid* **61**: 94-109.

Weaver, K.E., Rice, L.B., and Churchward, G. (2002) Plasmids and transposons, in The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance. Edited by M.S. Gilmore. ASM Press: Washington, D.C.

Wirth, R., and Muscholl, A. (1996) Comparative analysis of 18 sex pheromone plasmids from *Enterococcus faecalis*: detection of a new insertion element on pPD1 and implications for the evolution of this plasmid family. *Mol Gen Genet.* **252**: 967-971.

APPENDIX

FORMULAS, BUFFERS, AND REAGENTS

Growth Media

Lysogeny Broth (LB) (Bertani, 1951, 2004)

10 g Bacto™ Tryptone (Catalog # 211705)

5 g Bacto™ Yeast Extract (Catalog # 212750)

10 g Fisher® Sodium Chloride (Catalog # S271-3)

q.s. to 1 L with dH_2O

Autoclave for 35 minutes at 121°C.

LB Agar

10 g Bacto™ Tryptone (Catalog # 211705)

5 g Bacto™ Yeast Extract (Catalog # 212750)

10 g Fisher® Sodium Chloride (Catalog # S271-3)

15 g (1.5%) Difco™ Agar, Granulated (Catalog # 214530)

q.s. to 1 L with dH_2O.

Autoclave for 35 minutes at 121°C.

Terrific Broth (Tartof and Hobbs, 1987)

12 g Bacto™ Tryptone (Catalog # 211705)

24 g Bacto™ Yeast Extract (Catalog # 212750)

4 mL Fisher® Glycerol (G33-1)

q.s. to 900 mL with dH_2O.

Autoclave for 35 minutes at 121°C.

Mix Potassium phosphate solution (below) when both solutions are below 60°C.

Potassium phosphate solution for Terrific Broth

0.17 M Sigma® KH_2PO_4 (Catalog # P-5379)

0.7 M Fisher® K_2HPO_4 (Catalog # P288-500)

q.s. to 100 mL with dH_2O.

Autoclave for 20 minutes at 121°C.

Todd Hewitt (TH) Broth

30 g Bacto™ TH powder (Catalog # 249240)

q.s. to 1 L with dH_2O

Autoclave for 35 minutes at 121°C.

TH Agar

30 g Bacto™ TH powder (Catalog # 249240)

15 g (1.5%) Difco™ Agar, Granulated (Catalog # 214530)

q.s. to 1 L with dH_2O

Autoclave for 35 minutes at 121°C.

DNA Manipulation Solutions

1X TAE Buffer

242.2 g Fisher® Tris-Base (Cat #BP152-1)

57.1 mL Fisher® Glacial Acetic Acid (Catalog # A38-212)

100 mL of 0.5 M Fisher® EDTA (Catalog # BP120-1), pH 8.0

q.s. to 1 L with dH_2O.

The above formula generates 50X TAE. Dilute 1:50 with dH_2O to obtain the 1X working solution.

6X Loading Dye

300 µL Fisher® Glycerol (Catalog #G33-1)

25 µL FisherBiotech® Bromophenol Blue (Catalog # BP115-25)

q.s. to 1 mL with dH_2O.

Lysis and SDS-PAGE Solutions

Protein Lysis Buffer

1.95 mL (0.5 M) Sigma® KH_2PO_4 (Catalog # P-5379)

3.05 mL (0.5 M) Fisher® K_2HPO_4 (Catalog # P288-500)

50 µl (1 M) Fisher® Manganese Chloride (Catalog # M87-500)

25 mg (500 µg/ml) Calbiochem® DNase (Catalog # 260913)

12.5 mg (250 µg/ml) Calbiochem® RNaseA (Catalog # 556746)

45 mL Baxter® Sterile H_2O (Catalog # 2F7114)

Stacking Buffer

28 g Fisher® Tris-HCl (Catalog # BP153-1)

2.2 g Fisher® Tris-Base (Cat #BP152-1)

4 g Fisher® SDS (Catalog # BP166-500)

q.s. to 1 L with dH_2O.

Adjust to final pH of 6.8 with HCl.

Resolving Buffer

60 g Fisher® Tris-HCl (Catalog # BP153-1)

30 g Fisher® Tris-Base (Cat #BP152-1)

4 g Fisher® SDS (Catalog # BP166-500)

q.s. to 1 L with dH_2O.

Adjust to pH 8.8 with HCl.

10X SDS-PAGE Tank Buffer

30.3 g (25 mM) Fisher® Tris-Base (Cat #BP152-1)

144.1 g (192 mM) Fisher BioReagents® Glycine (Catalog # BP381-1)

10 g (0.1%) Fisher® SDS (Catalog # BP166-500)

q.s. to 1 L with dH_2O.

Adjust to a final pH of 8.3 with HCl.

5X Loading Dye

757 mg Fisher® Tris-HCl (Catalog # BP153-1)

10 mL 10% Fisher® SDS (Catalog # BP166-500)

10 mL Fisher® Glycerol (Catalog #G33-1)

0.65 mL FisherBiotech® Bromophenol Blue (Catalog # BP115-25)

Adjust to a final pH of 6.8 with NaOH.

Coomassie Blue Stain

500 mL (50%) Fisher® Methanol (Catalog # A411-4)

100 mL (10%) Fisher® Glacial Acetic Acid (Catalog # A38-212)

600 mL dH_2O

0.05% Bio-Rad Coomassie® Brilliant Blue R-250 (Catalog # 161-0400)

Coomassie De-staining Solution

300 mL (30%) Fisher® Methanol (Catalog # A411-4)

100 mL (10%) Fisher® Glacial Acetic Acid (Catalog # A38-212)

600 mL dH_2O

Western Blot Solutions

Towbin Buffer

1 g Fisher® Tris-HCl (Catalog # BP153-1)

2 g Fisher® Tris-Base (Cat #BP152-1)

14.4 g Fisher BioReagents® Glycine (Catalog # BP381-1)

200 mL Fisher® Methanol (Catalog # A411-4)

q.s. to 1 L with dH_2O

Adjust to a final pH of 8.3 with NaOH.

1X TBS Wash buffer (1 X TBS with 0.1% Tween)

100 mL 10X TBS

10X TBS

12.1 g (20 mM) Fisher® Tris-Base (Cat #BP152-1)

40 g Fisher® Sodium Chloride (Catalog # S271-3)

19 mL (1N) Fisher® Hydrochloric Acid (Catalog # A144-212)

q.s. to 500 mL with dH_2O

Adjust to a final pH of 7.6 using HCl.

q.s. to 1 L with dH_2O

Mix in 1 mL Fisher® Tween 20 (Catalog #BP337-500) while stirring.

Blocking Buffer

100 mL 1X TBS Wash Buffer (see above)

5.0 g Nestle® Carnation® Instant Nonfat Dry Milk (Catalog # 01-4280-00)