RICE UNIVERSITY

Evolutionary Trajectories to Daptomycin Resistance in *Enterococcus faecalis*

by

**Corwin Andrew Miller**

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Yousif Shamoo, Professor, Thesis Advisor, Biochemistry & Cell Biology

John S. Olson, Professor, Committee Chair, Biochemistry & Cell Biology

Matthew R. Bennett, Assistant Professor, Biochemistry & Cell Biology

Volker H. W. Rudolf, Associate Professor, Ecology & Evolutionary Biology

Yizhi Jane Tao, Professor, Biochemistry & Cell Biology

HOUSTON, TEXAS November 2013
Abstract

Evolutionary Trajectories to Daptomycin Resistance in Enterococcus faecalis

by

Corwin Andrew Miller

With increasing amounts of hospital-acquired antibiotic resistant infections each year and staggering healthcare costs, there is a clear need for new antimicrobial agents, as well as novel strategies to extend their clinical efficacy. While genomic studies have provided a wealth of information about the alleles associated with adaptation to antibiotics, they do not provide essential information about the relative importance of genomic changes, their order of appearance, or potential epistatic relationships between adaptive changes. In this thesis, I have combined experimental evolution, comparative whole genome sequencing, and allelic frequency measurements to study daptomycin (DAP) resistance in the vancomycin resistant clinical pathogen Enterococcus faecalis strain S613.

Maintaining cells inside a turbidostat, a single polymorphic culture was grown sustaining both planktonic and non-planktonic (e.g. biofilm) populations in co-culture as the concentration of antibiotic was raised, facilitating the development of more ecological complexity than is typically observed in laboratory evolution. This approach revealed a clear order and hierarchy of genetic changes leading to resistance, the signaling and metabolic pathways responsible, and the relative importance of these mutations to the evolution of DAP resistance. Quantitative genetic comparisons between resistant isolates identified convergent evolutionary trajectories, suggesting a common biochemical mechanism of resistance. Phenotypic analysis of resistant strains also revealed
convergent morphological traits associated with resistance, including increased membrane thickness and cell division abnormalities. Increased production of biofilm was also a conserved phenotypic trait, suggesting a possible role for biofilm as a general mechanism of DAP resistance. Despite the relative ecological simplicity of this approach compared to the complexity of the human body, I show that experimental evolution can be used to rapidly identify clinically relevant adaptive molecular pathways and new targets for drug design in pathogens.
Acknowledgements

I would like to first thank all the advisors and professors at Rice who have helped with my project. Members of my committee, John Olson, Matt Bennett, and Jane Tao, have consistently provided insight and encouragement as my project matured. Outside of my committee, my early graduate career would not have been the same without the contributions of Kevin MacKenzie, who would never allow me to get away with not thinking critically; more recently, the involvement of Cesar Arias at UT Medical School has played a vital role in the success of my project. Last but certainly not least, I would like to thank my advisor Yousif Shamoo, who has not merely ensured my project kept progressing, but also that I continued to mature as a scientist. Dr. Shamoo’s advice has consistently prioritized what is best for me as a student during my time at Rice, a mark of a true mentor.

I also would like to acknowledge all the members of my lab, without whom this work would not have been possible. As I joined the Shamoo lab, there was never a shortage of people willing to answer my questions and teach me new techniques despite my bumbling, most notably my senior lab members Matthew Peña, Kasia Walkiewicz, Milya Davlieva, Andrés Benitez, Erol Bakkalbasi, and Sarah Wu. Members who have joined after my arrival have never failed to give helpful insight and make the Shamoo lab a pleasant place to work, and so I would like to thank Kathryn Beabout, Xu Wang, Troy Hammerstrom, Gerda Saxer, and Anisha Perez. I would also like to acknowledge two undergraduate students, Jiayi Kong and Siegfried Hirczy, who played a significant role in the advancement of this project. Outside of the Shamoo lab here at Rice, I cannot go without mentioning Aaron Collier and George Huang, whose comments and conversations helped both my project and sanity.

I would also like to thank the members of the Arias lab across the street at University of Texas Medical School, who have been friends as well as collaborators in the last few years. In particular I would like to thank Cecilia Tran, Diana Panesso, José Muñita, and Lorena Diaz.

Lastly, I would like to thank all my friends and family for their support, both those here in Houston and in Tennessee. In particular, I would like to thank my sister Melanie and my mother Merry. Although he passed away before the completion of this thesis, I would also like to thank my father Barney, for helping to instill a love of science.
Table of Contents

Title.................................................................................................................. 1
Abstract............................................................................................................ 2
Acknowledgements........................................................................................... 4
Table of Contents............................................................................................... 5
List of Figure..................................................................................................... 7
List of Tables.................................................................................................... 8
List of Abbreviations......................................................................................... 9

Chapter 1. Introduction
1.1 Antibiotic resistance as a rising problem................................................. 10
1.2 Rate of antibiotic discovery and drugs of last resort are decreasing...... 11
1.3 Daptomycin structure and mechanism of action.................................. 13
1.4 Enterococci as opportunistic pathogens................................................. 15
1.5 History of E. faecalis strains S613 and R712 in previous studies........... 18
1.6 Resistance-linked mutations observed in enterococci......................... 20
1.7 Goals of this project................................................................................. 22

Chapter 2: Materials and Methods
2.1 Notes on bacterial isolates and growth media...................................... 28
2.2 Flask-based adaptation of E. faecalis strains OG1RF and S613 to daptomycin.. 28
2.3 Turbidostat construction and components.......................................... 29
2.4 Turbidostat operation and adaptation.................................................... 29
2.5 Isolation and screening of TDR strains.................................................. 31
2.6 Comparative whole-genome sequencing.............................................. 32
2.7 Purification of FREQ-Seq barcode adaptor primers............................ 33
2.8 FREQ-Seq analysis of intermediary turbidostat populations............. 35
2.9 Analysis of FREQ-Seq data................................................................. 38
2.10 Biofilm formation assays...................................................................... 39
2.11 Growth rate assays............................................................................... 41
2.12 RNA isolation, cDNA synthesis, and operon mapping...................... 42
2.13 Construction of reporter plasmids....................................................... 43
2.14 Preparation of competent Escherichia coli EC1000 cells.................... 44
2.15 Transformation of E. coli EC1000......................................................... 45
2.16 Preparation of electrocompeent E. faecalis CK111 cells.................... 45
2.17 Transformation of plasmid DNA into Enterococcus faecalis CK111..... 46
2.18 Conjugation of E. faecalis CK111 with E. faecalis S613 sub strains..... 46
2.19 Fluorescent reporter assays................................................................. 47

