

**BIOCHEMICAL ANALYSIS OF MODULATION OF
SEX PHEROMONE PRODUCTION BY PRGY, A
STRUCTURAL HOMOLOG OF A
METALLOPROTEASE (TIKI) THAT MODULATES
WNT SIGNALING IN EUKARYOTES**

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE
UNIVERSITY OF MINNESOTA

BY

THINH CANH LE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DR. GARY M. DUNNY, ADVISOR

MAY 2017

© Think Canh Le 2017

TABLE OF CONTENTS

Table of Contents.....	i
List of Tables.....	iii
List of Figures.....	iv
1. Introduction.....	1
2. Materials and Methods.....	17
Bacterial Strains.....	17
Culture Conditions.....	17
Plasmids.....	18
DNA Manipulation and Cloning.....	18
Determination of Pheromone Activity.....	19
Proteins Expression and Purification.....	19
Production of PrgY Polyclonal Antibody.....	20
Protein Binding Assays.....	21
Pheromone Activity in the Presence or Absence of PrgY.....	22
SDS-PAGE and Western Blot Analyses.....	22
Crosslinking Assays.....	23
In-Gel Trypsin Digest for Mass Spectrometry Analysis.....	23

Shotgun MS Analysis.....	24
Computer Analysis and Structural Modeling.....	25
Mass Spectrometry Analysis of Culture Supernatants.....	25
3. Results.....	27
I. Purification and Binding between PrgY and cCF10.....	27
Expression and purification of PrgY and variants.....	27
N-Y specifically reduces the amount of pheromone flow-thru in pull-down assays, demonstrating the direct binding between cCF10 and N-Y <i>in vitro</i>	32
The amino-terminal fragment of PrgY specifically interacts with cCF10 in solution, but does not directly degrade or modify cCF10.....	42
II. Examination of the Mechanistic Models for PrgY Function.....	51
Searching for host protease that interacts with PrgY.....	51
Comparative modeling of PrgY: a Structural Homolog of a Metalloprotease (TIKI) that modulates Wnt Signaling in Eukaryotes.....	60
PrgY cleaves cCF10 after the 2 nd Leucine.....	70
4. Discussion.....	75
5. References.....	84

LIST OF TABLES

1. Proteolytic processing of the pheromone precursors.....	6
2. Affinities of peptides for N-Y.....	40
3. Purified amino-terminal fragment of PrgY specifically interacts with cCF10 in solution, but does not directly degrade or modify cCF10.....	44
4. Mass spectrometry analysis of synthetic peptides.....	71
5. Mass spectrometry analysis of culture supernatants.....	73

LIST OF FIGURES

1. Transfer of <i>Enterococcus faecalis</i> plasmid pCF10.....	3
2. Proteolytic processing of cCF10.....	7
3. Control of self-induction of conjugation by the plasmid-encoded protein PrgY.....	10
4. Structure and characterization of PrgY.....	11
5. Model #1 of endogenous pheromone control by the pCF10 encoded protein PrgY.....	14
6. Purified N-Y and variants.....	30
7. Purified full length PrgY.....	31
8. Western blot of PrgY polyclonal antibody.....	33
9. Pheromone flow-thru in pull-down assay.....	35
10. Surface plasmon resonance kinetic analysis of cCF10 binding to N-Y...38	
11. N-Y protects the soluble pheromone from host protease digestion in the culture medium by binding to the free peptide.....	47
12. Model #2 of endogenous pheromone control by the pCF10 encoded protein PrgY.....	48
13. Pull-down assay of N-Y and OG1RF membrane fractions.....	53
14. Shotgun MS analysis of unique gel bands of N-Y pull-down assay.....	55
15. Pull-down assay of full-length PrgY and OG1RF membrane fractions...58	
16. Shotgun MS analysis of unique gel bands of PrgY pull-down assay.....	59
17. Crosslinking assay with <i>E. faecalis</i> OR1RF cell lysate over expressing PrgY.....	61

18. Mechanism of action by TIKI and the proposed mechanism of action by PrgY.....	63
19. Comparative modeling of PrgY.....	64
20. 3D structure model of PrgY.....	66
21. PrgY shares significant homology to the eukaryotic metalloprotease TIKI that regulates Wnt signaling.....	68

INTRODUCTION

Enterococci are opportunistic gram-positive cocci that inhabit the gastrointestinal tract, the oral cavity, and the vagina in humans [1]. They can cause a variety of life-threatening diseases in humans including infection of the urinary tract, bloodstream, endocardium, abdomen, biliary tract, and burn wounds. Currently Enterococci rank among the top three nosocomial bacterial pathogens and up to 90% of enterococcal infections in humans are caused by *Enterococcus faecalis*. *Enterococcus faecalis* is known to be the most frequently detected species in root canals with failed endodontic treatment as well as in saliva and subgingival samples of periodontitis patients [2, 3]. In addition to its pathogenic potential, *E. faecalis* is also intrinsically resistant to many commonly used antimicrobial agents including aminoglycosides, aztreonam, cephalosporins, clindamycin, the semi-synthetic penicillins nafcillin and oxacillin, and trimethoprim-sulfamethoxazole [4]; it also frequently exhibits resistance mediated by genes acquired via horizontal gene transfer, mainly conjugation [5-7]. This poses a great threat in hospital settings, particularly to patients with long-term hospitalization requiring use of a variety of antibiotics.

Enterococcus faecalis contains mobile genetic elements that can rapidly spread antibiotic resistance and virulence genes through its population by conjugation [8]. Several virulence and antibiotic resistance traits can accumulate on a single plasmid, which can transfer and disseminate through populations. In many cases, plasmid transfer is mediated by pheromone inducible conjugation

systems. Unlike canonical quorum sensing in which one cell type can both produce and detect a signal, pheromone-inducible conjugation involves two distinct cell types: donor cells which contain plasmid-encoded signal detectors and recipient cells that function as signal producers [9]. These pheromone inducible plasmids have evolved a highly specific and sensitive response to pheromone to allow efficient communication between cells and are tightly regulated in order to minimize fitness costs. The transfer functions of pheromone plasmids are induced by small, 7-8 amino acid peptides, encoded by the chromosome of most enterococcal strains. Each peptide is highly specific for a cognate plasmid. The best-studied pheromone inducible plasmids are pCF10, pPD1, pAD1, and pAM373, with corresponding sex pheromones cCF10, cPD1, cAD1, and cAM373, respectively [10, 11]. These plasmids are often found in clinical isolates, and up to three different pheromone responding plasmids have been identified in a single isolate [12].

Extensive research has been done to understand the mechanism of pheromone-inducible conjugation. An overview of the pheromone-induced conjugative plasmid transfer of pCF10 will be presented here and the details will be described later; the other *E. faecalis* pheromone responsive plasmids encode similar regulatory gene orthologs, which are described in [7, 8]. The chromosomal *ccfA* gene encodes a predicted secreted lipoprotein whose signal peptide is processed to the heptapeptide pheromone cCF10 by a membrane protease Eep (Figure 1)[13]. Mature pheromone is then released into the culture medium where it can be imported by a donor cell through the pCF10-encoded

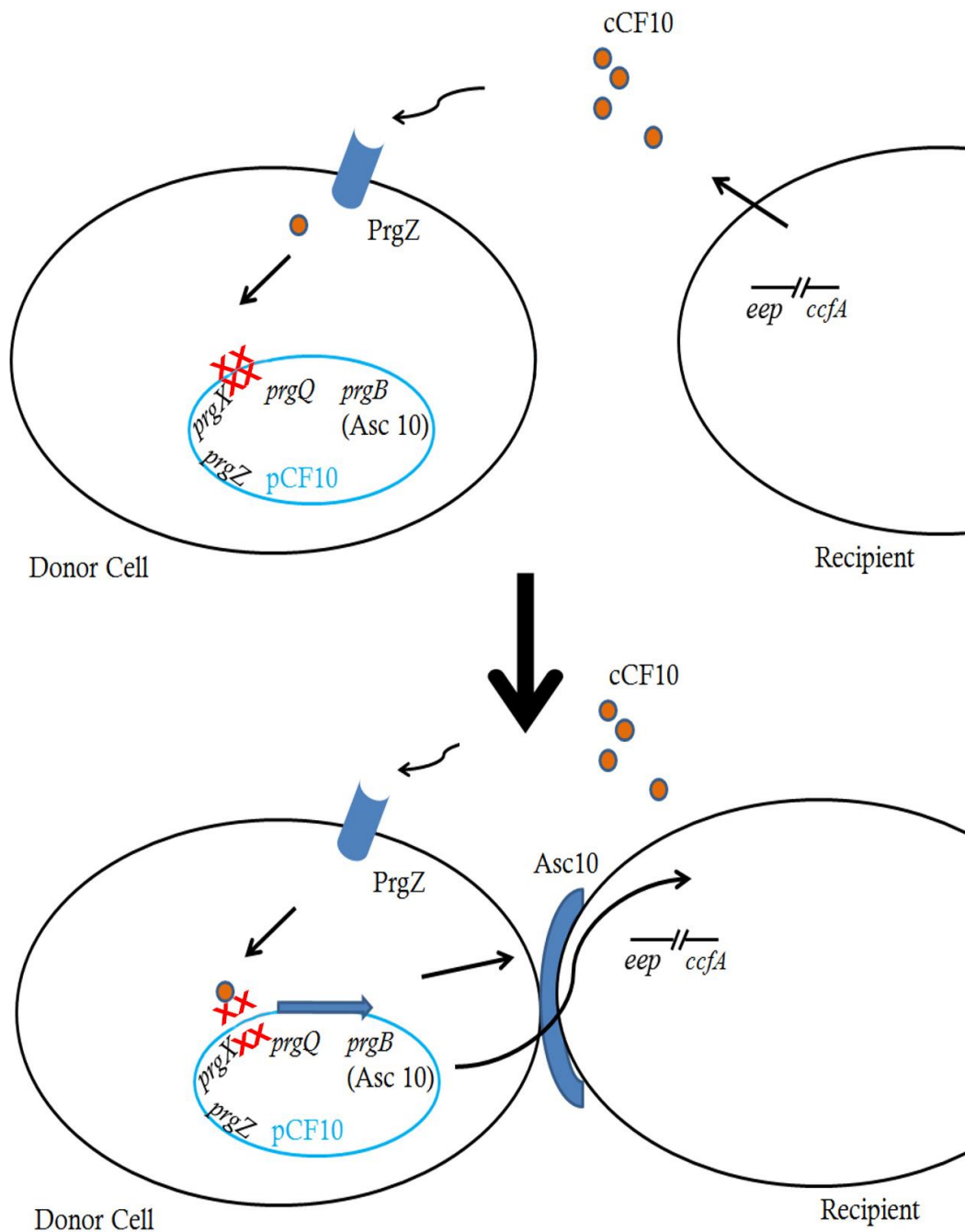


Figure 1. Transfer of *Enterococcus faecalis* plasmid pCF10. The signal peptide portion of the chromosomally encoded *ccfA* gene product contains a heptapeptide pheromone cCF10 which is processed from the full length polypeptide by Signal Peptidase II and a membrane protease Eep (upper right). Mature pheromone is released into culture medium where it gets imported by a donor

cell through a pheromone binding protein PrgZ (upper left). Once cCF10 gets in the donor cell, it then binds to and impedes the function of the negative regulator PrgX (lower left). This allows transcription of aggregation substance Asc10 which allows the two cells to stick together, enabling subsequent formation of a mating channel, and conjugative transfer of pCF10 (lower diagram).

pheromone binding protein PrgZ. PrgZ acts in concert with the chromosomally encoded oligopeptide permease system (Opp) to import the pheromone into the donor cell. Once cCF10 gets in the donor cell, it then binds to and alters the structure of the negative regulator PrgX [13]. This allows transcription of aggregation substance Asc10 and other conjugation proteins, which allows the two cells to stick together, form a mating channel, and conjugation occurs [14].

Each of the pheromones appear to be derived from the C-terminal segment of the signal sequences of predicted lipoprotein genes encoded in the chromosome of all *E. faecalis* [15]. Bioinformatics analysis of *E. faecalis* V583 genome sequence reveals greater than 50 putative lipoproteins with predicted signal sequences that may be associated with the *E. faecalis* membrane [16]. The best-studied pheromone precursors and processed peptides are listed in Table 1. The mature pheromone peptides for each plasmid system are very similar in length and are very hydrophobic, despite the specificity of each plasmid's response to its cognate pheromone. Proteolytic processing of the pheromone precursors must occur for the release of the mature pheromone into the culture supernatant (Figure 2). Interestingly, all of the precursors have a conserved cysteine residue downstream of the sequence where signal peptidase II cleaves, releasing the signal peptide from the lipoprotein [15]. In the case of cCF10, the LVTLVFV pheromone sequence is three residues upstream of this processing site and therefore requires further processing by a host exoprotease. A putative membrane protease Eep has been shown to mediate endoproteolytic cleavage at the upstream processing site, liberate the pheromone peptide from

Plasmid	Pheromone/precursor	Precursor gene name and amino acid length
pCF10	...MAG <u>L</u> VTLVFVLSAC G ...	ccfA, 275
pPD1	...GLLFLVMFLS G CV...	cpd, 234
pAD1	...AIALFSLVLAG C G...	cad, 309
pAM373	...LLGAIFILAS C G...	cam, 166

Table 1. Proteolytic processing of the pheromone precursors. Underlined letters represent amino acids in the mature peptides. Bold C indicates cysteine residues upstream of predicted signal peptidase II cleavage site.

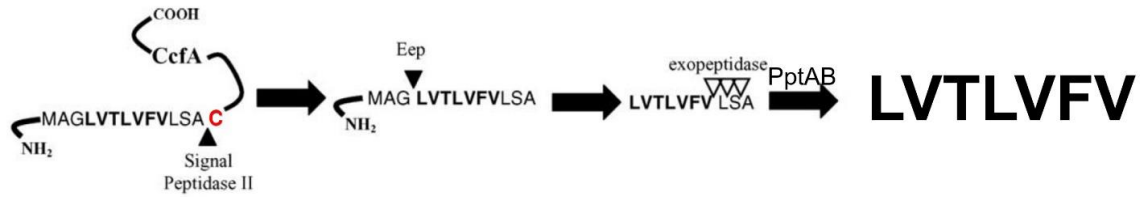


Figure 2. Proteolytic processing of cCF10. Lipoprotein CcfA produced in the cytoplasm is exported out through the secretory apparatus. A signal peptidase 2 cleaves the N-terminal overhang from the conserved cysteine, then Eep, a membrane-embedded zinc metalloprotease further cleaves the N-terminal overhang to produce pheromones. Exopeptidase cleaves several residues at the C-terminus to produce the mature pheromone. PptAB, the peptide transporter AB, actively transports mature pheromones outside the cell.

its cognate lipoprotein signal sequence [9]. Eep is a predicted metalloprotease that has been shown to be involved in production of cCF10, cPD1, and cAD1. A peptide pheromone transporter (PptAB), an ABC transporter, has been shown to actively transport the mature pheromones out of the cell [17].

The most unique phenotype of pCF10-containing cells when exposed to exogenous cCF10 is visible aggregation of cell culture that results from upregulation of the expression of aggregation substance Asc10 from the pCF10-encoded *prgB* gene [18]. Aggregation substance mediates close contact between donor and recipient cells, facilitates mating pair formation in liquid cultures and allows efficient transfer of the plasmid. This clumping response serves as a useful phenotype in quantifying pheromone activity using microtiter dilution assay (clumping assay) [19]. Asc10 has been shown to increase internalization and intracellular survival in polymorphonuclear leukocytes, adhere to several extracellular matrix proteins and has been shown to increase virulence of *E. faecalis* in several rabbit endocarditis models [20]. Studies by Saringen *et al* have demonstrated that aggregation substance increases adherence and internalization of *E. faecalis* through different intestinal epithelial cells in vitro [21], suggesting that these aggregation substances may be involved in the translocation of *E. faecalis* through the intestinal epithelia, resulting in systemic infection [22].

Once pCF10 is transferred into a new host strain, this recipient cell becomes a potential donor and the related pheromone is no longer detectable in culture supernatants. Since *ccfA* is encoded in the chromosome of all *E. faecalis*

strains, pCF10 containing cells can potentially continue to secrete pheromone after they have acquired the plasmid, thus there must be a mechanism to shut down induction by endogenously produced pheromone in these cells. Two different pCF10 encoded gene products have been identified through genetic studies to prevent donor cells from self-induction of transfer genes by their endogenously produced pheromone [8]. One gene, *prgQ*, encodes a heptapeptide iCF10 (AITLIFI) that gets secreted into the extracellular medium to neutralize the cCF10 secreted from the same cell and competes with the pheromone for binding to the surface protein PrgZ [23], and to PrgX. However, this is not sufficient to completely block the spontaneous self-induction in donor cells. In addition to iCF10, a membrane protein PrgY has been identified that appears to sequester or degrade the endogenous pheromone in pCF10 containing cells (Figure 3) [9]. The balance of cCF10 and iCF10 is carefully regulated such that the level of iCF10 produced is just sufficient to neutralize endogenous cCF10 in donor cells. The balance between cCF10 and iCF10 can overcome by addition of exogenous cCF10 at less than 5 molecules/responder cell [24].

Previous genetic experiments and sequence analysis done in our lab indicated that the likely functional domains of the protein correspond to the N-terminal region of PrgY, which is predicted by sequence analysis and functional studies to be on the outside of the cell membrane [9]. The N-terminal region is responsible for recognition of the pheromone, while the C-terminus anchors PrgY to the membrane via four predicted transmembrane domains [8, 9, 14] (Figure 4).