Chapter 3. Experimental Evolution adapting E. faecalis to DAP
3.1 Introduction.............................................................................................. 49
3.2 Flask based adaptation.......................................................................... 52
3.3 Turbidostat based adaptation.................................................................. 57

Chapter 4. Isolation and Sequence analysis of DAP resistant strains
4.1 Introduction.............................................................................................. 60
4.2 TDR strain isolation and screening....................................................... 62
4.3 Comparative whole genome sequence analysis and putative functions of resistance associated alleles......................................................... 66
4.4 Whole genome sequencing of mixed populations from days 3, 10 and 20.... 76
4.5 Observation of non-canonical contingency loci ........................................ 77

Chapter 5. Allelic Frequency Analysis of Intermediate Populations
5.1 Introduction ......................................................................................... 80
5.2 FREQ-Seq analysis and relative success of mutant alleles ..................... 81
5.3 Inference of strain phylogenies using endpoint genotypes ..................... 85
5.4 Cellular hierarchy of adaptive changes .............................................. 88

Chapter 6. Phenotypic Characterization of TDR strains
6.1 Introduction ......................................................................................... 91
6.2 Increased biofilm production in DAP resistant strains .......................... 94
6.3 Effects of DAP on growth rates ......................................................... 96
6.4 Correlation between biofilm formation and DAP resistance .................. 98
6.5 TEM of Membrane alterations observed in DAP resistant enterococci .... 102

Chapter 7. Conclusions and Future Studies
7.1 Discussion ......................................................................................... 105
7.2 Future work ...................................................................................... 112
Bibliography .......................................................................................... 117
Appendices A-H ...................................................................................... 125
List of Figures

Figure 1.1................................................................................. 12
Figure 1.2................................................................................. 14
Figure 1.3................................................................................. 25
Figure 2.1................................................................................. 37
Figure 3.1................................................................................. 53
Figure 3.2................................................................................. 58
Figure 4.1................................................................................. 70
Figure 4.2................................................................................ 78
Figure 5.1................................................................................. 82
Figure 5.2................................................................................. 86
Figure 5.3................................................................................. 89
Figure 6.1................................................................................. 95
Figure 6.2................................................................................. 97
Figure 6.3................................................................................. 99
Figure 6.4............................................................................... 101
Figure 6.5............................................................................... 102
Figure 6.6............................................................................... 103
Figure 7.1............................................................................... 107
Figure A.1............................................................................... 126
Figure B.1............................................................................... 129
Figure C.1............................................................................... 130
Figure D.1............................................................................... 134
Figure D.2............................................................................... 135
Figure D.3............................................................................... 136
Figure D.4............................................................................... 137
Figure E.1............................................................................... 140
Figure E.2............................................................................... 141
Figure G.1............................................................................... 146
Figure G.2............................................................................... 148
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1</td>
<td>34</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>36</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>54</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>56</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>59</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>63-64</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>67</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>68</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>73</td>
</tr>
<tr>
<td>Table 7.1</td>
<td>109</td>
</tr>
<tr>
<td>Table G.1</td>
<td>144</td>
</tr>
<tr>
<td>Table G.2</td>
<td>147</td>
</tr>
<tr>
<td>Table H.1</td>
<td>149</td>
</tr>
</tbody>
</table>
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. faecalis</em></td>
<td>Enterococcus faecalis</td>
</tr>
<tr>
<td><em>E. faecium</em></td>
<td>Enterococcus faecium</td>
</tr>
<tr>
<td><em>B. subtilis</em></td>
<td>Bacillus subtilis</td>
</tr>
<tr>
<td><em>S. aureus</em></td>
<td>Staphylococcus aureus</td>
</tr>
<tr>
<td><em>S. pneumoniae</em></td>
<td>Streptococcus pneumoniae</td>
</tr>
<tr>
<td>VRE</td>
<td>Vancomycin Resistant Enterococcus</td>
</tr>
<tr>
<td>cls</td>
<td>Cardiolipin synthase</td>
</tr>
<tr>
<td>BHI</td>
<td>Brain Heart Infusion</td>
</tr>
<tr>
<td>LB</td>
<td>Lysogeny Broth</td>
</tr>
<tr>
<td>BEA</td>
<td>Bile-esculin azide</td>
</tr>
<tr>
<td>WGS</td>
<td>Whole Genome Sequencing</td>
</tr>
<tr>
<td>CGS</td>
<td>Comparative Genome Sequencing</td>
</tr>
<tr>
<td>DAP</td>
<td>Daptomycin</td>
</tr>
<tr>
<td>MIC</td>
<td>Minimum Inhibitory Concentration</td>
</tr>
</tbody>
</table>
Chapter 1. Introduction

1.1 Antibiotic resistance as a rising problem. Over the past several decades, antibiotic resistance has increased at an alarming rate, particularly in hospital settings. In 2005, the CDC estimated that methicillin resistant Staphylococcus aureus (MRSA) infections alone led to roughly 19,000 deaths in the United States, accounting for “more… than emphysema, HIV/AIDS, Parkinson's disease, and homicide combined.”

In a more recent estimate spanning 2012, the CDC estimates that antibiotic resistant microbes caused over 2 million illnesses, leading to over 23,000 deaths. The U.S. Department Health & Human Services reports that Healthcare-Associated Infections (HAIs) are the most common complication of a hospital stay and one of the top ten leading causes of death in the United States.

A group of six organisms, collectively known as the ‘ESKAPE’ pathogens, have been singled out by the Infectious Disease Society of America as organisms against which new antibiotics are urgently needed, and have similarly been labeled a ‘Serious Hazard’ by the Center for Disease Control. Responsible for the majority of nosocomial infections, the six ‘ESKAPE’ organisms Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumonia, Acinetobacter baumannii, Pseudomonas aeruginosa, and Enterobacter species frequently “escape” current antibacterial therapy, making them especially challenging to the healthcare system. Vancomycin-resistant enterococci (VRE) are of particular concern due to lack of reliable bactericidal therapeutic options, with over 20,000 estimated cases in 2012 leading to over 1,300 deaths. As a consequence, the cyclic lipopeptide daptomycin (DAP) is often used ‘off-label’ as an antibiotic of last resort for enterococci and other drug resistant Gram-positive infections,
as resistance rates are currently low. Unfortunately, cases of DAP resistance in gram-positive bacteria, including \textit{E. faecium} and \textit{E. faecalis}, have already been reported clinically.\textsuperscript{3,9-11}

\textbf{1.2 Rate of antibiotic discovery and drugs of last resort are decreasing.}

Since their discovery, antibiotics have played a remarkable role in revolutionizing medicine. Apart from simply treating existing bacterial infections, antibiotics allow prevention of secondary infections and thus allow more invasive surgery options and immune-compromising procedures (such as chemotherapy). Despite their remarkable utility, development of novel antibiotics has been experiencing market failure, as the rate of new antibiotics being developed has progressively slowed over the last thirty years. \textit{Figure 1.1}, from a publication by R. Guidos and the Infectious Disease Society of America (IDSA)\textsuperscript{2}, demonstrates this developmental stagnation by showing the number of new antibiotics approved by the US FDA in 5 year increments since 1983.
As Figure 1.1 above demonstrates, if the decreasing rate of antibiotic development persists, within the next decade very few new antibiotics can be expected to reach the market. Antibiotic resistance in microbes, however, will continue to increase as long as antibiotics are being prescribed, as bacteria continually evolve resistance. Unlike most drug products, antibiotics are not indefinitely useful once developed but rather represent a more finite commodity, requiring continuous development of new drugs. As scientists and physicians are increasingly having difficulty keeping pace with rapidly evolving bacterial populations, there is a pressing need for novel drug development strategies.
**1.3 Daptomycin structure and mechanism of action.** Daptomycin (DAP), a cyclic lipopeptide antibiotic, was first discovered by Eli Lilly and Company (Lilly) in the early 1980’s \(^\text{12}\). Like many antibiotics, Daptomycin is a fermentation product of soil bacteria *Streptomyces*, identified in species *Streptomyces roseosporus*. Bactericidal against a range of Gram-positive organisms, a twice-daily antibiotic treatment regimen was developed by Lilly and brought through Phase 2 clinical trials in the late 1980’s. However, disappointed by the drug’s performance, Lilly licensed DAP to Cubist Pharmaceuticals Inc. (Cubist) in 1997 \(^\text{12}\). Considered the latest drug developed within a broader class of cyclic lipopeptide antibiotics \(^\text{13}\), Daptomycin was approved by the FDA in 2003 for treatment of Gram-positive skin infections, *Staphylococcus aureus* (*S. aureus*) bacteraemia, and right-sided *S. aureus* endocarditis \(^\text{14}\), although it is often prescribed for enterococcal endocarditis as well. Significantly, Daptomycin has been shown to have bactericidal activity against both VRE and methicillin-resistant *Staphylococcus* (MRSA)\(^\text{15,16}\), establishing DAP as a popular drug of last resort.

While a peptide, Daptomycin contains a number of noncanonical structural features, shown below in *Figure 1.2*. 

---

832
Figure 1.2. Structure of Cyclic Lipopeptide Antibiotic Daptomycin. Figure above reproduced from 17.

Shown in Figure 1.2, Daptomycin contains thirteen amino acids in all, ten of which form a cyclic ring. Included among the thirteen amino acids are three non-proteinogenic residues: L-kynurenine (Kyn), L-3-methylglutamic acid (mGlu), and L-ornithine. Additionally, two amino acids are present in their atypical dexrotatory enantiomers (D-serine and D-alanine). On the right of Figure 1.2, medium chain fatty acid decanoic acid can be seen branching off of an exocyclic tryptophan residue. This decanoic acid moiety, combined with the net positive charge of the cyclic domain, likely accounts for the detergent properties of Daptomycin, which is thought to lead to binding of pulmonary surfactant thereby preventing any capacity to treat pneumonial infections 18.

Though prescribed since 2003, many details of the mechanism of Daptomycin have only recently come to light. Shown to require calcium for bactericidal activity, spectral studies indicate that Daptomycin forms a complex with calcium cations before
entering its active conformation. Following formation of a Ca\(^{2+}\)-DAP complex, FRET data suggests that Daptomycin oligomerizes on the membrane surface forming a 6-7 member complex on a single membrane leaflet. Fluorescent studies using unilamellar liposomes observed that DAP undergoes a structural transition in the presence of membrane lipid phosphatidylglycerol (PG), making this lipid the most likely binding target on the membrane. As DAP resistance is associated with decreased phosphatidylglycerol (PG) content in Bacillus subtilis, S. aureus, and E. faecalis, in vivo data is in good agreement with studies in model membranes showing that PG is required for cell surface binding. Following oligomerization and membrane surface binding, fluorescent spectroscopy studies on model membranes suggest that the complex then inserts into the membrane, a process perhaps abetted by an increase in amphipathicity upon binding of Ca\(^{2+}\). After membrane insertion, fluorimetric and flow cytometry data show a loss of membrane potential, an event correlated with both loss of potassium ions and cell viability. Taken together, the above studies suggest that Ca\(^{2+}\)-dependent DAP oligomers likely form membrane lesions which lead to loss of membrane potential and eventual cell death, similar to other bactericidal antibiotics.