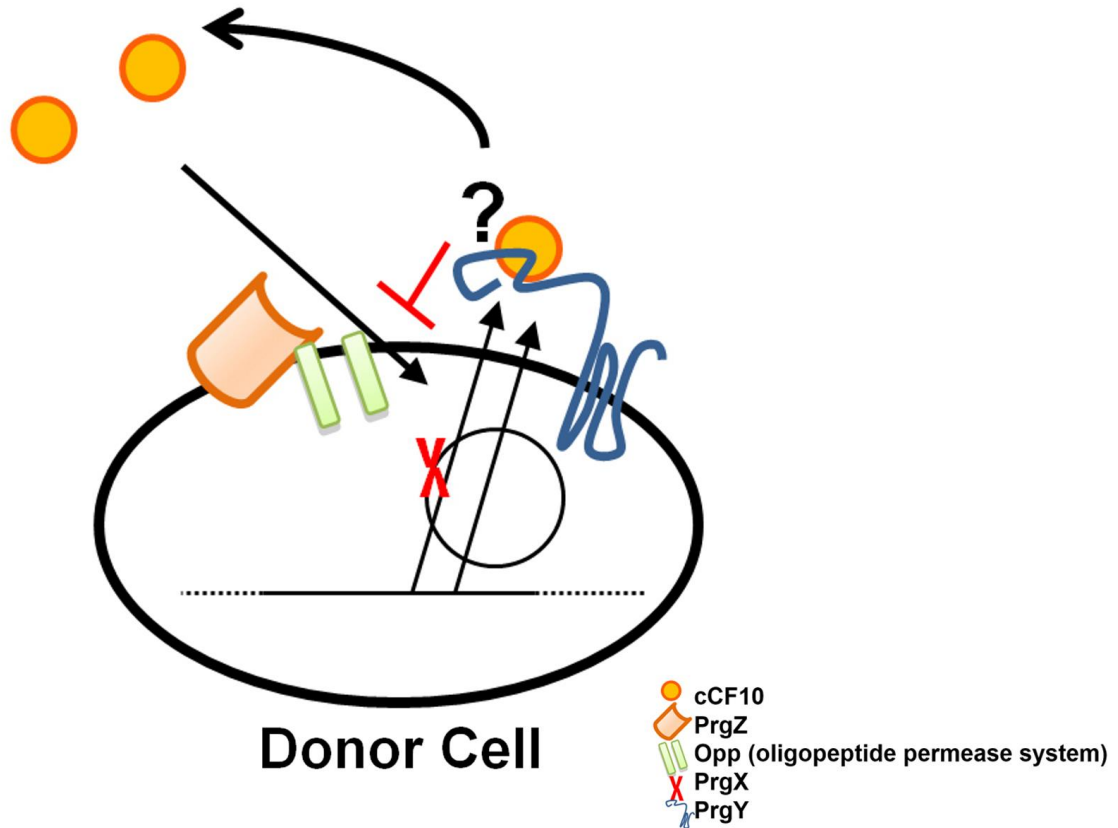


Figure 3. Control of self-induction of conjugation by the plasmid-encoded protein PrgY. Because recipient cells can potentially continue to secrete pheromone after acquiring the plasmid, thus there must be a mechanism to shut down self-induction of donor cells by endogenously produced pheromone. The plasmid encodes a membrane protein called PrgY that apparently plays a role in preventing donor cell from self-induction. PrgY only affects endogenously produced pheromone and not pheromone added exogenously or from recipient cells.

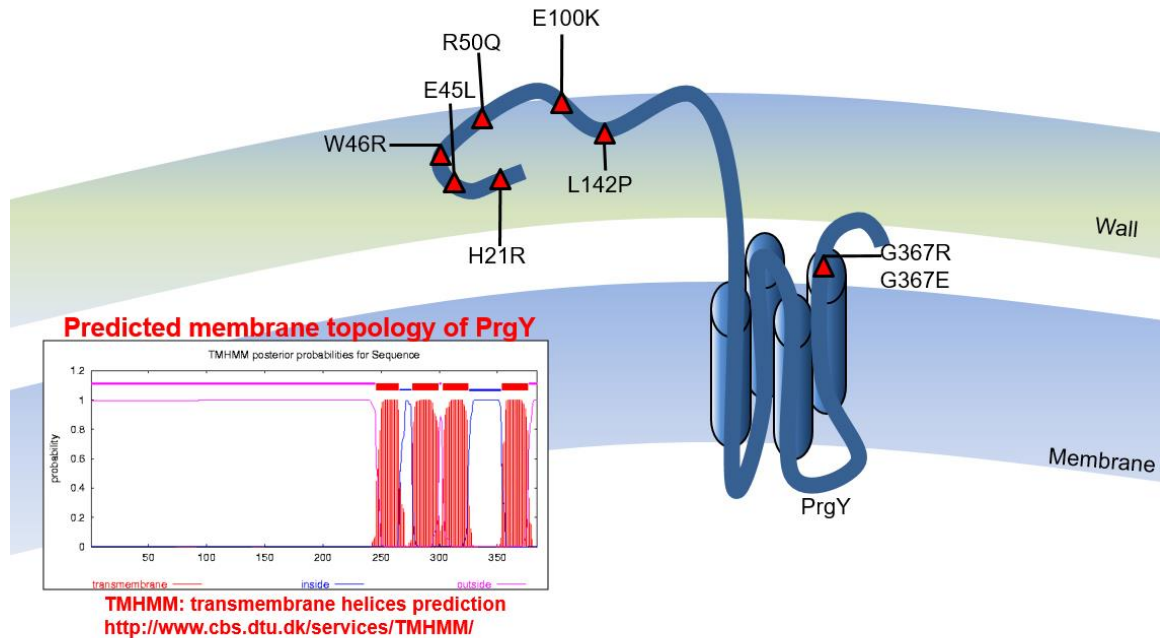


Figure 4. Structure and characterization of PrgY. Previous experiments and sequence analysis done in our lab indicated that the likely topology of the protein includes an extracellular N-terminal domain and a C-terminal membrane anchoring domain. The functional N-terminal domain was identified through random mutagenesis and later confirmed with site-directed mutagenesis. Most of the *prgY* point mutations that failed to reduce endogenous pheromone activity (red triangles) are located at the N-terminus of PrgY. These mutations map to conserved residues in many species including gram-negative and gram-positive bacteria, archaea, plants, and mammals. Furthermore, our lab has shown that residues 123 to 241 of PrgY are required for specific regulation of cCF10.

The N-terminal functional domain was identified through random mutagenesis in which *prgY* mutations were screened and those that failed to prevent self-induction were further confirmed with site-directed mutagenesis. These mutations in *prgY* that abolished control of endogenous pheromone in donor strains mapped to conserved residues in many species including gram-negative and gram-positive bacteria, archaea, plants, and mammals. Two of the random mutations (E45L and R50Q) mapped to residues that are conserved throughout the entire PrgY family [14]. Furthermore, our lab has generated chimeric pPD1 TraB/PrgY proteins to identify segments of the PrgY N-terminal domain that confers specificity for cCF10. Amino residues 125 to 241 of PrgY are required for specific regulation of cCF10 *in vivo* [9]. PrgY recognizes the mature cCF10 peptide sequence, after processing by Eep, rather than the N-terminal sequence residues. However, Eep is not required for PrgY to function. Additionally, it was previously identified that the fourth residue (L4) in cCF10 peptide is important for PrgY recognition and that an L4I peptide variant of cCF10 abolished PrgY ability to reduce pheromone activity in donor cells [9]. Previous experiments have shown a significant amount of pheromone lingering in the cell wall [8]. In fact, titrations of active cCF10 in supernatant and cell fraction revealed more than twice as much of pheromone in the cell wall compared to pheromone in supernatant. PrgY was found previously to reduce endogenous pheromone activity in the cell wall strain OG1RF while not affecting supernatant pheromone activity [8]. Based on these findings, we hypothesized that PrgY reduces endogenous pheromone production in donor cells by specifically binding and

likely degrading cCF10 as it is secreted across the cytoplasmic membrane (Figure 5). The lipoprotein CcfA produced in the cytoplasm is exported out through the secretory apparatus (Sec) as the signal peptide gets processed by signal peptidase II in the membrane. Next, the membrane protease Eep recognizes specifically the N-terminal domain of the cleaved signal peptide in the membrane and processes it further to enable the release of mature cCF10 through a peptide pheromone transporter (PptAB), an ABC transporter [17]. Previous genetic experiments done in our lab indicated that Eep functions through the recognition of the peptide sequence N-terminal to the pheromone [9]. Chimeric fusions of the *prgQ* promoter with N-terminal peptide sequence derived from the iCF10 or cAM373 pre-processed sequence and a pheromone at the C-terminal end were generated and assayed for the relative amounts of pheromone produced by each construct. Eep was previously demonstrated to not be involved in production of cAM373 [25]. Pheromone production from cAM373 signal peptide construct was the same for both Eep+ and Eep deficient strains, indicating that pheromone from the construct containing cAM373-derived N-terminal peptide is Eep independent, while the cCF10-derived N-terminal peptide was enhanced by Eep, indicating that the N-terminal peptide is important for recognition by Eep [9]. PrgY anchors in the cell membrane, probably in close proximity to Eep and in a precise orientation, such that it readily captures mature cCF10 immediately upon release from the membrane and before reimport by PrgZ. It is unclear whether this inhibition occurs through degradation, modification, or simply blocking the release of mature cCF10 from the cell surface.

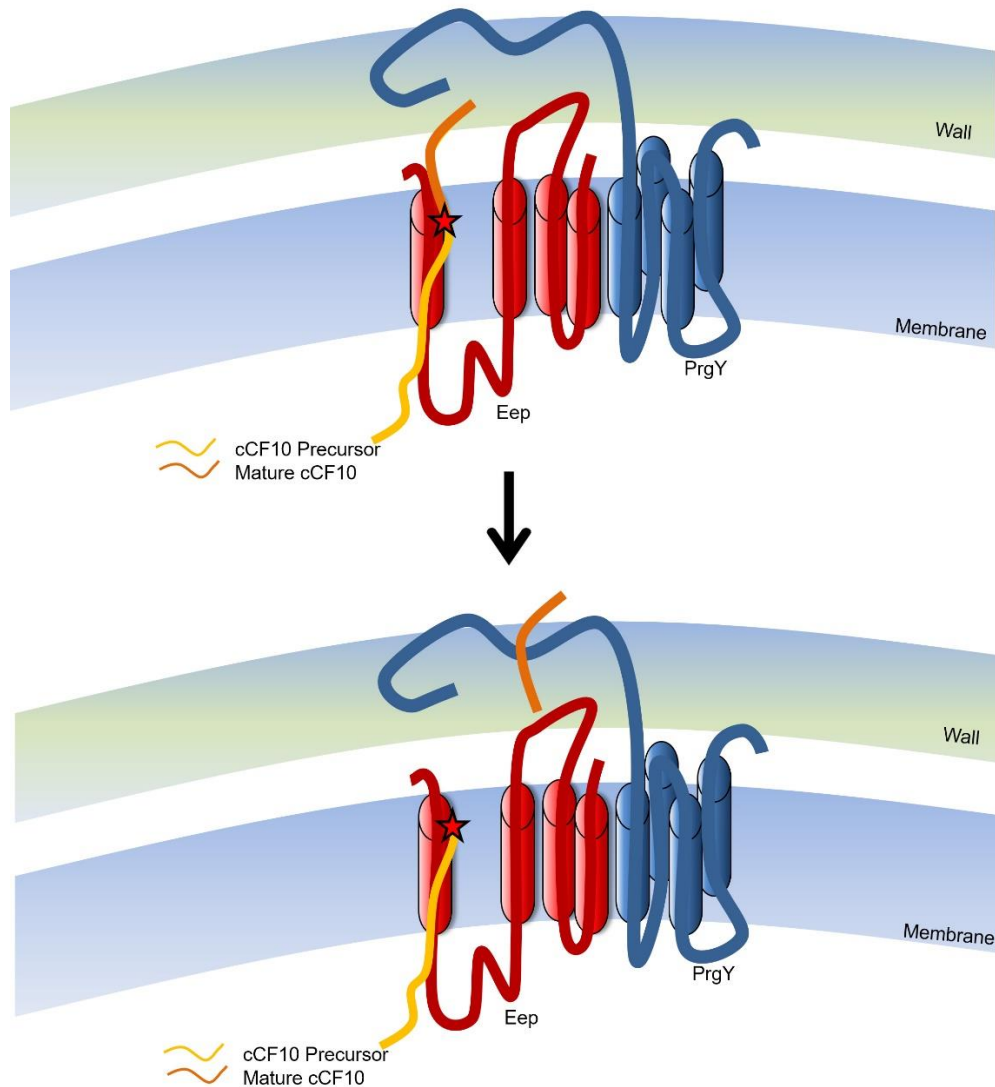


Figure 5. Model #1 of endogenous pheromone control by the pCF10 encoded protein PrgY. The membrane protease Eep recognizes N-terminal domain of signal peptide and processes it to enable release of mature cCF10. PrgY anchors in the membrane, probably in close proximity to Eep, and specially targets cCF10 through amino acids 125-241 in PrgY and reduces cCF10 activity through an N-terminal domain. It is unclear whether this inhibition occurs through degradation, modifying, or simply blocking the mature cCF10 from being released into the extracellular medium.

PrgY-like proteins are found in all biological kingdoms. No functions were previously known for PrgY-like proteins when I began my research project and the mechanism by which PrgY controls endogenous pheromone had not been elucidated. PrgY is a membrane protein and due to the hydrophobicity of membrane proteins, biochemical and structural studies of PrgY have been difficult, and thus very little is known about the biochemical mechanism of PrgY. For this reason, it is crucial to further investigate the roles of PrgY in this pheromone inducible conjugation system because whatever its biochemical mechanism is, it clearly appears to be important to all biological kingdoms.

Our long term goal has been to understand the control of pheromone activity in donor and recipient cells as well as the molecular mechanism of pheromone response in *Enterococcus faecalis*. The first objective of my research was to express the cloned *prgY* gene and mutant derivatives, and develop a method to purify membrane protein PrgY and variants, with adequate yields and stability in solution. My second objective was to characterize the biochemical activity of PrgY using the purified PrgY and variants. To test the working model, affinity chromatography and Surface Plasmon Resonance were used to demonstrate the direct interaction between PrgY and cCF10 and their binding affinities. During the course of these studies, another hypothesis was also developed and tested. We hypothesized that PrgY reduces endogenous pheromone production in donor cells by specifically binding to cCF10 as it is secreted across the cytoplasmic membrane and delivering it to one or more host proteins that degrade or modify the pheromone. To further investigate this,

purified His tag PrgY protein was used as the “bait” to capture and ‘pull down’ a protein-binding partner, the “prey” from OG1RF lysate that binds to PrgY and used mass spectrometry to identify the identity of the prey protein. Crosslinking assays were also performed in an attempt to capture PrgY interacting partner *in vivo*; unfortunately, these experiments were not informative in elucidating the functional mechanism of PrgY. Comparative modeling of PrgY was also carried out with Phyre2 and analyzed with PyMOL. Mass spectrometry was used to identify possible degradation products of cCF10 in culture supernatants by comparing supernatants from a strain expressing PrgY to a strain lacking PrgY. The cumulative results of this research provide important insights into the molecular mechanism of PrgY. These findings give us a better understanding of a new biological function, as well as providing useful information about signaling peptide degradation and modification in enterococci. Furthermore, results from this work could also advance our understanding on the function of each of the PrgY family members found in a diverse range of species.

MATERIALS AND METHODS

Bacterial Strains

Strain JRC 105 (OG1RF Δ gelE-sprE) was previously made by allelic exchange using conjugative delivery and the Δ gelE-sprE allele was confirmed by plating on THB + 3% gelatin and observed the absence of the characteristic halo around isolated colonies denoting gelatinase activity [9]. *E. coli* strain DH5 α was used to clone and purify full length PrgY, the N-terminal portion of PrgY containing residues 1-241 (N-Y), and variants for biochemical studies and polyclonal antibody production.

Culture Conditions

Enterococcus faecalis was grown at 37°C in Todd-Hewitt broth (THB; Difco), in M9-YE glucose medium [26], a semi-defined M9-based medium supplemented with 0.3% yeast extract, 1% casimino acids, 0.1% glucose, 1mM MgSO₄, and 0.1 mM CaCl₂, or in chemically defined medium (Ref). Antibiotics were used where appropriate at the following concentrations: tetracycline (tet) 10 ug/ml, spectinomycin (spect) 1000 ug/ml, rifampicin (rif) 200 ug/ml, erythromycin (erm) 10 ug/ml, kanamycin (kan) 100ul/ml. Nisin (Sigma) was used at 25 ng/ml in broth culture. *Escherichia coli* DH5 α was grown in Luria-Bertani (LB) broth at 37°C, shaking, for cloning purposes with the following antibiotics: erm 200 ug/ml, kan 50 ug/ml, spect 50 ug/ml.

Plasmids

A PrgY expression construct for *E. faecalis* was previously made by fusing the *prgY* gene in-frame with the first 3 codons of *nisA* gene in the nisin inducible promoter of the pMSP3545S expression vector [14]. The PCR amplified *prgY* gene from pCF10 with primer-encoded *NcoI* and *XbaI* restriction sites was cloned into the *NcoI/XbaI* sites of pMSP3545. Furthermore, *prgY* gene was previously PCR amplified to incorporate random mutations, cloned into pMSP3545S and screened for nonfunctional PrgY. The nonfunctional PrgY variants were identified on plate by their dry-colony phenotype that corresponds to clumping in broth culture. Plasmids pMSP3545S containing wildtype *prgY* as well as point mutations R50Q and L142P were isolated, PCR amplified, and digested with *NcoI* and *XhoI*. The amplicons were cloned into pET28b (Novagen, USA), an IPTG-inducible expression vector featuring 6-His Tag at the C-terminus for protein purification. The resulting plasmids were transformed into *E.coli* DH5 α for protein expression and purification.

DNA Manipulation and Cloning

Plasmids were isolated with Qiagen mini kit (Qiagen), digested DNA products were purified with Qiaquick gel extraction kit (Qiagen), and PCR products were purified with Qiaquick PCR purification kit (Qiagen). Plasmid constructs were verified with DNA sequencing by SimpleSeq (Eurofins). Restriction enzymes were purchased from Promega and New England BioLabs. PCR was performed with a Techne TC-412 Thermal Cycler.