This mechanism of cell death induced through loss of potassium, and thereby membrane potential, is similar to that of previously characterized antibiotic valinomycin, gramicidin, and nonactin. However, the data described above suggests a distinct mechanism of PG-dependent membrane insertion in the case of Daptomycin.

1.4 Enterococci as opportunistic pathogens. Enterococci are a genus of gram-positive bacteria most commonly found as commensals in insects and mammals,
including humans. A robust organism, enterococci are non-motile facultative anaerobes capable of withstanding a wide range of temperature (10-45°C), pH (4.5-10), and salinity capable of withstanding a wide range of temperature (10-45°C), pH (4.5-10), and salinity

28. Although non-sporulating, enterococci do form biofilms which likely play a role in their hardiness. Typically found in the digestive track, two enterococcal species predominate in the intestines of humans: *Enterococcus faecalis* (90-95%) and *Enterococcus faecium* (5-10%), although other species can be found under rare circumstances

29.

Though usually acting as benign gut commensals in humans, enterococci can act as an opportunistic pathogen in other areas of the body. Aided by numerous surface pili which facilitate attachment to host surfaces, enterococci are a well known cause of endocarditis, bacteremia, and other tissue infections

30. *E. faecalis* is considered the most virulent among enterococci species, frequently causing of urinary tract infections, which are, in turn, the most common form of nosocomial infection

31. *E. faecalis* infections are also problematic in dental surgeries, having been identified as the most common bacteria isolated from root-filled teeth with apical periodontitis (infection in the pulp of the tooth following a root-canal)

32. Although *E. faecalis* infections are more common than those from *E. faecium*, often times *E. faecium* can be more difficult to treat, with over 60% of clinical isolates carrying vancomycin resistance in 2003

33.

Antibiotic resistance within *Enterococcus* has been a rising problem in the last decade, particularly among strains resistant to vancomycin (vancomycin resistant *Enterococcus* abbreviated VRE). According to the Center for Disease Control, enterococci accounted for approximately one in every eight infections in hospitals between 2006 and 2007, with about 30% of these arising from vancomycin resistant
strains. Since first being identified in the United States in 1988, VRE has spread remarkably quickly, arising as a leading cause of nosocomial infections. Two of the most significant factors for such rapid spread of VRE include presence of extraordinarily mobile genetic elements among resistant strains, and appearance of multiple independent modes of resistance. As vancomycin functions through inhibiting cell wall synthesis by binding D-alanine-D-alanine dimers, resistance can evolve through several mechanisms, frequently involving alteration of membrane synthesis to replace D-alanine-D-alanine within the membrane.

To address growing medical concerns, enterococci have increasingly become the subject of study. Among virulent isolates, vancomycin resistant strain V583 has been the subject of the most study, notably observing that more than a quarter of its genome is comprised of mobile DNA elements, more than any other organism to date. Presence of virulence factors and numerous antibiotic resistance determinants, however, make strain V583 difficult to use in the laboratory and manipulate genetically. As a consequence, most laboratories study strain OG1 or derivatives of this strain, most notably OG1RF. A derivative of strain OG1 with rifampicin and fusidic acid resistance, OG1RF lacks resistance to other antibiotics, does not carry plasmids, is readily transformable, and possesses virulence in model rodent and C. elegans organisms. Whole genome sequencing of strain OG1RF revealed that while the genomes of strain OG1RF and V583 aligned syntenically (same ordering of genes with respect to the chromosome), OG1RF contained numerous genomic differences, most notably a stark lack of mobile elements and the presence of two CRISPR loci. As this comparison illustrates, despite sharing a common species,
enterococcal strains can vary considerably. Indeed, another *E. faecalis* strain Symbioflor 1 is not only non-pathogenic, but is sold as a probiotic supplement by SymbioPharm Ltd. Initially isolated from a stool sample, whole genome sequence analysis suggests Symbioflor 1 can likely trace its lack of pathogenicity to large deletions in virulence factors. As the diverse nature of strains V583, OG1RF, and Symbioflor 1 demonstrate, newly isolated strains of *E. faecalis* are likely to have variable genetic and phenotypic properties, highlighting the need to continually study new strains as they are identified in the clinic.

### 1.5 History of *E. faecalis* strains S613 and R712 in previous studies.

Used as a drug of last resort, most bacteria remain susceptible to DAP. However, in 2005 the first instance of observed clinical resistance to daptomycin in *E. faecalis* was reported by Munoz-Price *et al.* In this correspondence, a patient history is presented for a case of bacteremia, arising from vancomycin-resistant *E. faecalis* (VRE). Although the VRE strain was susceptible to β-lactam antibiotics such as ampicillin, the patient was allergic to penicillin derivatives. Thus, the patient was initially prescribed linezolid, however after development of thrombocytopenia linezolid therapy was discontinued. Following unsuccessful antibiotic treatment with oral ciprofloxacin, the patient was prescribed daptomycin and appeared to have a positive response. The patient was then discharged with a prescribed regimen of daptomycin and amikacin (another antibiotic targeting gram-negative organisms), however the infection recurred and the patient returned to the hospital after two weeks, again presenting with severe bacteremia. After blood cultures
again tested positive for *E. faecalis*, the patient began ampicillin desensitization treatment while treatment with linezolid was resumed.

Unfortunately, the patient died shortly thereafter; however, throughout the course of infection, multiple *E. faecalis* isolates were obtained from the patient. Most significant to our studies are two such isolates: an DAP-susceptible strain isolated immediately prior to daptomycin treatment in the patient, designated strain S613, and a DAP-resistant strain isolated following failure of DAP therapy, designated strain R712.

In the lab of Dr. Cesar Arias, an investigator at the University of Texas Medical School, the genomes of both strains S613 and R712 were sequenced and compared to one another in order to gain a better understanding of the resistance process on a genetic level. The Arias lab identified three genes containing mutations linked to daptomycin resistance in strain R712, identified in genes *liaF*, *cardiolipin synthase (cls)*, and *glycerophosphoryl-diester-phosphodiesterase (gdpD)*.45

Though the association between the mutated genes and DAP resistance is not immediately apparent, each of the altered proteins are involved with cell envelope or membrane chemistry on some level. GdpD and Cls are involved in membrane biosynthesis, with GdpD hydrolyzing glycerophosphodiesters while Cls catalyzes the dimerization of two phosphotidylglycerol (PG) molecules to form cardiolipin. LiaF, in turn, is part of the *liaFSR* three-component regulatory system, which regulates cell envelope stress response across many organisms, including *Bacillus subtilis* and *Staphylococcus aureus* (known as VraSR system in staphylococci).