Determination of Pheromone Activity

Pheromone induction can be detected by the formation of visible clumps in donor cell cultures exposed to the peptide [26]; this phenotype served as the basis for the Microtiter Assay (Clumping Assay) used to determine pheromone activity. One hundred microliters of either *E. faecalis* OG1RF culture supernatant, eluted fractions from binding assays, or synthetic pheromone was added to the first well of a microtiter plate, and twofold serial dilutions were made across a row of wells. Then 10 μ l of an overnight culture of *E. faecalis* OG1RF(pCF10) was added to each well, and the plate was incubated at 37°C with shaking for 2 hours. Pheromone activity was reported as the titer, which is the reciprocal of the highest twofold dilution showing a positive clumping results.

Proteins Expression and Purification

An overnight culture of Escherichia coli DH5a expressing a desired protein was grown in Luria-Bertani (LB) broth at 37°C with kan 50 μ g/ml, then diluted into 1 L of LB medium with kan 50 μ g/ml, incubated at 37°C shaking for about 3 hours. At exponential phase, culture was induced with 1 mM IPTG final concentration and continue to incubate at 37°C shaking for another 1 hour. Cells were collected by centrifugation and treated with lysozyme solution (0.5 mg/L) in 1X Bind buffer (5 mM imidazole, 500 mM NaCl, 20 mM Tris HCl) with 20% glycerol for 15 minutes at 37°C. Cells were sonicated on ice and centrifuged at 45,000 rpm for 45 minutes at 4 °C. Pellets were collected and resuspended in 1x Bind buffer with 6 M guanidine HCl and incubated on ice for 1 hour, then centrifuged at 38,000 rpm for 40 minutes and the supernatant (membrane

fraction with protein of interest) was collected. Ni-NTA agarose (ThermoFisher Scientific) and a purification column were used to purify His-tagged proteins. Columns were prepared with 1.5 mL of Ni-NTA agarose, charged with 7.5 ml of 1X Charge buffer (50 mM NiSO₄), washed with water, and then saturated the column with 1X Binding buffer with 6M guanidine HCl. Membrane fractions with protein of interest were added to the column and allowed to bind at 4 °C for 1 hour. Columns were then washed with 1X Wash buffer (60 mM imidazole, 500 mM NaCl, 20 mM Tris HCl) with 6 M guanidine HCl, then with gradually reduced guanidine HCl concentrations with a final wash containing only 1X Wash buffer without guanidine HCl. Samples were eluted with 1X Elution buffer (1 M imidazole, 500 mM NaCl, 20 mM Tris HCl) into several fractions.

For anti-PrgY polyclonal antibody production, protein was purified as described above, except 1% of n-Dodecyl B-D-Maltoside detergent was used to solubilize proteins instead of 6M guanidine HCl.

Production of PrgY Polyclonal Antibody

A pre-immune bleed (5 mL) was collected on Day 0 as control serum. On Day 1, about 0.25 ml (0.5 mg) of purified full length PrgY was injected into a rabbit subcutaneously. On Day 14, a booster injection of 0.25 mg of purified PrgY was administered subcutaneously. On Day 28, a second booster injection of 0.25 mg of purified PrgY was administered subcutaneously. Day 35, 5 mL of serum was collected and assayed by Western blotting. On Day 42, a third booster of 0.25 mg of purified PrgY was administered subcutaneously. On Day 56, 50 mL of serum was collected. The antibody was absorbed with 100x

volume of OG1RF cell extract overnight at 4°C and assayed with Western blot at an antibody concentration of 1:10,000.

Protein Binding Assays

Pull-down Assay: Recombinant cellular lysates were prepared by sonication and lysozyme lysis (see Protein Expression and Purification). The cell membrane fraction was recovered by high-speed centrifugation. This mixture was incubated with 3 mL of Nickel Sepharose (ThermoFisher Scientific) and then washed with phosphate buffer to remove non-specific proteins. Then synthetic cCF10 or a peptide variant was added to the column to allow binding to N-Y for 1 hour at 4°C. The unbound flow thru fractions were collected and cCF10 activity was detected by microtiter clumping assays.

Surface Plasmon Resonance (SPR): These assays were carried out in collaboration with Dr. Albert Tai of Tufts University. The purified His tag N-Y was captured in the NTA chip (GE Life Sciences) using a Biacore T100 (GE Life Sciences). For each cycle, a 2 mM Ni²⁺ solution was first injected onto flow cell, followed by the injection of N-Y onto the flow cell for capturing. Once the capturing was completed, the peptide was injected across both the reference flow cell and experimental flow cell for 4 minutes. Once the injection was done, the system was allowed to dissociate for 15 minutes, during which the bound peptide will start to dissociate. At the end, there are regeneration cycles to remove all Ni²⁺/protein/peptide from the surface prior to the next binding cycle. For each cycle, one concentration of peptide is tested, and each concentration is tested twice. The range of peptide tested in this experiment was 0.9 nM – 630

nM. Binding signals were corrected for non-specific binding by subtracting the signal from the control flow cell using Biacore T100 Evaluation Software version 1.1 (<http://www.biacore.com>). The affinities between N-Y and peptides were calculated by using the 1:1 Langmuir binding model. The fit analysis of the model to the data was determined by χ^2 value, which is the square of the differences between the theoretical ideal curve to the actual curve, calculated according to this equation:

$$\chi^2 = \sum(r_f - r_x)^2 / (n - p)$$

where r_f is the fitted value at a given point, r_x is the experimental value at that point, n is the number of data points, and p is the number of fitted parameters.

Pheromone Activity in the Presence or Absence of PrgY

A microtiter assay was used to detect effects of purified N-Y protein on pheromone activity. A 96 well plate was filled with M9-YE medium. A 1:2 serial dilutions of synthetic cCF10 was made across a 96 well plate, starting with 2 ul of 50 ng/ml cCF10. Purified N-Y or PrgY derivatives containing amino acid substitutions were added to appropriate rows and incubated for one hour at 37°C. Responder cells were added and incubated for 2-4 hours. Relative pheromone induction was detected by the formation of visible clumps in donor cell cultures exposed to the synthetic cCF10 peptide in the presence or absence of purified N-Y.

SDS-PAGE and Western Blot Analyses

Cell cultures were grown overnight (14 hrs) in THB or M9-YE at 37°C, diluted into fresh THB with 25 ng/ml nisin, then grown for 2.5 - 3hr (1.2 - 1.5 O.D.). Cells were collected by centrifugation and treat with lysozyme solution or mutanolysin (10 mg/ml lysozyme and/or 250 U/ml mutanolysin, 10 mM Tris-HCl (pH 8.0), 50 mM NaCl, 10 mM EDTA) at 37°C for 15 minutes.

Crosslinking Assays

OG1RF pMSP3545 YSN overnight culture was prepared in M9-YE or THB with 10 ug/mL Erm at 37°C, diluted into fresh M9-YE or THB with 25 ng/mL nisin and grown for 3 hr at 37°C. Cells were collected by centrifugation and washed with 100 ul of PBS buffer for each 100ul of cell culture. Cells were resuspended with 100 ul of PBS and DSG at 0mM, 1 mM, 1.5 mM, and 2 mM. Cells were incubated at room temperature for 30 minutes and 15 mM Tris was added to each sample and continue to incubate at room temperature for 15 minutes to quench the reaction. Cells were collected by centrifugation and were resuspended with 50 ul of 5 mg/ml lysozyme solution for 30 minutes at 37°C. Cells were resuspended into SDS buffer for Western blot analysis.

In-Gel Trypsin Digest for Mass Spectrometry Analysis

Protein bands from polyacrylamide gels were excised with a sterile razor blade and were cut into ~2 x 2 mm cubes. Gel bands were washed and de-stained (removed silver stain) with 1:1 solution of 30 mM potassium ferricyanide and 100 mM sodium thiosulphate made freshly before use. Excised gel bands were incubate with solution for 8 minutes, then washed 4 x with 1 mL of Mill-Q

water. Gels were then incubated with 75 ul 1:1 100 mM ammonium bicarbonate:acetonitrile for 15 min at room temperature then washed with 75 ul of 100% acetonitrile for 30 sec – 1 min. After pieces shrunk and turned white and semi-opaque, acetonitrile was removed and gels were rehydrated with 75 ul 10 mM DTT in 50 mM ammonium bicarbonate, incubated for 1 hr at 56 °C in water bath. Then DTT solution was removed and incubated with 75ul of 55 mM iodoacetamide in 50 mM NH₄HCO₃ for 30 min at room temperature in the dark (iodoacetamide is light sensitive). The iodoacetamide solution was removed and washed with 75 ul 1:1 acetonitrile:100 mM ammonium bicarbonate and rehydrated in digestion buffer at 4 °C (50 mM NH₄HCO₃, 5 mM CaCl₂, 5 ng/ul trypsin) on ice for 15 min. Replaced solution with 70 ul 50 mM NH₄HCO₃, 5 mM CaCl₂ and incubated at 37 °C overnight in a warm-air incubator. The supernatant from overnight incubation was recovered and purified with C-18 column (Thermo Scientific) prior to submission for mass spec analysis.

Shotgun MS analysis

The protein sequence database was obtained from UNIPROT (<http://www.uniprot.org>). MASCOT software (<http://matrixscience.com>) was used to match the fragmentation spectra obtained from a trypsin digested protein sample to a protein sequence database. In general, the trypsin digested peptides are mostly between 7 and 20 amino acids long, with an average of 11 amino acids. Scaffold software (<http://www.proteomesoftware.com>) was used to combine the most important data produced by MS with the identification process, and to distribute the results in an organized table. Results listed in the table are

the identified proteins, the accession number, predicted molecular weight, the probability, and the number of assigned spectra. The numbers of assigned spectra give an idea of the confidence of the identification. In general, a good identification should have at least 4 or more spectra that correspond to a particular protein [27]. The filtering parameter was set at 95% probability with 1 peptide minimum.

Computer Analysis and Structural Modeling

Comparative modeling of PrgY was carried out with Protein Homology/analogy Recognition Engine V 2.0 (Phyre2) hosted at <http://www.sbg.bio.ic.ac.uk/phyre2/>. Comparative modeling of the human TIK11 active site was previously carried out by MODELLER, followed by KoBaMIN structural refinement [28]. Comparative modeling of PrgY active site was carried out with Phyre2 and analyzed with PyMOL.

Mass Spectrometry Analysis of Culture Supernatant

A PrgY expression construct was previously made by fusing the *prgY* gene in-frame with the first 3 codons of *nisA* gene in the nisin inducible promoter of pMSP3545S vector [14]. The plasmid pPCR4 was previously constructed to carry the *prgQ* Orf with an N-terminal peptide sequence derived from the iCF10 pre-processed sequence and a pheromone cCF10 at the C-terminal end. A cell culture was grown overnight (14 hrs) in M9-YE at 37°C. Cells were collected by centrifugation and were washed 3x with fresh M9-YE. Cells were then diluted into fresh M9-YE to an O.D. = 0.1 and grown for about 2 hr. At an O.D. = 0.3,

culture was induced with 25 ng/ml nisin and incubated for about 3 hr. At an O.D. = 1.2, cells were spun down by centrifugation and supernatant was collected for mass spectrometry analysis.

The culture supernatant was purified with C-18 column (C-18 Hypersep Thermo Scientific). Then sample was acidified with Trifluoroacetic acid (TFA) to bring the pH to 2. About 20 ml of culture supernatant was added to the C-18 column. Column was washed with 3ml of HPLC water with 0.1% TFA, then washed with 3 ml of HPLC water with 0.1 % TFA and 20% acetonitrile (ACN), and finally eluted with 1.5 ml of HPLC water with 0.1% TFA with 80% ACN. Eluted fraction was placed in a SpeedVac to evaporate ACN. The sample was then resuspended with 20 ul of HPLC water and 2ul was loaded to QTRAP 5500 to analyze the sample.

RESULTS

I. Purification and binding between PrgY and cCF10

Expression and purification of PrgY and variants

In order to understand the mechanism of PrgY and to demonstrate the direct interaction between PrgY and its cognate peptide cCF10, it was imperative to purify the protein. A PrgY expression construct was previously made by fusing *prgY* gene in-frame with the first 3 codons of *nisA* gene in the nisin inducible promoter of pMSP3545S vector [14]. Furthermore, *prgY* gene was previously PCR amplified to incorporate random mutations, cloned into pMSP3545S and screened for nonfunctional PrgY. The nonfunctional PrgY variants were identified on plate by their dry-colony phenotype that corresponds to clumping in broth culture. This screen resulted in eight point mutations and two truncation mutations that failed to reduce endogenous pheromone activity and resulted in clumping phenotype (Figure 4) [14]. Six of the eight point mutations were located on the N-terminal portion of PrgY, which is predicted to lie outside the membrane. Additionally, our lab has generated chimeric pPD1 TraB/PrgY proteins to identify region of the PrgY domain that confers specificity for cCF10. Amino residues 125 to 241 of PrgY, which are predicted to lie outside the membrane, are required for specific regulation of cCF10 *in vivo* [9]. Arginine 50 is a conserved residue throughout the entire PrgY family, and L142P is in the specificity region that was previously identified to be important for interacting with

the pheromone. Of the eight point mutations, the PrgY alleles encoding two amino acid variants, R50Q and L142P, were selected for cloning and protein purification.

Membrane proteins are often very difficult to purify because they are embedded in the lipid bilayer and require detergents to become soluble in aqueous solution. Since *prgY* mutations affecting endogenous pheromone control were located primarily in the N-terminal region of PrgY, we initially expressed and purified the N-terminal portion of PrgY (N-Y) containing residues 1-241. Plasmids pMSP3545S containing wild type *prgY* or the R50Q and L142P variant alleles were isolated, PCR amplified, and digested with *NcoI* and *XhoI*, and the amplicons were cloned into pET28b (Novagen, USA), an IPTG-inducible expression vector featuring a His Tag at the C-terminus for protein purification (see materials and methods). Cells were sonicated on ice to break down cellular membranes and fragmentation of DNA until the sample was no longer viscous otherwise the extract would be so viscous that it would clog the purification column. Unfortunately, N-Y and variants were all found in the insoluble fractions, suggesting inclusion bodies were formed. Changes in IPTG concentrations and induction times were tested but inclusion bodies still formed. Lowering the incubation temperature to 30 °C during IPTG induction resulted in fewer inclusion bodies and solubilized about 30% of the recombinant protein. The remaining inclusion bodies were solubilized with 6M guanidine HCl (see materials and methods). Ni-NTA agarose (Thermo Fisher Scientific) and purification column was used to purified His-tagged proteins. Early preparations resulted in very low

yields with some low molecular weight contaminants. We initially thought that perhaps PrgY was degraded by host proteases or that PrgY was not stable at room temperature since the purification step was done at room temperature. Next, the Halt Protease Inhibitor Cocktail (ThermoFisher Scientific) was used and the column purification step was done at 4 °C. This resulted in cleaner purified protein but still with low yield. We refrained from using Halt Protease Inhibitor Cocktail in subsequent purifications because we did not want the protease inhibitor to interfere with the biochemical activity of PrgY. Next, batch binding purification attempts were made in which His-bind resin was placed in a tube instead of the column. Membrane fractions with protein of interest were added to tubes and allowed to bind at 4 °C for 1 hour. Tubes were then spun down to separate the resin from the purified protein. This method resulted in a higher yield but not very clean preparations, most likely because the resin was disturbed between washes when the tube was transferred in and out of the centrifuge, allowing contaminated resin to resuspend into the purified supernatant. Lastly, membrane fraction with protein of interest was added to the column and allowed to bind at 4 °C for 1 hour. The guanidine HCl concentration in wash buffers was gradually reduced with a final wash without guanidine HCl in order to minimize precipitations, and this resulted in the highest yield and the cleanest protein (Figure 6). The full length PrgY protein was purified as described above, except 1% of n-Dodecyl B-D-Maltoside detergent was used to solubilize proteins (Figure 7). Other detergents including Triton-X and n-Octyl- β -D-Glucopyranoside were tested but n-Dodecyl B-D-Maltoside gave the highest yield. A PrgY polyclonal

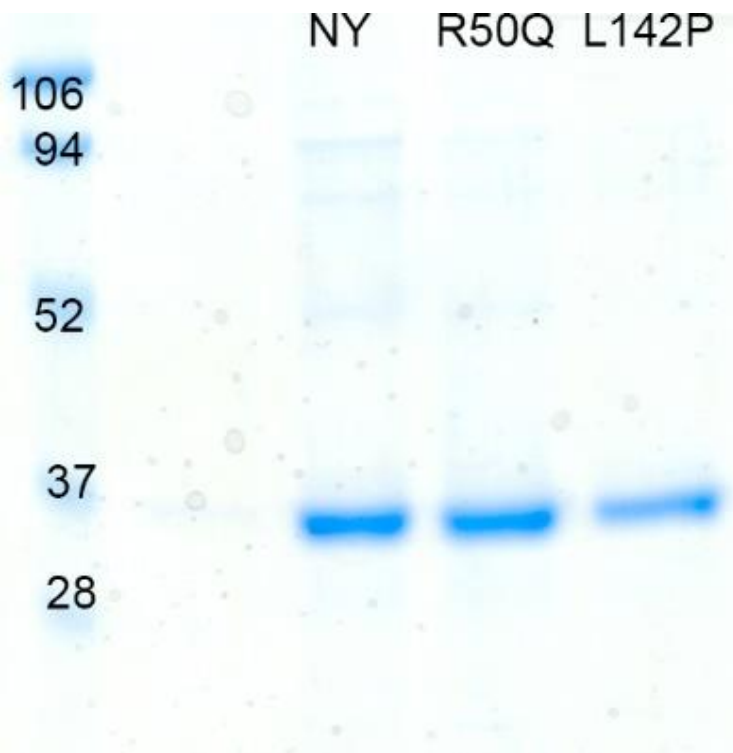


Figure 6. Purified NY and variants. The N-terminus portion of PrgY (N-Y), containing residues 1-241, and variants R50Q and L142P were expressed in *Escherichia coli* DH5 α , induced with 1mM IPTC final concentration, solubilized with 6M guanidine HCl, purified with Ni-NTA agarose (Thermo Fisher Scientific), and assayed with polyacrylamide gels with GelCode Blue Stain Reagent (ThermoFisher Scientific).