Subsequent analysis of the membrane composition of strains S613 and R712 revealed multiple changes in phospholipid content associated with DAP resistance, most
notably a decrease in phosphatidylglycerol (PG) \(^{25}\) analogous to a similar observation previously noted in \(B.\ subtilis\) \(^{23}\) (see section 6.1 for detailed discussion). While the mutations previously identified in strain R712 lead to this phenotype, the contribution and significance of each individual mutation has yet to be determined. Furthermore, while this study has focused on one path to daptomycin resistance identified in a single patient, there are likely other pathways to daptomycin resistance yet to be described. By studying adaptation \textit{in vitro}, we can more thoroughly explore this adaptive landscape without the need for an infected patient.

1.6 \textbf{Resistance-linked mutations observed in enterococci.} Though relatively few DAP resistant \(E.\ faecalis\) variants have been isolated to date, several have already been studied genetically. In 2012, Palmer \textit{et al.} described use of a serial flask transfer procedure (analogous to the procedure described in section 2.2) to adapt \(E.\ faecalis\) strain V583 to Daptomycin over roughly two weeks \(^{46}\). After isolating three resistant variants (labeled DAP-A, DAP-B, and DAP-C), Palmer \textit{et al.} used comparative whole genome sequencing to identify resistance-linked mutations. Among the several mutant alleles identified, strains DAP-B and DAP-C were found to have a mutation in \(yvlB\) homologue EF1753 (shown in this study to be regulated by the \(liaFSR\) system, see Appendix B), and all three strains were found to carry a mutation in cardiolipin synthase, similar to findings by Arias \textit{et al.} \(^{45}\). Palmer \textit{et al.} went on to postulate that mutations in \(cls\) alone were sufficient for resistance, demonstrating that insertion of a plasmid containing mutant gene \(cls^{R218Q}\) into \(E.\ faecalis\) strain OG1RF resulted in a DAP resistant phenotype. However, it is important to note that this \(cls\) allele was expressed from a multi-copy plasmid in a
different genomic context (strain OG1RF rather than V583), hence other mutations identified may still be required for resistance in strain V583

In a second study on DAP resistance in enterococci, Humphries et al. used DAP resistant *E. faecium* strains isolated from the blood of a patient with Vancomycin-resistant bacteremia. Initially, the patient was treated for twenty-six days with Daptomycin, after which highly DAP resistant strain 5938 was isolated. Four months after discontinuation of DAP therapy, DAP-susceptible strain 8019 was isolated, and after eight months moderately DAP resistant strain 5994 was subsequently isolated. Using comparative whole genome sequencing, Humphries et al. identified several mutations associated with DAP resistance, among them mutations in *cls* in both resistant strains (5938 and 5994) and a mutation in gene *pspC* (shown in this study to be on an operon regulated by the *liaFSR* system, see Appendix B) in moderately resistant strain 5994.

Following identification of DAP resistance-associated alleles in *E. faecalis* strain R712, Arias et al. sequenced resistance associated alleles *liaF, liaS, liaR,* and *cls* in clinically isolated DAP-susceptible *E. faecium* strain S447, as well as six clinically isolated resistant strains (including strain R446, the pair of strain S447 as they were each isolated from the same patient). Relative to ancestral strain S447, four of the six strains carried a mutation in the *liaFSR* system, and five carried a mutation in *cls*. In a subsequent screen of both DAP resistant and susceptible clinically isolated *E. faecium* strains, it was also observed that a majority of resistant isolates had mutations in the *liaFSR* system. In depth analysis of *E. faecium* strain R446, a resistant strain without a mutation in the *liaFSR* system, found that resistance evolved in this isolate primarily
through mutations in the *yycFG* two-component signaling system, with *cls* mutations not playing a significant role \(^{48}\).

Comparing the studies discussed above, mutations in the *liaFSR* system and *cls* appear to be consistently associated with DAP resistance. This association not only spans related species *E. faecalis* \(^{45}\) and *E. faecium* \(^{11}\), but both clinically \(^{47}\) and laboratory \(^{46}\) isolated strains. As *E. faecium* strain R446 demonstrates however, enterococci can evolve DAP resistance through different routes such as the *yycFG* system. Thus, it is key to confirm if *liaFSR* and *cls* mutations mediate the primary route to resistance and determine if other significant evolutionary trajectories remain uncharacterized.

### 1.7 Goals of this project

The primary goal of this project was to develop a molecular understanding of Daptomycin resistance in enterococci, elucidating mechanisms of resistance on a biochemical and cellular level. To meet this goal, this thesis had three primary aims, which together have led to a new understanding of the evolutionary trajectories that lead to DAP resistance. The first aim was to use experimental evolution to evolve DAP resistance in the laboratory, accomplished in this thesis by using a turbidostat to evolve a polymorphic culture containing numerous highly resistance *E. faecalis* strains. The second aim was to identify genes associated with resistance, accomplished using comparative whole genome sequencing. Lastly, the third aim was to quantitatively assess the evolutionary success of each mutant allele and elucidate the cellular mechanisms of genes critical to resistance, an aim accomplished through quantification of intermediary allelic frequencies throughout the course of adaptation followed by bioinformatic analysis.
To address the first aim of this thesis, *Enterococcus faecalis* strain S613 (discussed in section 1.5) was selected for laboratory adaptation. As a clinically isolated VRE strain, data gleaned from the evolution of strain S613 stands to be more medically relevant than data from a laboratory strain such as O1GRF. Additionally, as part of a sequenced clinical strain pair, strain S613 provides a unique opportunity to compare evolution of DAP resistant in a patient versus the laboratory. However, unlike strain pairs isolated in the clinic, evolving cells in a laboratory environment could enable analysis of drug resistance in a manner that is predictive rather than reactive, identifying resistance determinants before they become a health concern.