Figure 7. Purified full length PrgY. PrgY was expressed in *Escherichia coli* DH5 α , induced with 1mM IPTC final concentration, solubilized with purified with 1% of n-Dodecyl B-D-Maltoside, purified with Ni-NTA agarose (ThermoFisher Scientific), and assayed with polyacrylamide gels with Silver Stain Reagent (ThermoFisher Scientific).

antibody was developed by immunization of rabbits with purified full length PrgY (see Materials and Methods). The antibody was absorbed with OG1RF cell extract and assayed against OG1RF (pMSP3545S) and OG1RF(pMSP3545YSN) cell lysates, as control and PrgY, respectively, with Western blot at an antibody concentration of 1:10,000 (Figure 8). The PrgY antiserum reacted with a faint nonspecific band at ~70 kDa in both control lysate and lysate expressing PrgY, while an intense band migrating at the expected size for PrgY (~45 kDa) was detected in the lysate from the strain expressing PrgY. Although full length PrgY purified with n-Dodecyl B-D-Maltoside worked well for antibody production, it was very difficult to run biological assays with this protein, as it precipitated out of solution once the detergent was removed through dialysis. The full length PrgY was partially solubilized with 6M guanidine HCl, however, with very low yield.

N-Y specifically reduces the amount of pheromone flow-thru in pull-down assays, demonstrating the direct binding between cCF10 and N-Y in vitro.

Previous work suggests that cCF10 is specifically controlled by PrgY, however, a direct interaction between PrgY and cCF10 has not been shown. We initially used pull-down assays to demonstrate the direct and specific interaction between PrgY and cCF10. Cell membrane fractions containing His-tag N-Y (described above) were incubated with Nickel Sepharose which was then washed with phosphate buffer to remove weakly bound proteins. As a result, His-tag N-Y was bound to the column and allowed us to test its binding to cCF10.

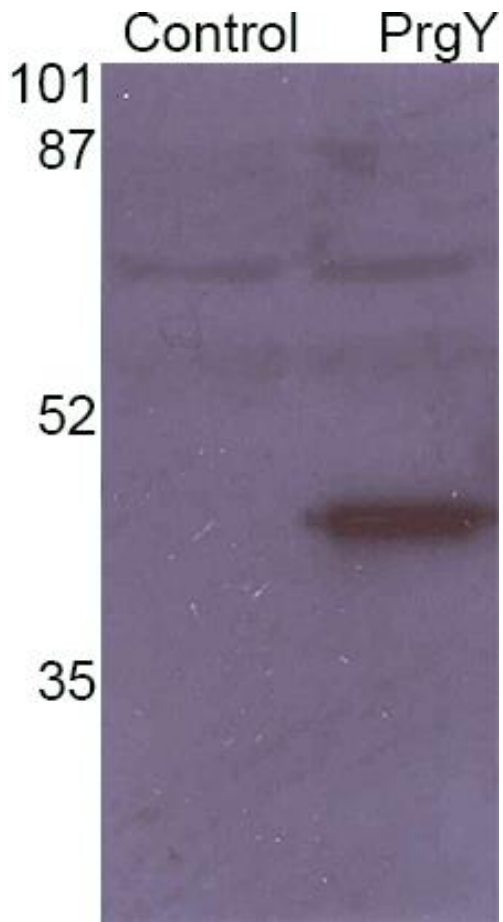


Figure 8. Western blot of PrgY polyclonal antibody. The PrgY antibody was absorbed with 100x volume of OG1RF cell extract overnight at 4°C and assayed with Western blot at an antibody concentration of 1:10,000. OG1RF (pMSP3545S) and OG1RF(pMSP3545YSN) cell lysates were assayed as control and PrgY, respectively.

Typically about 200 ug/ml of N-Y could be eluted from the Nickel column from 3 mL of cell membrane fraction containing His-tag N-Y. The molecular weight difference between N-Y and cCF10 is about 34-fold. Assuming a 1:1 molar ratio of binding between N-Y and cCF10, for every 3ml of cell membrane fraction that was added to the nickel column, 5.9 ug/mL of cCF10 added to the column and the flow through fractions containing unbound cCF10 were collected. The relative amounts of unbound cCF10 were detected by micro-titer assays. The His-tag N-Y bound to the nickel column was then eluted to determine total N-Y bound to column via Nano Spec. The cCF10 flow-thru fraction from the column without N-Y was assayed and normalized to 100 to account for nonspecific binding of cCF10 to the resin (Figure 9A). In the presence of N-Y, cCF10 in flow-thru fraction was significantly reduced to 0.4, indicating that most cCF10 that was added to the column was bound to N-Y (Figure 9A). The fourth residue (L4) of the cCF10 peptide was previously identified to be important for PrgY recognition [9], and that an L4I peptide variant of cCF10 abolished PrgY ability to reduce pheromone activity in donor cells. Similarly, when L4I was added to the column, fewer L4I was bound to N-Y as indicated by higher L4I in the flow-thru fraction (Figure 9B). The peptide cPD1 pheromone which induces transfer of the pPD1 plasmid [29], was previously shown to have no pheromone activity in the pCF10 system but does interact with PrgY [14], partially blocked binding of and cCF10 to N-Y bound to the column, resulting in more cCF10 in flow thru fractions (Figure 9C). Although the absolute amount of cCF10 recovered from each column varied between experiments, the ratio of cCF10 recovered between N-Y bound

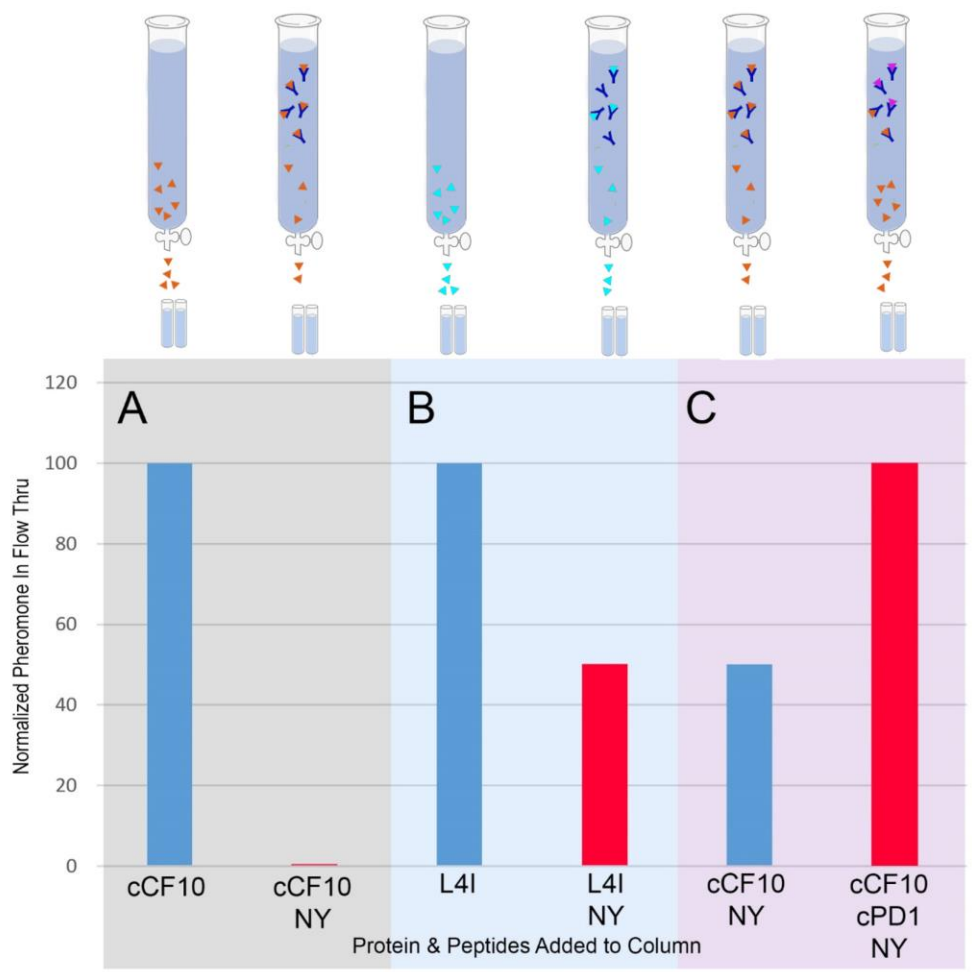
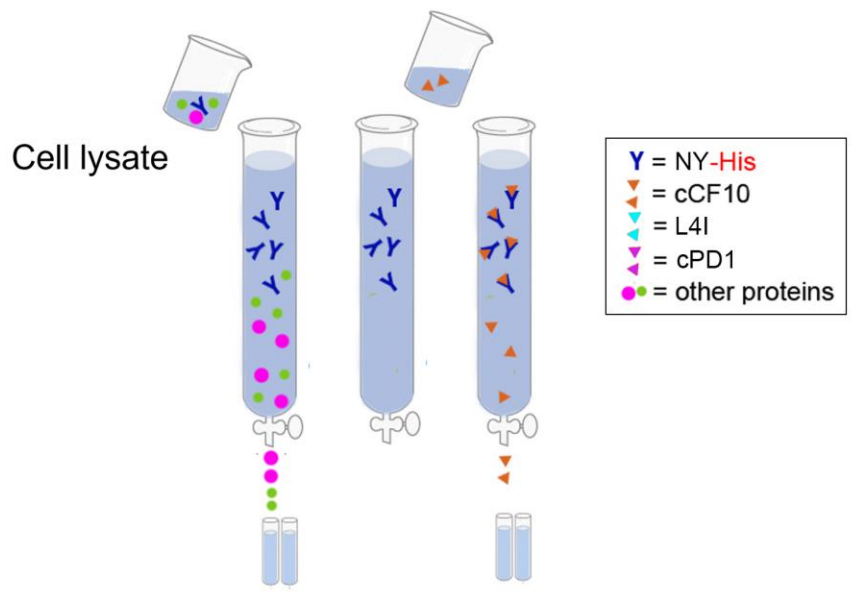


Figure 9. Pheromone flow-thru in the pull-down assay. Recombinant cellular lysates were prepared by sonication and lysozyme lysis. Cell membrane fraction was recovered by high-speed centrifugation. This mixture (contained about 200 ug/ml of N-Y) was then incubated with Nickel Sepharose and then washed with phosphate buffer to remove non-specific proteins. Either 5.9 ug/mL of synthetic cCF10 or a peptide variant was added to the column to allow binding to N-Y and the unbound flow thru fractions were collected and cCF10 activity was detected by microtiter clumping assays. **(A)** In the absence of N-Y, cCF10 flow-thru fraction was assayed and normalized to 100. In the presence of N-Y, cCF10 in flow-thru fraction was significantly reduced, indicating that most cCF10 that was added to the column was bound to N-Y. **(B)** The fourth residue of the cCF10 peptide was previously shown to be important for PrgY recognition, and when L4I was added to the column, less L4I bound to N-Y indicated by higher L4I in the flow-thru fraction. **(C)** With 5.9 ug/ml of competing peptide cPD1 (which has no pheromone activity) partially blocked binding of NY and cCF10, resulted in more cCF10 in flow thru fractions as indicated by higher titer. The data shown are representative of at least three independent experiments.

columns versus control columns remained the same.

Although the experiments described above confirmed the direct binding between cCF10 and N-Y *in vitro*, I experienced some difficulties quantify the binding interactions due to some non-specific binding of cCF10 to the resin. In order to determine the affinities between N-Y and cCF10, Surface Plasmon Resonance (SPR) kinetic analysis was performed. For this particular assay, I was working in collaboration with Dr. Tai, an expert in SPR, at the Biacore core facility at Tufts University. Purified his tag N-Y was immobilized to the NTA sensorchip and increasing concentration of synthetic cCF10 was injected over the sensor chip (see materials and methods). Once the injection was done, the bound peptides were allowed to dissociate. Binding signals were calculated with BIA evaluation software as previously described [30]. The data suggests that cCF10 binds to N-Y slowly (low on rate), and it also falls off very slowly (slow off rate) (Figure 10A). The dissociation of cCF10 from the immobilized N-Y on the NTA sensorchip was very slow, and had a monophasic morphology, consistent with a single receptor-ligand interaction. Furthermore, the binding of cCF10 to N-Y can be saturated at a level comparable to the calculated theoretical maximum (60RU) assuming 1:1 binding. As expected, no significant binding was detected between cCF10 and N-Y variant R50Q since R50 is located within the functional domain and is a conserved residue throughout the entire PrgY family (Figure 10B). Synthetic cCF10 variant L4I (633nM) was also injected stepwise into a flow cell containing NTA sensor chip with 2000RU of purified His-tagged N-Y

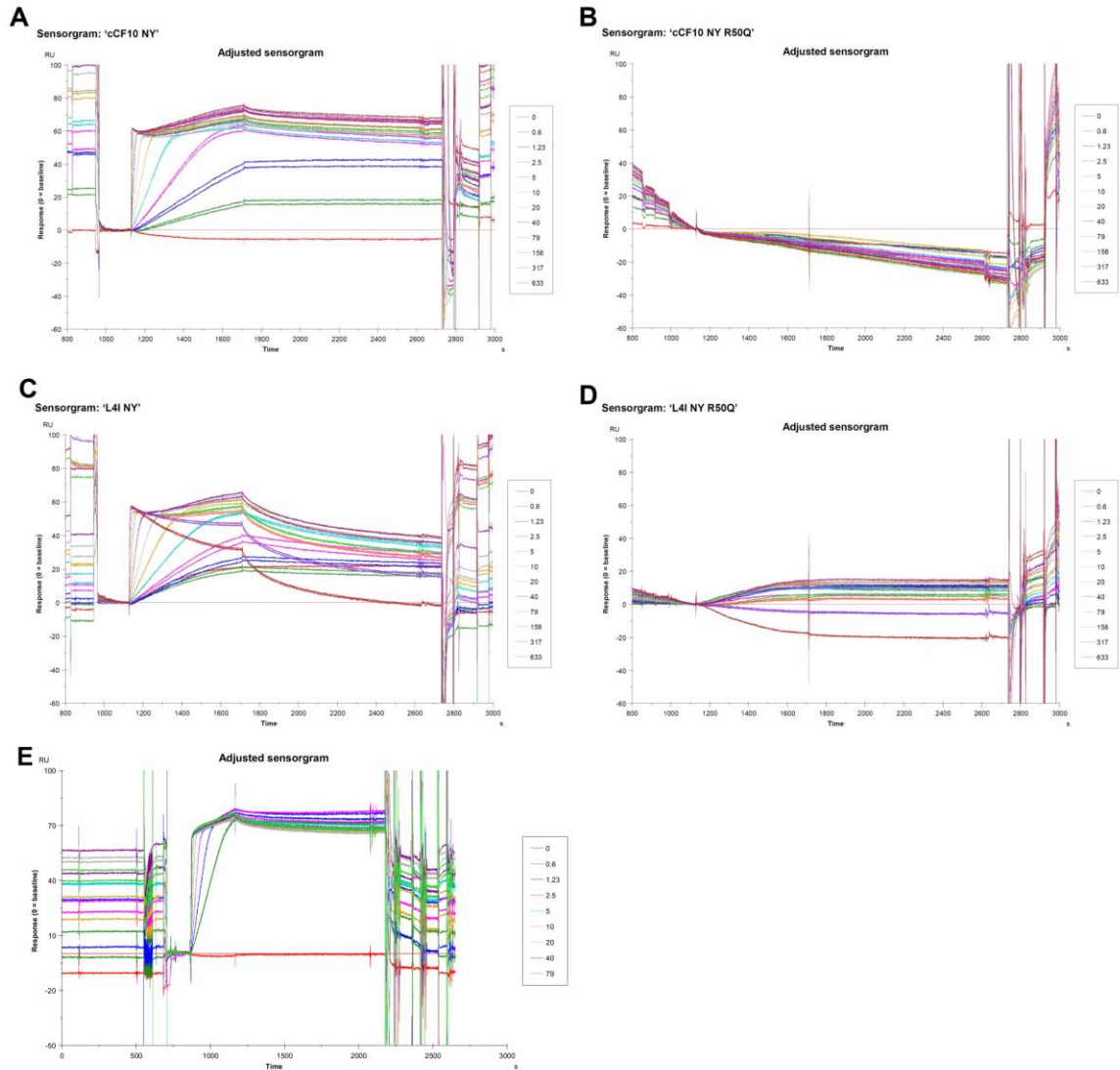


Figure 10. Surface plasmon resonance kinetic analysis of cCF10 binding to N-Y. Synthetic cCF10 (633nM) was injected stepwise into a flow cell containing NTA sensor chip with 2000RU of purified His-tagged N-Y (**A**) or N-Y mutant (**B**). The binding of cCF10 to N-Y can be saturated at a level comparable to the calculated theoretical maximum (60RU) assuming 1:1 binding (**A**). No significant binding between cCF10 and N-Y variant R50Q was detected (**B**). Synthetic peptide variant L4I (633nM) was also injected stepwise into a flow cell containing

NTA sensor chip with 2000RU of purified His-tagged N-Y **(C)** or N-Y variant R50Q **(D)**. The binding affinity of L4I to N-Y was lower than cCF10 to N-Y but can still be saturated at 60RU. No significant binding between L4I and N-Y variant R50Q. For comparison, synthetic cCF10 (79nM) was injected stepwise into a flow cell containing NTA sensor chip with purified His-tagged PrgX, which has a stronger binding affinity ($K_D \sim 10^{-13}$) **(E)**.

	ka (M⁻¹ s⁻¹)	kd (s⁻¹)	KD (M)	χ²
cCF10/NY	6.126x10⁶	1.637x10⁻⁴	2.672x10⁻¹¹	4.41
L4I/NY	2.000x10⁶	1.472x10⁻³	7.358x10⁻¹⁰	1.79

Table 2. Affinities of peptides for N-Y. Binding affinities yielded by kinetic modeling using lower concentration range (0.6 nM – 10nM). ka, association rate constant; kd, dissociation rate constant; KD, dissociation constant.

(Figure 10C) or N-Y variant R50Q (Figure 10D). The binding affinity of L4I to N-Y was lower than cCF10 to N-Y (Figure 10C), which was expected since fourth residue of the cCF10 peptide was previously shown to be important for PrgY recognition [31], but could still be saturated at 60RU. No significant binding between L4I and N-Y variant R50Q was detected (Figure 10D), which confirmed that the conserved arginine 50 was important for interacting with the pheromone. Binding signals were corrected for non-specific binding by subtracting the signal from the control flow cell. The pCF10 encoded PrgX, which controls the initiation of transcription of conjugation genes, also interacts with cCF10. Higher binding affinity of PrgX to cCF10 was observed compared to N-Y and cCF10 (Figure 10E). These observations indicate potential limitations of this *in vitro* assay as the purified PrgY lacks the transmembrane domain and consequently N-Y could be less able to effectively capture cCF10 than the membrane-anchored full-length version found in cells, while full length purified PrgX (a cytoplasmic protein) was used. The Langmuir model was used to determine the kinetic constants of these bimolecular interactions through SPR analysis. This model describes a 1:1 interaction in which one ligand molecule interacts with one analyte molecule as we hypothesized for PrgY and cCF10. The dissociation constant (KD) of N-Y and cCF10 calculated by this model was 2.672×10^{-11} M, suggesting a strong interaction between cCF10 and N-Y (Table 2). The binding curves fit well to the 1:1 Langmuir binding model ($\chi^2 = 4.41$). The KD of N-Y and L4I was 7.358×10^{-10} M, suggesting weaker interaction between N-Y and L4I compare to N-Y and cCF10. The binding curves also fit well to the 1:1 Langmuir

binding model ($\chi^2 = 1.79$). These affinities yielded by kinetic modeling using lower concentration range (0.6 nM – 10nM). No significant binding between N-Y variant R50Q and cCF10 as well as L4I and N-Y variant R50Q, thus no kinetic modeling was calculated for these interactions.

The amino-terminal fragment of PrgY specifically interacts with cCF10 in solution, but does not directly degrade or modify cCF10.

We initially hypothesized that PrgY reduces endogenous pheromone production in donor cells by specifically binding and degrading cCF10 as it is secreted across the cytoplasmic membrane. If this was the case then in the SPR experiment when synthetic cCF10 was injected over the sensor chip with immobilized N-Y, N-Y would have degraded cCF10 and should not give us a binding kinetic result. Results from the pull-down assay and SPR provided great binding kinetic data between N-Y and cCF10, however, these binding results described above did not indicate degradation. For this reason, we wanted to examine this further using microtiter assays to detect pheromone activity in the presence or absence of N-Y (see materials and methods). Briefly, the purified N-Y was either pre-incubated with cCF10 in M9-YE medium for 1 hour prior to adding responder cells OG1RF or was added at the same time as responder cells, and examined the effects of the purified N-Y protein on pheromone activity. We hypothesized that PrgY degrades or modifies cCF10 as it is released from cell membrane by PptAB after it has been processed by the protease Eep; these data suggested that the purified amino-terminal fragment of PrgY specifically

interacts with cCF10 in solution, but did not irreversibly sequester, or directly degrade or modify cCF10 (Table 3). Pheromone activity was not reduced in the presence of purified N-Y but rather increased when N-Y was incubated with synthetic pheromone for 1 hr prior to adding responder cells (Table 3A). No change in pheromone activity was observed if N-Y was added along with responder cells without pre-incubation with synthetic pheromone. Incubation of N-Y with the heterologous peptide pheromone cAD1 had no effect on cAD1 activity, indicating that the interaction between PrgY and cCF10 was specific (Table 3B). Point mutations (R50Q, L142P) that failed to reduce endogenous pheromone activity *in vivo*, also reduced or eliminated the *in vitro* binding interaction (Table 3C). Rather than detecting pheromone degradation, I found an apparent binding activity of N-Y for cCF10 that could potentially protect the pheromone from degradation by secreted enterococcal proteases in liquid medium (Table 3). The addition of N-Y enhanced clumping rather than reducing clumping. In the presence of N-Y, incubation of N-Y with the heterologous peptide pheromone cAD1 had no effect on cAD1 activity, indicating that the interaction between PrgY and cCF10 was specific. The specificity of the interaction is also supported by the fact that point mutations that abolished the ability of PrgY to control endogenous pheromone *in vivo*, also reduced or eliminated the *in vitro* binding interaction. The apparent increase in the titer of cCF10 incubated with N-Y probably resulted from N-Y protecting the soluble pheromone from protease digestion in the culture medium by binding to the free peptide (Figure 11). Microtiter assays for cCF10 activity actually comprise a race between the peptide getting degraded by host protease in the culture medium

A

Addition of N-Y	Incubation Pre (1hr)	Titer (3hr)	Titer (4hr)
+	+	256	2048
-	+	32	512
+	-	32	1024
-	-	32	1024

B

Addition of N-Y	Incubation Pre (1hr)	Titer (3hr)	Titer (4hr)
+	+	8	8
-	+	8	8
+	-	16	16
-	-	16	16

C

Addition of N-Y variants	Incubation Pre (1hr)	Titer (3hr)	Titer (4hr)
WT	+	32	64
R50Q	+	8	16
L142P	+	8	32
-	+	8	16
WT	-	8	16
R50Q	-	8	16
L142P	-	8	16
-	-	8	16

Table 3. Purified amino-terminal fragment of PrgY specifically interacts with cCF10 in solution, but does not directly degrade or modify cCF10. Micro-titer assays were used to detect pheromone activity in the presence or absence of purified N-Y. The purified N-Y was either pre-incubated with cCF10 in M9 medium 1 hour prior to adding responder cells OG1RF or was added the same time as responder cells. The titer reported represent the reciprocal of the highest dilution that induced clumping. **(A)** When N-Y was incubated with synthetic

pheromone for 1 hr prior to adding responder cells, pheromone activity was not reduced but actually increased. **(B)** Incubation of NY with the heterologous peptide pheromone cAD1 had no effect on cAD1 activity. **(C)** Point mutations (R50Q, L142P) that abolished the ability of PrgY to control endogenous pheromone *in vivo*, also reduced or eliminated the *in vitro* binding interaction.

and N-Y protecting the peptide long enough to allow the peptide to re-enter the cell, resulting in an apparent enhancement of pheromone activity. The high-affinity pheromone binding protein PrgZ, expressed on the surface of the responder cells in the assays removed the N-Y-bound peptide, and imported it into the cell for induction, leading to an increase in the apparent pheromone activity. These data suggest that PrgY might not directly degrade pheromone as we previously thought but it may have a chaperone-like activity involving “presentation” of the pheromone to a cell-associated protease or peptidase for degradation (Figure 12). Alternatively, the C-terminal domain of PrgY may be important for maintaining a conformation that is needed for optimal pheromone-degrading enzymatic activity, but not for specific binding. In this assay, only the soluble purified His-tagged N-terminal fragment containing residues 1-241 was incubated with cCF10, and thus N-Y might not be able to maintain a conformation that is required for enzymatic activity. As a result, no reduction of pheromone activity was observed in this *in vitro* assay. Initial attempts to use full length PrgY in this type of assay were not successful due to immediate precipitation of PrgY during in the incubation with peptides. PrgY was later successfully purified with reasonable efficiency and stability (see materials and methods). Micro-titer assays were used to detect pheromone activity in the presence of purified full-length PrgY, and results were very similar to those obtained with N-Y. PrgY appeared to bind to but not degrade cCF10, leading to an increased in titer due to the protection of the soluble pheromone from protease digestion in the culture medium. This reinforces the idea that the

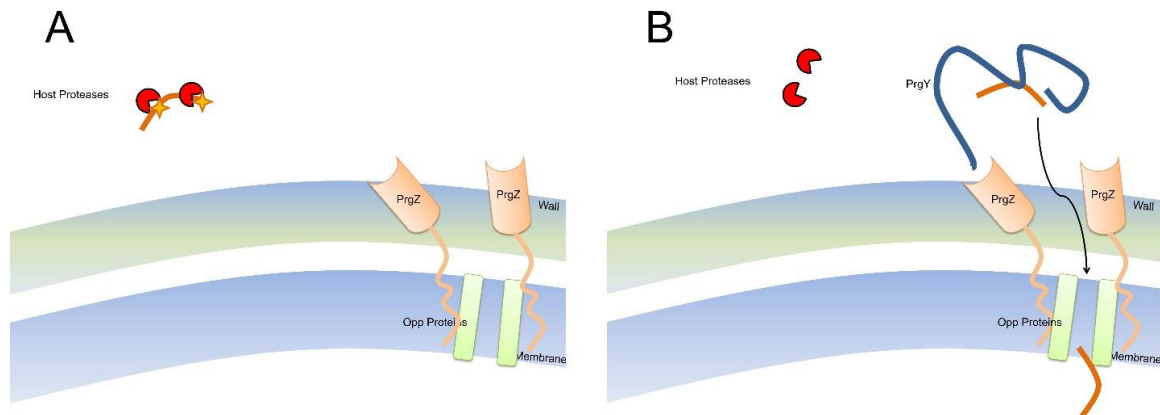


Figure 11. NY protects the soluble pheromone from host protease digestion in the culture medium by binding to the free peptide. (A) In the absence of purified N-Y, a portion of the peptide added exogenously was degraded by general host proteases. **(B)** N-Y was bound to the peptide but did not degrade the peptide and instead protected of the soluble pheromone from protease digestion in the culture medium. We hypothesized that the high-affinity pheromone binding protein PrgZ, expressed on the surface of the responder cells, removed the NY/peptide complex, and imported the protected pheromone, leading to an increase in the apparent pheromone activity.

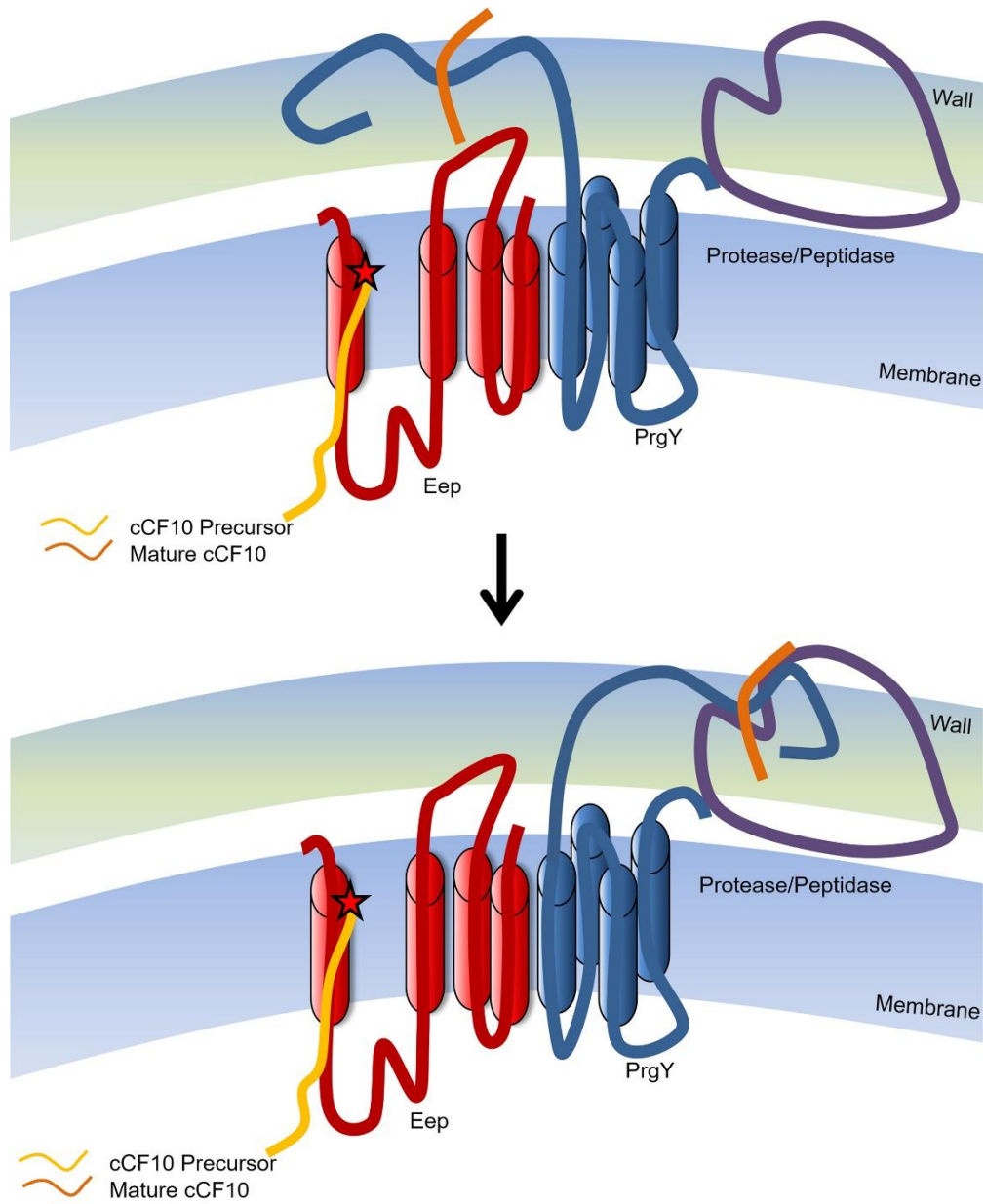


Figure 12. Model #2 of endogenous pheromone control by the pCF10 encoded protein PrgY. PrgY was thought to regulate endogenous induction by sequestration or degradation of pheromone within the cell wall, thereby reducing endogenous pheromone activity in donor cells. *In vitro* data suggests that PrgY specifically interacts with cCF10 in solution, but does not directly degrade or modify cCF10. An alternative model consistent with the data is that, PrgY may

have a chaperone-like activity involving “presentation” of the pheromone to a cell-associated protease or peptidase for degradation as depicted in this figure.

transmembrane portion of PrgY is not involved in catalytic activity. There is a possibility that the transmembrane protein was not active in solution or its interacting partner was missing when it was purified. Alternatively, the purified N-Y may lack a metal ion cofactor that plays a critical role in coordinating amino acids in the active site of PrgY.

II. Examination of the mechanistic models for PrgY function

Searching for a host protease that interacts with PrgY

The binding data with purified proteins suggests a model in which PrgY reduces endogenous pheromone production in donor cells by specifically binding to cCF10 and delivering it to one or more host proteins that degrade or modify the pheromone (Figure 12). To test this model, purified His tag PrgY protein was used as the “bait” to capture and ‘pull down’ a protein-binding partner, the “prey” from OG1RF lysate that binds to PrgY. His tag PrgY was mixed with an immobilized affinity ligand specific for the tag, thus creating a secondary affinity support to select other proteins that interact with PrgY. The prey protein sample was then poured into the column where bait/prey interactions will cause the prey to be immobilized to the column. DNA fragment encoding PrgY was amplified and cloned into the vector pET28b to obtain an N-terminal His-tag fusion expression plasmid. The fusion protein PrgY was induced with isopropyl β -D-1-thiogalactopyranoside (IPTG) and over expressed in *Escherichia coli* BL21(DE3) for purification. Briefly, N-Y-His was bound to nickel column, and OG1RF cell lysate (the membrane fraction which contain host proteases that may interact with PrgY) was added to the column and allowed for binding to N-Y. A negative control consisting of a non-treated affinity column (minus bait, plus prey protein sample) was used to help identify false positives caused by nonspecific binding of proteins to the Nickel column. Columns containing bound protein complexes were washed with high ionic strength buffers to eliminate any false positive result

due to nonspecific binding prior to elution with elution buffer. The eluted fractions were assayed with Silver Stain (ThermoFisher Scientific) on 15% acrylamide SDS gel. The N-terminus PrgY (N-Y) appeared to pull down some proteins from OG1RF cell lysate as indicated by the extra bands in elution fractions containing purified N-Y (Figure 13). These extra bands only appeared in the N-Y/OG1RF elution fractions but not in the N-Y or OG1RF control fractions. Four unique bands around 15-27 kDa and one blank spot (for control) were excised. These bands were de-stained, digested with trypsin, and extracted for mass spectrometry analysis (see materials and methods). Fragmentation spectra obtained from a trypsin digested protein sample were matched to a protein sequence database obtained from UNIPROT (<http://www.uniprot.org>) using MASCOT software (<http://matrixscience.com>). In general, the trypsin digested peptides are mostly between 7 and 20 amino acids long, with an average of 11 amino acids. Scaffold software (<http://www.proteomesoftware.com>) was used to combine the most important data produced by MS with the identification process, and to distribute the results in an organized table. Figure 14 contains the identified proteins, the accession number, predicted molecular weight, the probability, and the number of assigned spectra. The filtering parameter was set at 95% probability with 1 peptide minimum. Interestingly, bands #1,2, and 4 contained sequences of cCF10 (not shown), suggesting that cCF10 may be bound tightly to N-Y and was pulled down with N-Y, which agreed with our binding kinetic data with slow off rate (Figure 10). Gel bands #1 and #4 (Figure 13) likely contain nonspecific degradation product of N-Y since N-Y has a

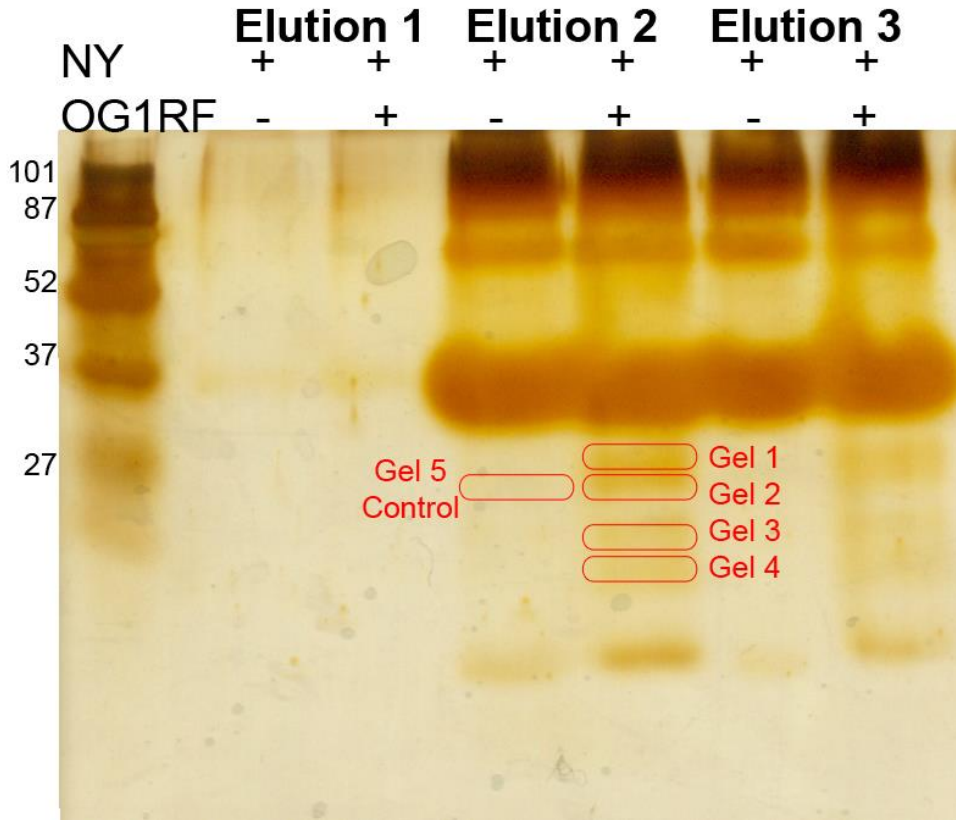


Figure 13. Pull-down assay of N-Y and OG1RF membrane fractions. The N-Y-His was bound to nickel column, then OG1RF cell lysate (membrane fraction) was added to the column and stored at 4°C for 1hr to allow binding. The column was then washed with buffer to remove unbound materials and finally eluted into several fractions. The eluted fractions were assayed with Silver Stain (ThermoFisher Scientific) on 15% acrylamide SDS gel. Four unique bands around 15-27 kDa and one control band were excised, de-stained, digested with trypsin, and extracted for mass spectrometry analysis.

molecular weight of 34 kDa and that bands #1 and #4 have 3 and 2 spectra, respectively, that correspond to PrgY sequence (Figure 14). GelE (20.7 kDa) and SprE (31.8 kDa) are host proteases that may be interacting with PrgY and lysozyme (14.7 kDa) or mutanolysin (23 kDa) are possible contaminants since they were used in OG1RF cell lysate preparation. The numbers of assigned spectra give an idea of the confidence of the identification. In general, a good identification should have at least 4 or more spectra that correspond to a particular protein [27]. With the exception of PrgY, all other proteins identified with Shotgun MS (Figure 14) had only 1 spectrum, with a 95% probability, suggesting that those proteins were most likely not present in the excised band samples.