For the laboratory adaptation process, a continuous culture turbidostat was used to perform experimental evolution. Using a modified bioreactor (see sections 2.3 and 3.4), continuous evolution is markedly different from the established flask-transfer methods that have revolutionized experimental evolution\(^\text{10,46,49-58}\). By growing cells in a turbidostat, a device maintaining constant population size, cells are able to grow exceedingly rapidly as they are held in constant exponential growth phase. More importantly, previous studies have found that turbidostats are able to culture a polymorphic population\(^\text{58-60}\), as well as maintain both planktonic and biofilm populations in co-culture\(^\text{61}\).

Maintenance of a polymorphic population is key to addressing the second aim of this thesis: identifying genes and pathways associated with DAP resistance. By analyzing multiple cell lines that have evolved resistance independently, the degree of divergence or convergence between lines will inform on the degree of constraint within the process of DAP adaptation, while also having a greater chance of identifying novel
trajectories to resistance. In this thesis, I show that such an analysis can be performed by isolating individual strains from a DAP-resistant population, screening for distinct genotypes using previously identified genetic markers \(^4^5\), and using comparative whole genome sequencing to identify mutations in each strain.

As mentioned above, the third aim of this thesis entails measuring allelic frequencies of resistance-associated alleles throughout the course of adaptation. While observing convergence within a particular gene or pathway between cells adapting to the same stressor does provide an indication of its significance, sequences from a relatively large number of strains are required to make such an observation quantitative. As isolating and whole-genome sequencing a large number of strains was not a tractable endeavor at the time these experiments were performed, in this thesis I will discuss use of a next-generation sequencing technique known as FREQ-Seq \(^6^2\) to measure frequencies of resistance alleles (identified as part of the second goal) within DNA samples purified from mixed populations throughout the course of adaptation. By measuring allelic frequencies, one can generate a quantitative metric of evolutionary success, thereby assessing the significance of each resistance-associated allele allowing them to be prioritized for further study. Measuring allelic frequencies at intermediary time-points will inform on the order of mutations arising within the population, facilitating a detailed look at the evolutionary dynamics within an evolving population.

As identifying resistance-associated alleles is dependent upon first evolving a resistant population, and measuring allelic frequencies is dependent upon first identifying resistance-associated alleles, the objectives described above comprise an interwoven process, termed quantitative experimental evolution. This process, as applied in this
work to *E. faecalis* strain S613, is depicted graphically below in Figure 1.3, and is painted in contrast to evolution of resistance within a patient.

![Graphical representation of evolved resistance and antibiotic therapy](image)

**Figure 1.3. Quantitative Experimental Evolution Process.** Quantitative experimental evolution and validation against clinically derived strain *E. faecalis* R712. (A) Arias et al., (2011) used comparative whole-genomic sequencing of clinical strain pair S613 and R712 to show that mutations to liaF, cls and gdpD were sufficient to confer DAP resistance. (B) Quantitative experimental evolution combines experimental evolution, genomic sequencing and measurements of allelic frequencies to identify and deconstruct the adaptive network of genomic changes responsible for resistance. A turbidostat is used for experimental evolution, maintaining a single polymorphic population in exponential growth with the unique capacity to culture both planktonic and biofilm forming cells (Box 1). Phenotypic and genotypic cluster analysis is then used on numerous end-point strains isolated from the turbidostat (Box 2) to identify genetically
distinct strains for comparative whole genome sequencing (Box 3). Allelic frequencies of mutant genes identified in these strains are then quantified across numerous intermediary populations using DNA-barcoding technique FREQ-Seq (Box 4). Intermediary frequency measurements allow us to identify the most successful alleles, the order of their appearance in the population, and infer their associated mechanisms.

I show in this work that use of quantitative experimental evolution provides a robust and scalable approach to elucidate genes and molecular pathways conferring resistance in clinical pathogens, without the need for infection in a patient. There is a common expectation that the evolution of antibiotic resistance in a patient is ‘too complicated’ to be recapitulated in vitro with any accuracy. While pathogenesis will likely hinge on the complex environment within a patient and interactions between the pathogen and the patient, the biochemical foundations of antibiotic resistance need not. As quantitative experimental evolution only requires that an organism be culturable, it is particularly useful for organisms such as *E. faecalis* S613 for which molecular genetics is challenging.

Identification of the most significant alleles associated with resistance has a number of applications. Genes identified as primary resistance determinants provide new protein targets for structural or biochemical studies, which allow a better understanding of mechanisms of resistance on a molecular level. In turn, studying phenotypes of resistant strains allows helps clarify mechanisms of resistance-associated alleles on a cellular level, as is shown in Chapter 6 of this thesis. Genes involved in resistance can also be used as targets for rational drug design, as drugs disabling systems involved in
resistance could be used as anti-evolution adjuvants to prolong the efficacy of DAP. In a similar manner, DAP resistant *E. faecalis* isolates could also be used to test new variants of DAP to ensure their efficacy as resistance spreads. By demonstrating the efficacy of this approach, it is my hope that use of evolution experiments as a tool for prioritizing adaptive genetic changes will be applied to other pathogens, furthering the fields of biology and medicine by improving understanding of antibiotic resistance and providing targets for future drug design.
Chapter 2: Materials and Methods

2.1 Notes on bacterial isolates and growth media. *E. faecalis* strains were grown in a broth mixture comprised of 20% brain-heart infusion (BHI) and 80% lysogeny broth (LB), abbreviated LBHI. Isolation of clinical *E. faecalis* strains S613 and R712 has been described previously by Arias *et al* \(^44\), as well as their whole genome sequencing and characterization \(^10\). Isolation of *E. faecalis* strain OG1RF has also been previously described \(^38\).