The full length PrgY was also tested with pull-down assays. The full length PrgY appeared to pull down a ~35 kDa protein from OG1RF cell lysate, labeled “Band 7” (Figure 15). The unique band was extracted for mass spec analysis. Fragmentation spectra obtained from a trypsin digested protein sample were matched to a protein sequence database, as described above, to identify specific host proteins that may be interacting with PrgY. Unfortunately, band #7 corresponded to a degradation product of PrgY; lysozyme and trypsin were probably contamination from cell lysate and trypsin digest (Figure 16). Keratin type I and type II are known common contaminants and could safely ignored. There is a possibility that the binding of PrgY and cCF10 changes PrgY’s conformation that is required prior to binding to a host protease and degrades cCF10. Attempts were made to complex PrgY with synthetic cCF10 prior to

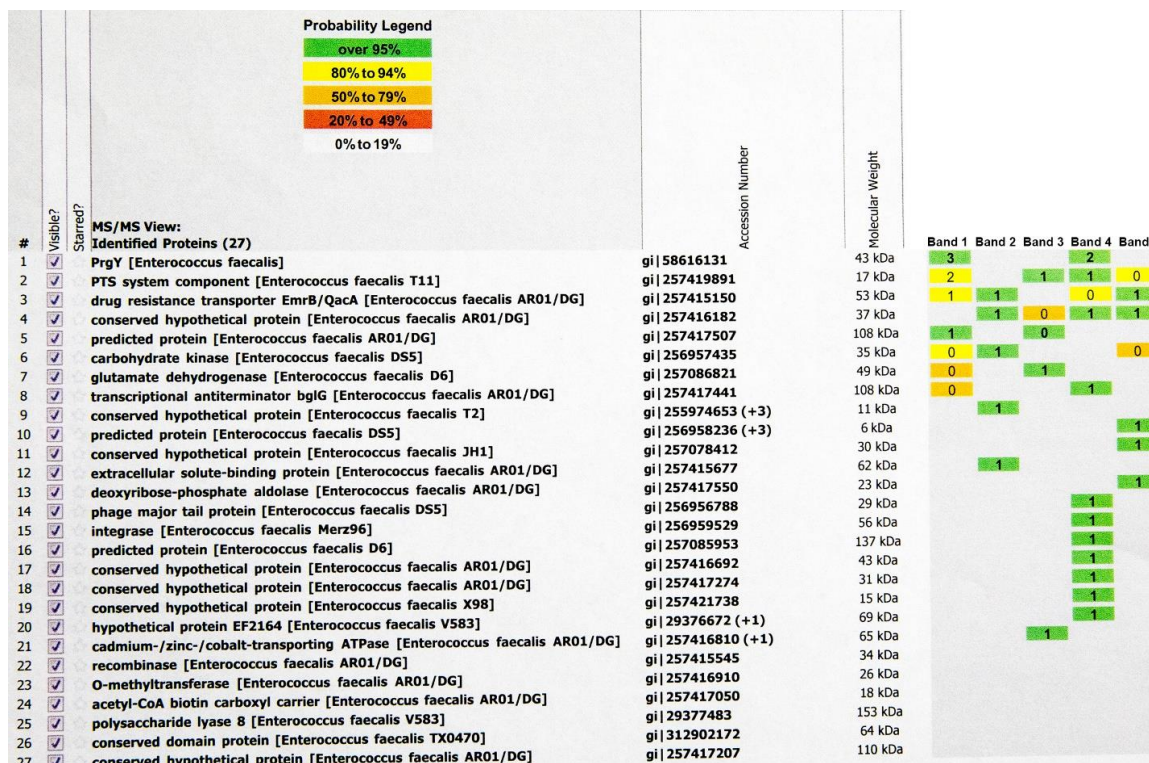


Figure 14. Shotgun MS analysis of unique gel bands of N-Y pull-down assay. Five unique gel bands were excised, trypsin digested, and fragmentation spectra obtained from these bands were matched to a protein sequence database obtained from UNIPROT (<http://www.uniprot.org>) using MASCOT software (<http://matrixscience.com>). Scaffold software (<http://www.proteomesoftware.com>) was used to combine the most important data produced by MS with the identification process, and to distribute the results in an organized table containing the identified proteins, the accession number, predicted molecular weight, the probability, and the number of assigned spectra. Filtering parameter was set at 95% probability with 1 peptide minimum. Gel band #1 and band #4 likely contained nonspecific degradation products of N-Y. In addition, band #1 and band #4 have 3 and 2 spectra, respectively, that

correspond to PrgY sequence, while all other proteins identified with Shotgun MS had only 1 spectrum, with a 95% probability, suggesting that those proteins were false positives and were most likely not present in the excised band samples.

pull-down with OG1RF lysate, however, results (not shown) were indifferent. It is interesting that PrgY, an important pheromone shut down protein, was very sensitive to degradation in the presence of OG1RF lysate. Perhaps the “real” mechanism of PrgY is to soak up all of the excess cCF10 and complex eventually gets degraded by various host proteases to reduce pheromone activity. This may explain why cCF10 binding to PrgY has a long off rate to allow PrgY to be saturated and degraded. Alternatively, PrgY’s target protein may be located in the membrane and is not soluble under the experimental condition of the cell lysate preparation, or the complex between PrgY and host protease may have dissociated during the pull-down assay if the interactions were weak.

Chemical cross-linking is another tool for detection and characterization of protein-protein interactions. In general, individual protein members in a noncovalent complex assembly remain in close proximity which is within the reach of the two reactive groups of a cross-linker. Thus cross-linking reactions have potential for linking two interacting proteins which exist in close proximity. Briefly, chemical cross-linking experiments was carried out by first linking the interacting proteins through covalent bonds using Disuccinimidyl glutarate (DSG), a water-insoluble homobifunctional crosslinker based on the amine-reactive N-hydroxysuccinimide (NHS) ester group which was designed for membrane proteins. (see materials and methods). The DSG has a spacer arm of 7.7 Å, and is cell membrane-soluble. OG1RF pMSP3545 YSN culture was prepared in M9-YE. Cells were collected by centrifugation, crosslinked with DSG and assayed with Western blot. A unique crosslinked band at ~95kDa was identified (Figure

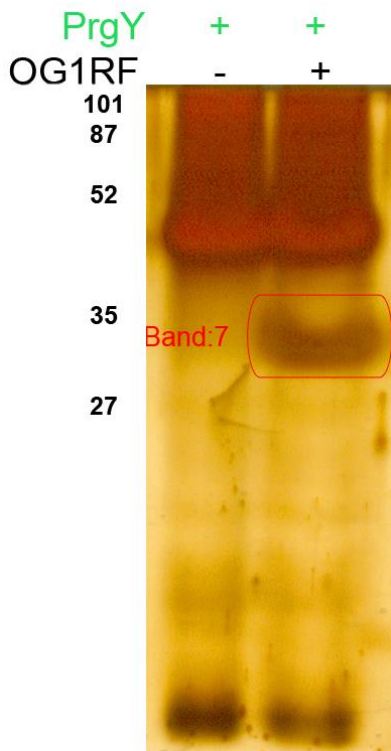


Figure 15. Pull-down assay of full-length PrgY and OG1RF membrane fractions. N-Y-His or Full-length PrgY-His were bound to nickel column, OG1RF cell lysate (membrane fraction) was added to the column and stored at 4°C for 1hr to allow binding. The column was then washed with buffer to remove unbound materials and finally eluted into several fractions. The eluted fractions were assayed with Silver Stain on 15% acrylamide SDS gel. Unique bands were excised, de-stained, digested with trypsin, and extracted for mass spec analysis.

#	Visible?	Starred?	Identified Proteins (8)	Accession Number	Molecular Weight	Band 7
1	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Lysozyme C precursor (1,4-beta-D-acetylmuramidase C) (Allergen Gal d 4) (Allergen Gal d IV) cRAP	gi 126608 sp P00698.1 LYSC_CHICK	16 kDa	100%
2	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Trypsin precursor cRAP	gi 136429 sp P00761.1 TRYP_PIG	24 kDa	100%
3	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Pheromone cCF10 shut down protein O5=Enterococcus faecalis GN=prgY PE=4 SV=4	tr Q51644 Q51644_ENTFL	43 kDa	100%
4	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Keratin, type I cytoskeletal 15 (Cytokeratin-15) (CK-15) (Keratin-15) (K15) cRAP	gi 75058787 sp O77727.1 K1C15_SHEEP	49 kDa	57%
5	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Keratin, type I cytoskeletal 10 (Cytokeratin-10) (CK-10) (Keratin-10) (K10) cRAP	gi 147744568 sp P13645.4 K1C10_HUMAN	60 kDa	100%
6	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Keratin, type I cytoskeletal 9 (Cytokeratin-9) (CK-9) (Keratin-9) (K9) cRAP	gi 81175178 sp P35527.2 K1C9_HUMAN	62 kDa	100%
7	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Keratin, type II cytoskeletal 2 epidermal (Cytokeratin-2e) (K2e) (keratin-2) cRAP	gi 547754 sp P35908.1 K22E_HUMAN	66 kDa	100%
8	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Keratin, type II cytoskeletal 1 (Cytokeratin-1) (CK-1) (Keratin-1) (K1) (67 kDa cytokeratin) (Hair alpha protein) cRAP	gi 1346343 sp P04264.5 K2C1_HUMAN	66 kDa	100%

Figure 16. Shotgun MS analysis of unique gel bands of PrgY pull-down assay. A unique gel band was excised, trypsin digested, and fragmentation spectra obtained from a unique gel band was matched to a protein sequence database, to identify specific host proteins that may be interacting with PrgY. Band #7 matched a degradation product of PrgY while lysozyme and trypsin were probably contamination from cell lysate and trypsin digest. Keratin type I and type II are known common contaminants and thus were ignored.

17), however, when analyzed with mass spectrometry (as described above), PrgY was not identified in the band. PrgY was probably cross-linked in such a way that mass spec could no longer identify the protein. In theory, that unique band should contain PrgY along with its binding partners. A possible binding partner is Eep, however, when performed the same crosslinking assay with a Δ Eep strain, the same result was observed, suggesting that the observed crosslinked species was probably not related to Eep. Attempts were made with immunoprecipitation to isolate PrgY from background proteins using PrgY antibody, however, there were too many background proteins, which caused mass spec column to clog. When the sample was diluted, the background noise was high and the spectrum corresponding for PrgY was not identified.

Comparative modeling of PrgY: a Structural Homolog of a Metalloprotease (TIKI) that modulates Wnt Signaling in Eukaryotes

Results from pull-down assays and crosslinking assays were inconclusive, and suggested that N-Y and full-length PrgY were easily degraded in the presence of OG1RF cell lysate, under the experimental conditions. Meanwhile, recent data suggested that PrgY shares significant homology to the eukaryotic metalloprotease TIKI that has been shown to cleave the amino terminus of mature Wnt proteins, thereby regulates the Wnt signaling. Using a structural prediction algorithm, Bazan *et al* [28] indicated that PrgY and TIKI share key conserved residues in the metal-binding catalytic core as well as overall

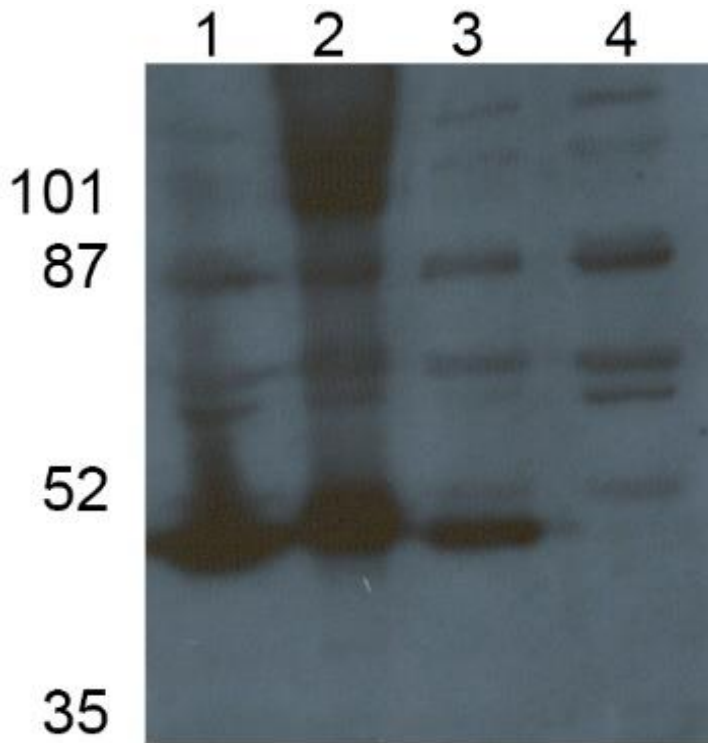


Figure 17. Crosslinking assay with *E. faecalis* OR1RF cell lysate overexpressing PrgY. Plasmid pMSP3545YSN expressing wild-type prgY in OG1RF was used in this crosslinking assay with Disuccinimidyl glutarate (DSG), a water-insoluble homobifunctional crosslinker based on the amine-reactive N-hydroxysuccinimide (NHS) ester group. Lanes 1 and 2 contain OG1RF (pMSP3545YSN) crosslinked with 1mM and 1.5mM of DSG respectively. Lanes 3 contains OG1RF (pMSP3545YSN) without DSG for positive control of PrgY expression. Lane 4 contains OG1RF (pMSP3545) without DSG for negative control. The dark band at ~48 kDa corresponds to PrgY monomer. A unique crosslinked band at ~95kDa was identified.

secondary structure. PrgY and TIKI share a common GX2H motif and a conserved glutamate. Furthermore, 6 conserved amino acids have been predicted in the putative active site core of PrgY. Although TIKI and PrgY proteins share limited sequence identity, their predicted secondary structures reveal a similar placement of β -strands and α -helices to anchor several potential conserved catalytic residues, consistent with a common evolutionary origin. Based on the structural similarity between PrgY and TIKI, we favor our original hypothesis in which PrgY reduces endogenous pheromone production in donor cells by specifically binding and degrading cCF10 as it is secreted across the cytoplasmic membrane (Figure 18). To gain insights into the biochemical mechanism of PrgY, additional structure predictions were used to construct a three-dimensional template for PrgY based on distantly related proteins with known crystal structures. Comparative modeling of PrgY was carried out with Phyre2 (see materials and methods) (Figure 19). The N-terminus PrgY (residues 1-241) is predicted to lie outside the membrane, and residues 125 to 241 of PrgY are required for specific regulation of cCF10 *in vivo* [9], shared significant homology to EreA/ChaN-like superfamily. EreA is an erythromycin esterase, which inactivates macrolide antibiotic erythromycin through the hydrolysis of the macrolactone ring [32]. Furthermore, conserved residues E43, H46, E74, and H285 of EreA have been identified in the active site cleft and have been shown to play a role in catalysis [32]. ChaN (*Campylobacter* heme acquisition) is a putative lipoprotein thought to associate with the outer-membrane and has been

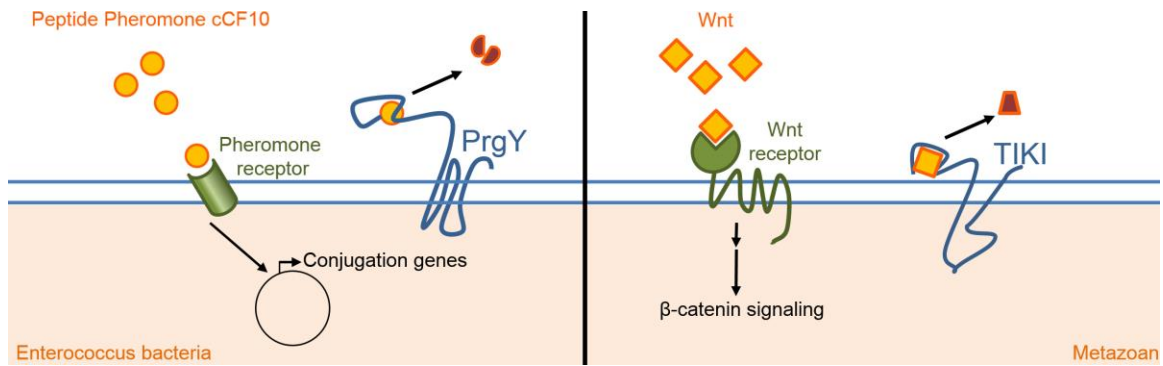


Figure 18. Mechanism of action by TIKI and the proposed mechanism of action by PrgY. TIKI is a metalloprotease that has been shown to cleave the amino terminus of mature Wnt proteins, thereby inactivating the Wnt ligand. Based on the structural similarity between PrgY and TIKI, we hypothesized that PrgY reduces endogenous pheromone production in donor cells by specifically binding and degrading cCF10 as it is secreted across the cytoplasmic membrane.





Template	Alignment Coverage	3D Model	Confidence	Template Information
d2g5gx1	<p>10-224</p>  <p>Alignment</p>		96.6	<p>Fold:EreA/ChaN-like Superfamily:EreA/ChaN-like Family:ChaN-like</p>
c4jkvA	<p>235-370</p>  <p>Alignment</p>		34.4	<p>PDB header:membrane protein PDB Molecule:cytochrome b562</p>

Figure 19. Comparative modeling of PrgY was carried out with Phyre2.

This figure describes the different parts of PrgY primary sequence homologous to other proteins with known structures. The N-terminus PrgY (residues 1-241), is predicted to lie outside the membrane, and residues 125 to 241 of PrgY are required for specific regulation of cCF10 *in vivo*, shared significant homology to EreA/ChanN-like superfamily. The C-terminus of PrgY is predicted to be in the membrane with four transmembrane domains, shared some homology to the membrane protein cytochrome b562. Other alignments with lower confidence level are not shown.

shown to be important in heme acquisition in the Gram-negative *Campylobacter jejuni* [33]. The transmembrane helices of C-terminus of PrgY shares some homology to the membrane protein cytochrome b562. These data suggests that PrgY may also contain four conserved residues in the putative active site similar to EreA: two His and two Glu. The 3D structure model of PrgY was deduced by comparison to the available structures of EreA, ChaN, and cytochrome b562 using Phyre2 and PyMOL (Figure 20). *PrgY* point mutants that failed to reduce endogenous pheromone activity (Figure 4) are not very close in proximity in the primary sequence, however, they are in very close proximity in the predicted 3D structure of PrgY (Figure 20), strongly suggesting that they may be important in both binding to cCF10 and catalysis. The N-terminus and C-terminus of PrgY are also predicted to be in very close proximity. Bazan *et al* used MODELLER, followed by KoBaMIN structural refinement, to perform comparative modeling of the human TIK1 active site [28] (Figure 21A). Comparative modeling of PrgY active site was generated with Phyre2 and analyzed with PyMOL (Figure 21B, C). The two proteins shared key conserved residues in the metal-binding catalytic core as well as overall secondary structure. Furthermore, four key conserved residues in the putative active site of TIK1 are H39, E66, E140, and H311 (Figure 21 A), map to the corresponding conserved residues in PrgY are H21, E45, D120, and H215 respectively (Figure 21B, C), implying a catalytic assembly for PrgY. The combination of the inconclusive results from pull-down assays and crosslinking assays, and recent data suggested PrgY shares significant homology to the eukaryotic metalloprotease TIK1, a membrane bound

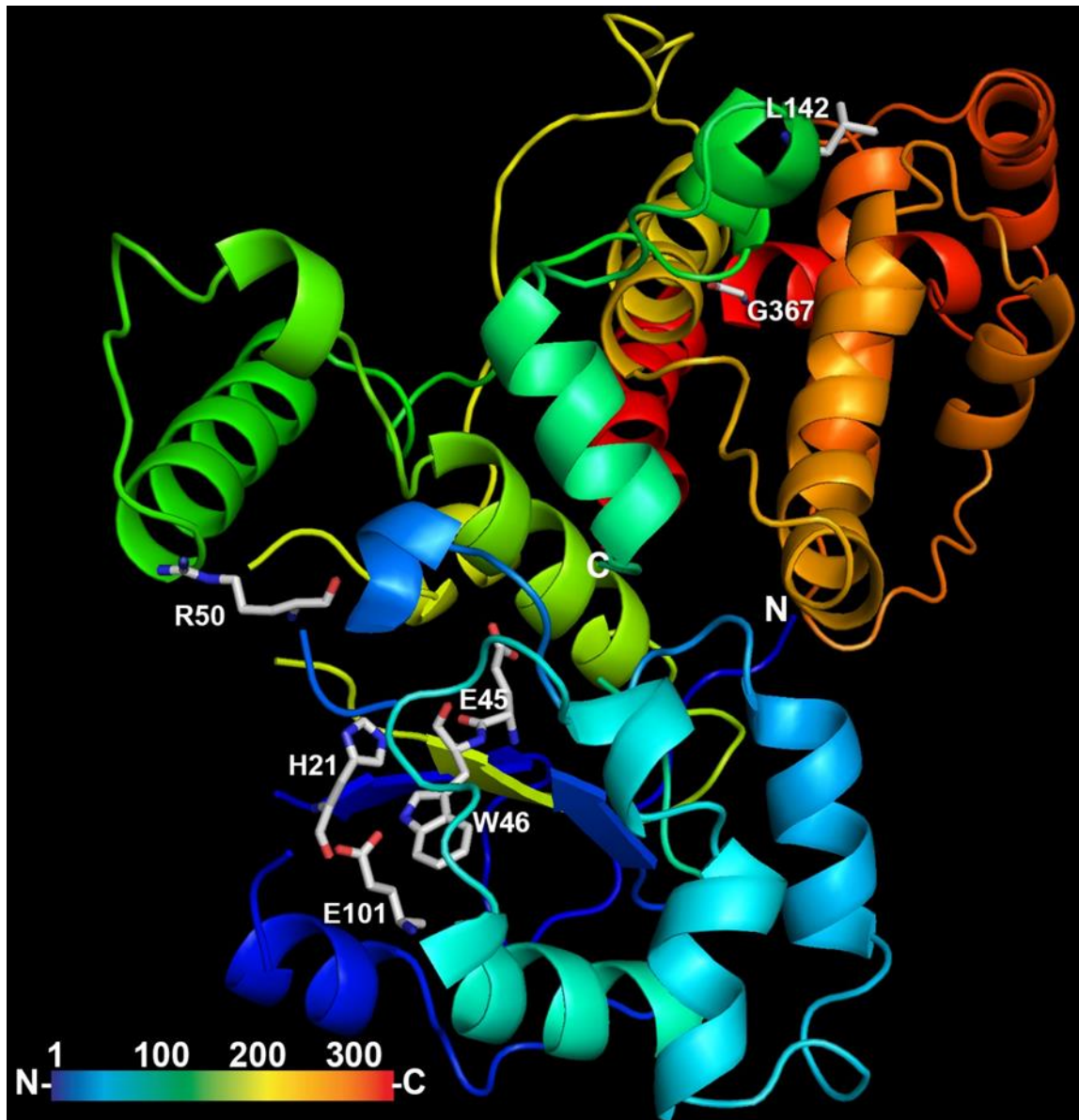


Figure 20. 3D structure model of PrgY. Structure predictions were used to construct a three-dimensional template for PrgY based on distantly related proteins with known crystal structures including EreA, ChaN, and cytochrome b562. *PrgY* point mutants that failed to reduce endogenous pheromone activity (Figure 4) are shown in “stick” form and are labeled. These residues are not very close in proximity in the primary sequence, however, they are in very close

proximity in the predicted 3D structure of PrgY. The N-terminus (dark blue) and C-terminus (red) of PrgY are also surprisingly in very close proximity.

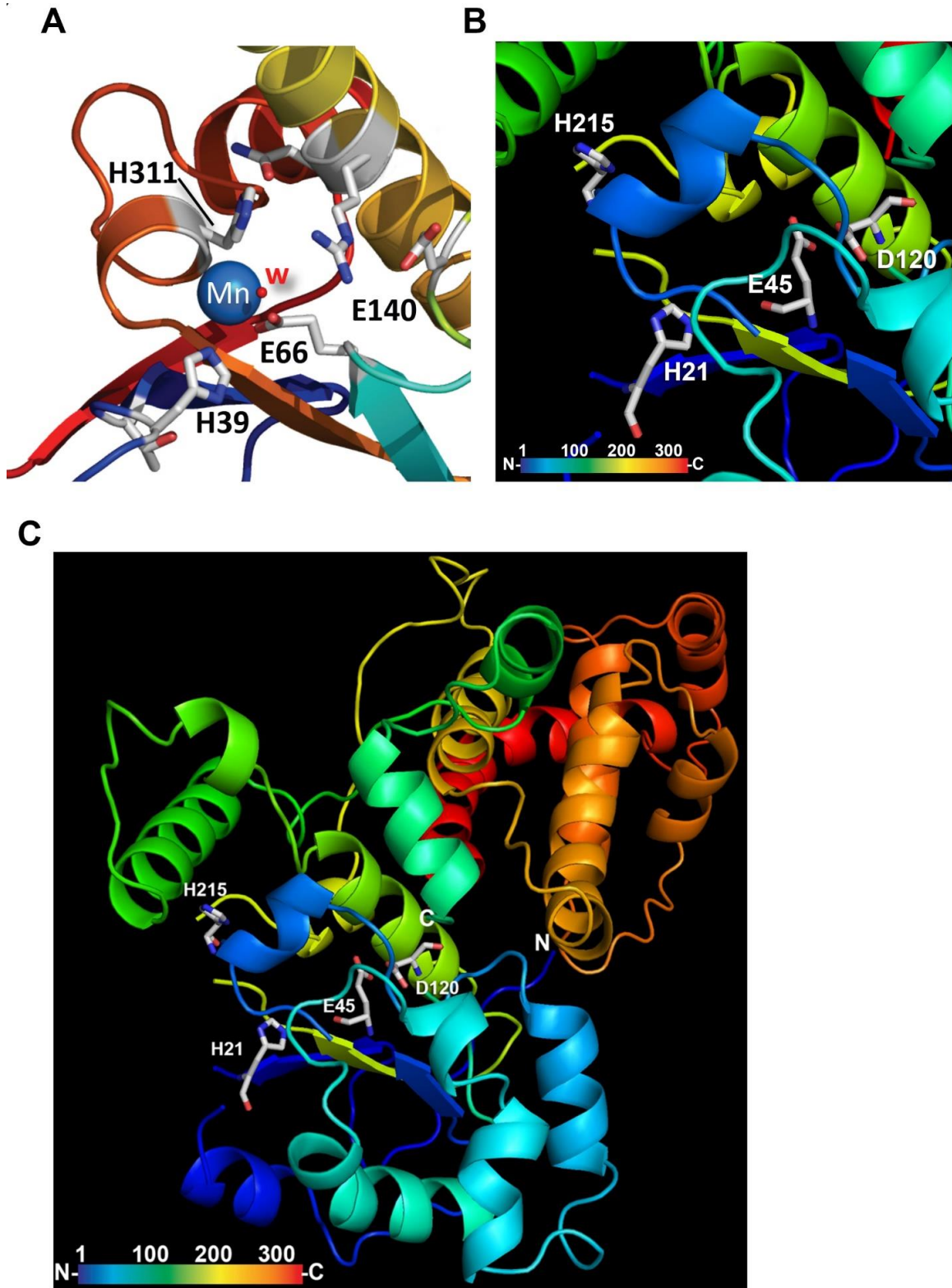


Figure 21. PrgY shares significant homology to the eukaryotic metalloprotease TIKI that regulates Wnt signaling. Comparative modeling of

the human TIKI1 active site was previously carried out by MODELLER, followed by KoBaMIN structural refinement **(A)**. Comparative modeling of PrgY active site was carried out with Phyre2 and analyzed with PyMOL. Four key conserved residues in the putative active site of TIKI are H39, E66, E140, and H311 **(A)**, and the corresponding conserved residues in PrgY are H21, E45, D120, and H215 respectively **(B)**. **(C)** Full 3D structure model of PrgY with the corresponding conserved putative active site residues.

protease that has been shown to cleave the amino terminus of mature Wnt proteins, thereby inactivating them. This led us to revisit our original model of PrgY in which PrgY was hypothesized to degrade cCF10 directly through its N-terminal domain.

PrgY cleaves cCF10 after the 2nd Leucine

There are 19 human Wnt proteins and TIKI was shown to cleave 10 human Wnt proteins [34]. Interestingly, the amino terminus of TIKI-cleavable Wnt are relatively hydrophobic, whereas the amino termini of non-cleavable Wnt are more polar with charged residues. These observations suggest that the hydrophobicity of Wnt amino terminus is a major determinant of TIKI cleavage susceptibility. This suggests that pheromones could be also susceptible to PrgY/TraB cleavage as pheromones are also very hydrophobic. Furthermore, TIKI was shown to cleave Wnt3a after leucine [34], suggesting that PrgY may also cleave cCF10 after leucine. Several possible degradation products of cCF10 were synthesized and analyzed with QTRAP 5500 to determine if it was able to reliably detect these small hydrophobic peptides. Transition states and retention times for each of the synthetic peptides were collected (Table 4). As expected, smaller fragments such as VFV, VTL, and LVTL had retention times of about 3 minutes and were eluted from the column around 30-40% ACN. Full length peptide had a retention time of 8 minutes and was eluted from the C-18 column at about 60 ACN.

Peptides	Transition states	Retention Time (min)
VTL	332 → 201, 233	3.75
VFV	364 → 247,265	4.60
LVTL	445 → 332, 233, 213	5.94
TLVFV	578 → 477 (not seen), 364, 461	8.42
LVTLVFV	790 → many transition states	11.24

Table 4. Mass spectrometry analysis of synthetic peptides. 50 ng of each synthetic peptides were added to M9-YE medium, purified with C-18 column, washed with 20% ACN with 0.1% TFA, and then eluted with 1.5 ml of 80% ACN with 0.1% TFA. QTRAP 5500 was used to identify each peptides, transition states, and retention times.

Next, mass spectrometry was used to examine the PrgY-dependent cCF10 degradation products in culture supernatants from a strain expressing PrgY compared to control culture supernatant from an isogenic PrgY⁻ derivative. A PrgY expression construct was previously made by fusing the *prgY* gene in-frame with the first 3 codons of *nisA* gene in the nisin inducible promoter of pMSP3545S vector [14]. Plasmid pPCR4 was previously constructed to carry the *prgQ* Orf with an N-terminal peptide sequence derived from the iCF10 pre-processed sequence and a pheromone cCF10 at the C-terminal end [9]. This strain produces about 400 ng/ml of cCF10, about 100-fold higher than cCF10 produced from the chromosomal *ccfA* allele in wild type strains. Culture supernatants were collected from two isogenic strains, one carrying both the PrgY and cCF10 expression plasmids, while the second strain only carried the cCF10 expression construct. We expected to find significant amount of cCF10 in both culture supernatants, but we anticipated that the strain carrying both expression plasmids would produce less cCF10, and possibly unique cCF10 fragments resulting from PrgY cleavage. Culture supernatants were purified with a C-18 column. Based on the transition states and the retention times from synthetic peptides (Table 4), QTRAP 5500 was used to examine PrgY-dependent cCF10 degradation products in donor culture supernatants. Peptide LVTL was identified only in culture supernatant from the strain expressing PrgY (Table 5), suggesting that PrgY cleaves cCF10 after the second leucine. Interestingly, VFV was identified in donor culture supernatant expressing PrgY, in control culture supernatant lacking PrgY, in control medium M9-YE +

	JRC105 pPCR4 (pMSP3545Y)	JRC105 pPCR4 (pMSP3545S)	M9-YE + LVTLVFV	M9-YE
Peptides Identified	LVTL VFV TLVFV LVTLVFV	VFV TLVFV LVTLVFV	VFV LVTLVFV	VFV

Table 5. Mass spectrometry analysis of culture supernatants. Mass spectrometry was used to examine the PrgY-dependent cCF10 degradation products in donor culture supernatant expressing PrgY (JRC105 pPCR4 (pMSP3545Y)) compare to control culture supernatant lacking PrgY (JRC105 pPCR4 (pMSP3545S)). Both strains lack *gelE* and *sprE*, the two major proteases in OG1RF. Peptide LVTL (bold) was uniquely identified in donor culture supernatant expressing PrgY.

synthetic LVTLVFV, and in M9-YE control medium. Both donor culture supernatant expressing PrgY and control supernatant lacking PrgY contained TLVFV, suggesting that a host exopeptidase in both culture supernatants may have cleaved two residues at the N-terminus of cCF10 (Table 5).

DISCUSSION

Based on previous genetic data [9], it was clear that cCF10 is specifically controlled by PrgY, however, a direct interaction between PrgY and cCF10 has not been demonstrated. Reduction of endogenous pheromone by PrgY pCF10-containing donor cells is critical for prevention of self-induction. PrgY was shown by genetic studies to target the mature cCF10 rather than other part of the signal peptide that it is processed from. It was also unknown whether PrgY directly degrades cCF10 or mediates its function indirectly through other proteins.

I have demonstrated that N-Y reduces amount of pheromone flow-thru in pull-down assays, confirming the direct binding between cCF10 and N-Y *in vitro* (Figure 9). The peptide cPD1 partially blocked binding of N-Y and cCF10 (Figure 9C). Surface Plasmon Resonance (SPR) spectroscopy was used to quantify the binding interaction between PrgY and cCF10. The data suggests that cCF10 binds to N-Y very slowly (low on rate), and it also falls off very slowly (slow off rate) (Figure 10A). This result is in agreement with my biochemical studies which required incubation of N-Y and synthetic pheromone for 1 hr prior to adding responder cells for activity (Table 3). The slow on rate suggests that in its natural environment, PrgY anchors in the cell membrane, probably in close proximity to Eep and in a precise orientation, such that it readily captures mature cCF10 immediately upon release from the membrane and before reimport by PrgZ. The slow on rate was most likely due to the limitation of this *in vitro* assay as it lacks

the transmembrane domain and consequently N-Y was unable to effectively capture cCF10.