2.2 Flask-based adaptation of *E. faecalis* strains OG1RF and S613 to daptomycin. Flask based adaptation experiments were conducted through twice daily serial passage through increasing sublethal doses of daptomycin. Cultures were passaged in 50 ml aliquots of LBHI media, inoculating with 50 µl culture after 12 hours of growth. Drug concentration was kept constant for a minimum of 24 hours (two passages); after 12 hours of growth at a given concentration, broth macrodilution MIC testing (5 ml test tubes) \(^63\) was used to determine the highest tolerable daptomycin concentration for the following day’s growth. This process was continued for each strain, ending at a growth concentration of 20 µg/ml daptomycin. For adaptations done in the presence of tetracycline, each flask of media was prepared with 25 µg/ml tetracycline.

For solid-surface based adaptations, a razor blade was suspended from the top of each flask and was allowed to move freely while submersed in the liquid. Instead of inoculating each subsequent flask with 50 µl of culture, the suspended razor was transferred to the next flask.
2.3 Turbidostat construction and components. To construct a turbidostat, a Sartorius Stedim Biostat Bplus bioreactor was used as a growth chamber. Media was sterilized using an INTEGRA Mediaclave 10, and was fed to the growth chamber from 9L glass carboys, drawn using peristaltic pumps built into the bioreactor control unit. Excess media was drawn from the growth chamber in a similar fashion and flushed into a large reservoir containing an excess of bleach. Before entering the growth chamber, media was filtered through a Millipore Opticap XL 2 filter (0.22 µm, catalogue # KVGLA02FF3) to ensure sterility.

2.4 Turbidostat operation and adaptation. The turbidostat described in Section 2.2 was inoculated from a single colony of *E. faecalis* strain S613 and maintained without DAP to allow cells to equilibrate turbidostat growth conditions. Cells were grown at steady-state conditions in LBHI media (250 ml culture volume, 37°C, 50 mg/ml Ca\(^{2+}\), air flow of 0.16-0.2 L/min). A constant optical density of OD\(_{600}\) = 0.3 (~3×10\(^8\) cells/ml) was maintained by automated dilution of the culture from a series of feed reservoirs. Within 12 hours following inoculation into the turbidostat, biofilms started to form in the vessel. Once biofilms formed, the total population size could no longer be measured accurately through automation using light scattering and therefore dilution control parameters were adjusted manually for the entirety of the experiment.

A turbidostat-adapted strain of S613 was produced by continuous culture in the turbidostat without antibiotic for 6 days 16.5 hrs. No endpoint genomic changes were identified following pre-culturing, suggesting that selection in the LBHI media under these conditions was very weak. Cells from the turbidostat-adapted population were plated...
on solid LBHI media, and a single colony was isolated and used to reinoculate the turbidostat. DAP concentration was raised to 20 µg/ml in seven intervals across approximately 24 days of continuous culture (*Figure 3.2*). A polymorphic population was maintained by increasing antibiotic concentration in empirically determined steps during the course of the adaptation. To measure the MIC of the population, we collected a sample of the turbidostat population every day and diluted it 100-fold into test tubes containing 10 ml of fresh LBHI media and varying concentrations of DAP. Growth rate of the sample was determined as a function of DAP concentration in a 10 ml LBHI liquid culture over approximately 12 hours. If the growth rate in a test population with elevated drug concentration matched the growth rate at the DAP concentration currently in use, then the DAP concentration in the bioreactor was increased. To ensure that contaminants were not present during continuous culture, samples from the turbidostat were plated each day on both LBHI and bile esculin agar (BEA), and colony morphology examined.

On the 10th day of the experiment, after changing the 0.2 µm filter through which all the media was pumped, a contaminant was observed coming out of the growth vessel. At this point the run was halted, and the growth vessel was then sterilized and refilled with media at the same DAP concentration (0.75 µg/ml) containing 20 µg/ml gentamycin. The growth vessel was then inoculated using a glycerol stock (thawed on ice) collected prior to changing the 0.2 µm filter. After the reinoculated culture reached an OD of 0.3, media inflow was reactivated, connected to a media tank that contained 0.75 µg/ml DAP, but no gentamycin, thereby rapidly flushing the gentamycin out of the bioreactor shortly thereafter. As a broad-spectrum antibiotic, gentamycin will kill most contaminating
bacteria, however *E. faecalis* is naturally resistant to it, thus a brief passage through gentamycin should kill contaminants and not interfere with the adaptation experiment.

Population samples were collected every 3-6 hours, supplemented with glycerol and stored at -80°C. Samples were of a relatively large size (roughly $5 \times 10^9$ cells), representing approximately 5% of the total population in the turbidostat. The biofilm was present as particulate material that was continuously sloughed off the vessel surfaces due to the mechanical shear forces produced by rapid stirring and was collected readily.

As mentioned above, a constant optical density was held within the turbidostat throughout the course of adaptation. To verify that the correlation between optical density and cellular density did not change as bacteria evolved DAP resistance, colony-forming units (CFUs) were measured using freezer stocks from five days over the course of the experiment (days 2, 6, 13, 17, and 22). These mixed-population freezer samples were thawed on ice, and OD$_{600}$ was confirmed to be approximately 0.3. Cells were serially diluted $10^3$, $10^4$, and $10^5$ fold, and 50 µl of each dilution were plated in triplicate on LBHI agar. The five time points were found to have an average CFU/(OD$_{600} \times$ ml) of $1.15 \times 10^9$, with values ranging from $2.72 \times 10^8$ (day 17 sample) to $3.14 \times 10^9$ (day 6 sample) and an average standard deviation of $3.17 \times 10^8$ CFU/(OD$_{600} \times$ ml) and showed no significant change over time (slope = -0.03, $F_{1,3} = 1.88$, $p = 0.26$).