The three best studied pheromone-inducible conjugative plasmids are pAD1, pPD1, and pCF10 [13]. These plasmids share a similar mechanism of regulation. There is partial cross reactivity between pPD1 and pCF10 system, however, there is no cross reactivity between the pAD1 and pCF10 systems. The pheromone protective activity of N-Y was specific and not observed with the pAD1-inducing cAD1 pheromone [35] (Table 3B). In addition, this activity was not observed with a L4I variant of cCF10 that is insensitive to PrgY inhibition *in vivo* [9]. Furthermore, the corresponding N-terminal fragments of two PrgY variants that failed to reduce pheromone production *in vivo* did not show this binding activity (Table 3C). Furthermore, the corresponding N-terminal fragments of two PrgY variants that failed to reduce pheromone production *in vivo* did not show this binding activity (Table 3C). This suggests that the extracellular N-terminal domain of PrgY has a specific cCF10 binding activity.

The results of this work revealed that N-Y directly interacts with cCF10. N-Y appeared to interact with cCF0 at the nM range and thus a long injection time was needed to increase the signal. Since the signal strength is directly proportional to the molecular weight, it might be possible to immobilize the peptide to the sensor chip and flow N-Y across; however, there may be some potential issues associated with immobilization of the small peptide, and perhaps a linker would be needed. Results from assays using various combinations of N-Y and peptide variants (Figure 9, 10) confirmed the binding specificity and

complemented subsequent *in vivo* and *in vitro* assays. The SPR setup conditions and fusion protein used in this study may be useful tools for subsequent kinetic analysis of proteins PrgY, PrgX, and PrgZ for cCF10 and iCF10 peptides. These three proteins are encoded on the same plasmid involved in conjugation and are responsible for the spread of antibiotic resistance genes. They serve different functions in regulating conjugation and yet all three proteins interact with the same seven amino acid peptide. Understanding how these three proteins evolved and how they interact with the peptides allows us to potentially design peptide mimetics that bind to these specific proteins and “freeze” the peptide/protein complex to a non-functional state in order to reduce or eliminate conjugation and the spread of antibiotic genes.

Results from the pull-down assay and SPR provided great binding kinetic data between N-Y and cCF10, however, no degradation was observed. These results suggest that the C-terminal domain of PrgY may be required to anchor PrgY in the membrane with a specific conformation that is needed for optimal pheromone-degrading enzymatic activity, but not for specific binding. For this reason, the absence of the C-terminal domain of PrgY in these *in vitro* assay only resulted in binding but not degradation. Initial attempts to use full length PrgY in this type of assay were unsuccessful due to immediate precipitation of PrgY during the assay. However, PrgY was later successfully purified with reasonable efficiency and stability and micro-titer assays were used to detect pheromone activity in the presence of purified full-length PrgY. Unfortunately, results with purified full-length PrgY were very similar to those obtained with N-Y. This

reinforces the idea that the transmembrane portion of PrgY is not involved in catalytic activity but is likely required to anchor PrgY in the membrane for optimal pheromone-degrading enzymatic activity.

The lack of peptidase activity observed with purified proteins may be related to folding problems in solution or its interacting partner was missing when it was purified. This led us to hypothesize that PrgY does not directly degrade pheromone as we previously thought but it may interact with a cell-associated protease or peptidase for degradation (Figure 11). We began searching for a host protease that may interact with PrgY using pull-down assays and crosslinking assays. Full length PrgY and N-Y appeared to pull down several unique proteins from OG1RF cell lysate (Figure 13, 15). Fragmentation spectra obtained from a trypsin digested protein sample were matched to a protein sequence database to identify specific host proteins that may be interacting with PrgY. Unfortunately, these bands corresponded to a degradation product of PrgY or contaminants from cell lysate and trypsin digest (Figure 16). In the case of N-Y, gel band #1 and band #4 likely contained nonspecific degradation products of N-Y since they had lower molecular weights than N-Y at 34 kDa (Figure 13, 14). In addition, band #1 and band #4 had 3 and 2 spectra, respectively, that corresponded to PrgY segments, while all other proteins identified with Shotgun MS had only 1 spectrum, with a 95% probability, suggesting that those proteins were false positives (Figure 14). The combination of these negative results and the new structural data on TIKI triggered a change in direction and led us to re-examine the protease activity of PrgY.

The striking similarity between the 3D structure of TIKI and the predicted structure of PrgY suggested that PrgY may have similar activity. The amino termini of TIKI-cleavable Wnt proteins are relatively hydrophobic compared to the amino termini of non-cleavable Wnt. Similarly, pheromones are also very hydrophobic, suggesting that pheromones could be also susceptible to PrgY/TraB cleavage. Mass spectrometer QTRAP 5500 was used to examine whether PrgY-dependent cCF10 degradation products can be detected in donor culture supernatants. Peptide LVTL was uniquely identified in donor culture supernatants expressing PrgY (Table 5). This suggests that PrgY cleaved cCF10 after the second leucine, and released the degraded peptide fragments LVTL and VFV that do not have any pheromone activity. TIKI was also shown to cleave Wnt3a after leucine [34]. This is an important piece of information as it suggests that PrgY is in fact a protease that degrades pheromone cCF10 rather than simply blocking the mature cCF10 from releasing into the cell medium. Peptide VFV was identified in all samples suggesting that VFV was most likely came from the casamino acids that was used to formulate M9-YE. Although both strains lack *gelE* and *sprE*, the two major proteases in OG1RF, there are likely other host exopeptidases in the culture supernatant that may have cleaved two residues at the N-terminus of cCF10, resulting in peptide fragment TLVFV in culture supernatants (Table 5).

We also examined the possibility of the degradation products of cCF10 for inhibitory activity similar to iCF10, and further controls endogenous pheromone. Arpan Bandyopadhyay, a graduate student in chemical engineering who works

on the pheromone project in collaboration with our lab, examined both the VFV and LVTL fragments of cCF10, and found that they do not have any inhibitory activity or pheromone activity. Therefore all of the functional activity of PrgY relates to reduction of endogenous pheromone and not to generation of an inhibitory peptide.

Proteases likely arose at the earliest stages of protein evolution as simple destructive enzymes essential for protein catabolism and the generation of amino acids. At a certain point, *E. faecalis* OG1RF (pCF10) acquired genes that coded for an early version of PrgY. This early version of PrgY likely did not contain the specificity domain for cCF10, and that the specific binding affinity to cCF10 was later acquired to enhance the regulation of the pheromone system and to minimize fitness cost. PrgY likely did not evolve from a single evolutionary line from one common ancestor. Interestingly, PrgY does not contain any structural motif that resembles known metalloprotease. Then at some point, metazoans, including humans and mammals, acquired the catalytic domain of PrgY to serve as a nonspecific enzyme necessary for protein catabolism, particularly degrading hydrophobic peptides. It is likely that the selective pressure and the extensive combinatorial activity served as a driving force in the protease transition from nonspecific enzymes such as early versions of PrgY and TIKI, to highly selective proteases responsible for critical proteolytic processing. The nonspecific degradative function of the early version of TIKI evolved and eventually links its catalytic domain to a specialized domain that provide substrate specificity, including the specificity region that recognizes Wnt proteins. It is interesting to

note that all known bacterial sex pheromones, which are putative PrgY/TraB substrates, are hydrophobic short peptides [10]. Similarly, TIKI-cleavable Wnt are relatively hydrophobic, whereas the amino termini of non-cleavable Wnt often contain more charged and polar residues [34]. These observations suggest that the hydrophobicity of the Wnt amino terminus is a major determinant of TIKI cleavage susceptibility. This allows TIKI to cleave at least 10 of the 19 Wnt proteins found in humans [34]. In contrast, PrgY is highly specific to cCF10 and that a L4I peptide variant of cCF10 abolished PrgY ability to reduce pheromone activity in donor cells [9]. These data suggest that TIKI's specificity region did not evolve from PrgY. Changes to PrgY occurred mostly at the sequence level leaving the structures and the catalytic domain rather conserved in TIKI. This can be used to explain the presence of homology between PrgY and TIKI, which share similar structures but have adapted to perform different functions. It would be interesting to clone the specificity region of *prgY* into *TIKI* and determine if TIKI could cleave cCF10, or vice versa. Based on previous experiments structural analysis, wild-type PrgY is unlikely to be able to cleave Wnt proteins as PrgY specificity region is quite specific to cCF10. Furthermore, PrgY may lack a cofactor that plays a role in coordinating amino acids in the active site of TIKI. Moreover, it would be interesting to consider how PrgY expression impacts on the host fitness, virulence, and the expression of aggregation substance, Asc10, as Asc10 has been shown to increase colonization and virulence by promoting biofilm formation and attachment to host tissues, and by increasing resistance to phagocytic killing [20-22].

TIKI is known to cleave Wnt, and the cleavage is enhanced by Co^{2+} or Mn^{2+} , but is inhibited by Ni^{2+} , Cu^{2+} , or Zn^{2+} *in vitro*. Unfortunately, I have been purifying PrgY with Ni^{2+} column which may inhibit its activity. For this reason, PrgY was later purified on a Cobalt column (ThermoFisher Scientific). The pheromone activity was assayed in the presence and absence of PrgY, purified with Cobalt column. Unfortunately, there was no significant difference in pheromone activity in the presence of PrgY purified with Cobalt column versus Nickel column (data not shown). Furthermore, PrgY was incubated with Co^{2+} , Mn^{2+} , Ni^{2+} , Cu^{2+} , or Zn^{2+} *in vitro* and assayed for its activity, however, no significant change in pheromone activity in the presence of these metal ion cofactors. It is very likely that these cofactors were readily available in M9-YE. Additional experiments, particularly in a chemically defined medium, are needed to elucidate the specific metal ion that is required and possibly enhance PrgY activity.

At the beginning of the post-genomics era, an enormous amount of information is available about the composition and organization of proteolytic systems in many living organisms. These broad genomic views have revealed that the protease landscape is vast and quite unexplored. For example, the recently identified TIKI represents a new metalloprotease family that is dependent on Co^{2+} or Mn^{2+} but lacks known metalloprotease motifs. The TIKI fold is considered a new addition to the metalloprotease pantheon. Likewise, structural modeling of PrgY using Phyre2 gives a model with a 96% confidence level over most of the extracellular domain (Figure 19), yet the model does not

contain any structural motif that resembles known metalloprotease. PrgY represents a new metalloprotease family similar to TIKI. The size of the different degradomes will continue to grow in the near future, as new enzymes with unique structural designs and catalytic mechanisms are identified and characterized. These protease studies will contribute to the elucidation of genetic diseases caused by mutations in protease loci and gene polymorphisms associated with an increased susceptibility to certain diseases. A deeper understanding of these interactions would open the door to the modeling of three-dimensional structures that may help us to design drugs targeting important interactions in various types of cancer and other diseases.

REFERENCES

1. Jett, B.D., M.M. Huycke, and M.S. Gilmore, *Virulence of enterococci*. Clin Microbiol Rev, 1994. **7**(4): p. 462-78.
2. Wang, L., et al., *Relationship of biofilm formation and gelE gene expression in Enterococcus faecalis recovered from root canals in patients requiring endodontic retreatment*. J Endod, 2011. **37**(5): p. 631-6.
3. Souto, R. and A.P. Colombo, *Prevalence of Enterococcus faecalis in subgingival biofilm and saliva of subjects with chronic periodontal infection*. Arch Oral Biol, 2008. **53**(2): p. 155-60.
4. Amyes, S.G., *Enterococci and streptococci*. Int J Antimicrob Agents, 2007. **29 Suppl 3**: p. S43-52.
5. Panesso, D., et al., *Transcriptional analysis of the vanC cluster from Enterococcus gallinarum strains with constitutive and inducible vancomycin resistance*. Antimicrob Agents Chemother, 2005. **49**(3): p. 1060-6.
6. Zheng, B., et al., *Isolation of VanB-type Enterococcus faecalis strains from nosocomial infections: first report of the isolation and identification of the pheromone-responsive plasmids pMG2200, Encoding VanB-type vancomycin resistance and a Bac41-type bacteriocin, and pMG2201, encoding erythromycin resistance and cytolysin (Hly/Bac)*. Antimicrob Agents Chemother, 2009. **53**(2): p. 735-47.
7. Hirt, H., P.M. Schlievert, and G.M. Dunny, *In vivo induction of virulence and antibiotic resistance transfer in Enterococcus faecalis mediated by the sex pheromone-sensing system of pCF10*. Infect Immun, 2002. **70**(2): p. 716-23.
8. Buttaro, B.A., M.H. Antiporta, and G.M. Dunny, *Cell-associated pheromone peptide (cCF10) production and pheromone inhibition in Enterococcus faecalis*. J Bacteriol, 2000. **182**(17): p. 4926-33.
9. Chandler, J.R. and G.M. Dunny, *Characterization of the sequence specificity determinants required for processing and control of sex pheromone by the intramembrane protease Eep and the plasmid-encoded protein PrgY*. J Bacteriol, 2008. **190**(4): p. 1172-83.
10. Clewell, D.B., *Tales of conjugation and sex pheromones: A plasmid and enterococcal odyssey*. Mob Genet Elements, 2011. **1**(1): p. 38-54.

11. Wirth, R., *Sex pheromones and gene transfer in Enterococcus faecalis*. Res Microbiol, 2000. **151**(6): p. 493-6.
12. Murray, B.E., F.Y. An, and D.B. Clewell, *Plasmids and pheromone response of the beta-lactamase producer Streptococcus (Enterococcus) faecalis HH22*. Antimicrob Agents Chemother, 1988. **32**(4): p. 547-51.
13. Chandler, J.R. and G.M. Dunny, *Enterococcal peptide sex pheromones: synthesis and control of biological activity*. Peptides, 2004. **25**(9): p. 1377-88.
14. Chandler, J.R., et al., *Specific control of endogenous cCF10 pheromone by a conserved domain of the pCF10-encoded regulatory protein PrgY in Enterococcus faecalis*. J Bacteriol, 2005. **187**(14): p. 4830-43.
15. Clewell, D.B., et al., *Enterococcal sex pheromone precursors are part of signal sequences for surface lipoproteins*. Mol Microbiol, 2000. **35**(1): p. 246-7.
16. Paulsen, I.T., et al., *Role of mobile DNA in the evolution of vancomycin-resistant Enterococcus faecalis*. Science, 2003. **299**(5615): p. 2071-4.
17. Varahan, S., et al., *An ABC transporter is required for secretion of peptide sex pheromones in Enterococcus faecalis*. MBio, 2014. **5**(5): p. e01726-14.
18. Kao, S.M., et al., *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, 1991. **173**(23): p. 7650-64.
19. Dunny, G.M., et al., *Plasmid transfer in Streptococcus faecalis: production of multiple sex pheromones by recipients*. Plasmid, 1979. **2**(3): p. 454-65.
20. Sava, I.G., E. Heikens, and J. Huebner, *Pathogenesis and immunity in enterococcal infections*. Clin Microbiol Infect, 2010. **16**(6): p. 533-40.
21. Sartingen, S., et al., *Aggregation substance increases adherence and internalization, but not translocation, of Enterococcus faecalis through different intestinal epithelial cells in vitro*. Infect Immun, 2000. **68**(10): p. 6044-7.
22. Olmsted, S.B., et al., *A plasmid-encoded surface protein on Enterococcus faecalis augments its internalization by cultured intestinal epithelial cells*. J Infect Dis, 1994. **170**(6): p. 1549-56.

23. Nakayama, J., et al., *The prgQ gene of the Enterococcus faecalis tetracycline resistance plasmid pCF10 encodes a peptide inhibitor, iCF10*. J Bacteriol, 1994. **176**(23): p. 7405-8.
24. Mori, M., et al., *Structure of cCF10, a peptide sex pheromone which induces conjugative transfer of the Streptococcus faecalis tetracycline resistance plasmid, pCF10*. J Biol Chem, 1988. **263**(28): p. 14574-8.
25. An, F.Y., M.C. Sulavik, and D.B. Clewell, *Identification and characterization of a determinant (eep) on the Enterococcus faecalis chromosome that is involved in production of the peptide sex pheromone cAD1*. J Bacteriol, 1999. **181**(19): p. 5915-21.
26. Dunny, G.M. and D.B. Clewell, *Transmissible toxin (hemolysin) plasmid in Streptococcus faecalis and its mobilization of a noninfectious drug resistance plasmid*. J Bacteriol, 1975. **124**(2): p. 784-90.
27. Searle, B.C., *Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies*. Proteomics, 2010. **10**(6): p. 1265-9.
28. Bazan, J.F., B.T. Macdonald, and X. He, *The TIKI/TraB/PrgY family: a common protease fold for cell signaling from bacteria to metazoa?* Dev Cell, 2013. **25**(3): p. 225-7.
29. Fujimoto, S., et al., *Physical mapping of the conjugative bacteriocin plasmid pPD1 of Enterococcus faecalis and identification of the determinant related to the pheromone response*. J Bacteriol, 1995. **177**(19): p. 5574-81.
30. Fivash, M., E.M. Towler, and R.J. Fisher, *BIAcore for macromolecular interaction*. Curr Opin Biotechnol, 1998. **9**(1): p. 97-101.
31. Fixen, K.R., et al., *Analysis of the amino acid sequence specificity determinants of the enterococcal cCF10 sex pheromone in interactions with the pheromone-sensing machinery*. J Bacteriol, 2007. **189**(4): p. 1399-406.
32. Morar, M., et al., *Mechanism and diversity of the erythromycin esterase family of enzymes*. Biochemistry, 2012. **51**(8): p. 1740-51.
33. Chan, A.C., et al., *Cofacial heme binding is linked to dimerization by a bacterial heme transport protein*. J Mol Biol, 2006. **362**(5): p. 1108-19.
34. Zhang, X., et al., *Characterization of Tiki, a New Family of Wnt-specific Metalloproteases*. J Biol Chem, 2016. **291**(5): p. 2435-43.

35. Clewell, D.B., *Properties of Enterococcus faecalis plasmid pAD1, a member of a widely disseminated family of pheromone-responding, conjugative, virulence elements encoding cytolysin*. *Plasmid*, 2007. **58**(3): p. 205-27.