### 2.5 Isolation and screening of TDR strains

Cells from the final day of the Turbidostat selection experiment (at a bioreactor DAP concentration of 20 µg/ml) were plated onto solid LBHI media. 50 individual colonies were isolated and designated TDR for Turbidostat-derived DAP Resistant colony (TDR 1-18, 19.1, 19.2, 20-48). To identify
novel end-point strains for whole genome sequencing (WGS), we determined the MICDAP using agar microdilution and sequenced liaF and cls, two genes previously linked to DAP resistance. From the 50 colonies, seven distinct DAP resistant colonies were selected for WGS (TDR 4, 7, 8, 13, 19, 22, 28). Genomic DNA was isolated using the UltraClean Microbial DNA Isolation Kit (MO BIO, Carlsbad, CA) following the manufacturer’s protocols, adding an additional lysis step using lysozyme (0.5mg/ml lysate) and mutanolysin (0.25µg/ml lysate).

2.6 Comparative whole-genome sequencing. Genomes of TDR strains mentioned above were sequenced as paired-end reads on an Illumina HiSeq sequencer (SeqWright, Houston, TX). The reads were aligned to the reference sequence of S613 (gff3 file dated 11/12/ 2010, downloaded from PATRIC at http://www.patricbrc.org) using Seqman NGen by DNAStar Inc as well as BWA and SAMtools. The average coverage across the seven clones was 496X. Mutation detection was performed using Seqman (DNAStar) and the BWA alignments (based on 90% agreement among reads for SNPs and 50% agreement for indels), and all reported mutations were confirmed using Sanger sequencing.

After closer investigation of genomic sequencing data, TDR19 turned out to be a mixture of three genetically distinct colonies that were renamed TDR19.1, TDR19.2, and TDR19.3. The last one, TDR19.3 was genetically identical to strain TDR8, and thus is not its own distinct strain. The frequencies of individual mutations in this mixed sample were compared to the results from Sanger sequencing of these loci in multiple individual colonies isolated from the mixed TDR19 frozen stock to get both linkage and frequency
information. Mutations identified in each of these eight distinct strains are listed in *Table 4.2*. Genomes of TDR strains are currently in submission to NCBI under submission number SUB157235.

2.7 Purification of FREQ-Seq barcode adaptor primers. Purification of barcode sequencing primers was begun using a 96-well puncture plate generously provided by the Marx lab 62. Culture in each well was grown overnight at 37°C in LB broth, and minipreps were used to purify plasmid. On each plasmid, a unique barcode identifier was located within string of nucleotides, which read as follows:

“AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTC
CGATCTNNNNNGTAAAACGACGGCCAGT”, where the NNNNNN is a unique sequence of nucleotides (the “barcode”) for each of the 96 wells, and the preceding nucleotides are an Illumina adaptor sequence.

Following plasmid purification, PCR was used to amplify the barcode adaptor primer region using the steps listed below. To amplify the adapter sequence from the plasmid, a PCR reaction was run using the following reaction mixture::10 µL Buffer (Phusion), 2 µL Template (plasmid), 0.5 µL Polymerase (Phusion), 1 µL dNTP, 0.5 µL Illumina Reverse Primer (sequence below), 0.5 µL Illumina Forward Primer (sequence below), and 35.5 µL ddH2O. The primer sequences are: Forward Primer: 5’ – AATGATACGGCGACCAC (Illumina Forward) and Reverse Primer: 5’ – ACTGGCCGTCGT TTTTAC. Cycling conditions are shown below in *Table 2.1.*
Table 2.1. Cycling conditions for Amplification of FREQ-Seq barcode primers

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
<th>Repetitions</th>
<th>Notes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>98 C</td>
<td>120 s</td>
<td>1x</td>
<td>The 53 C step has a temperature decrease of .2 C per cycle.</td>
</tr>
<tr>
<td>98 C</td>
<td>60 s</td>
<td>40x</td>
<td></td>
</tr>
<tr>
<td>53 C</td>
<td>30 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72 C</td>
<td>5 s</td>
<td>1x</td>
<td></td>
</tr>
<tr>
<td>72 C</td>
<td>10 s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 C</td>
<td>Hold</td>
<td>1x</td>
<td></td>
</tr>
</tbody>
</table>

The resulting PCR 81 base pair product was then purified using acrylamide gel extraction, as product is too small for PCR clean-up purification, and too similar in size to Illumina_F and Illumina_R primers for agarose gel purification. The PCR product was run on a 1.5% Acrylamide gel using a large size well. The band was cut out and placed into dialysis tubing, taking care to use tubing with a small enough molecular weight cut off such that the adapters do not diffuse out of the bag. Inside the dialysis bag, 200 µL of PBS buffer was added, and placed into PBS solution inside an agarose gel electrophoresis machine box, taking care to orient the dialysis bag so the clips are not blocking the current through the bag/gel, with the clips on the end of the dialysis bag perpendicular to the top of the gel and the dialysis bag parallel to the top. The sample was run at 100 V for 30 minutes, confirming that no bands remains in gel using UV light following electrophoresis. At this point, the purified primer product was dissolved in the buffer inside the dialysis bag.
The primer was then purified by ethanol precipitation. Fluid was removed from the dialysis bag and placed into a microcentrifuge tube. Ethanol was added at a volume 2.5 times greater than that of the sample, and sodium acetate was added to a final concentration of 0.2 M. Tubes were immersed tubes in ethanol in a Styrofoam container, taking care to label each tube with a wax pencil. Using a thermocouple to measure temperature, just enough dry ice was added to reach -40°C, at which point the contents of the eppendorf tube were viscous, but not solid. Tubes were incubated for one hour at -40°C, adding more dry ice as necessary to maintain temperature. When the product froze, it was allowed to warm back to a viscous liquid, and the timer on the incubation was reset. After incubating, the tubes were immediately centrifuged using a table top centrifuge at 14,000 rpm for 30 minutes inside the cold room. The supernatant was removed by pipetting, and the tubes were lyophilized until completely dry. Purified adaptor primer product was resuspended in 20 µL ddH₂O.

2.8 FREQ-Seq analysis of intermediary turbidostat populations. Following purification of FREQ-Seq adaptor primers, FREQ-Seq constructs were assembled using a series of PCR reactions to assess frequencies of mutant alleles. In the first PCR reaction, the gene containing a mutation of interest was amplified from a genomic DNA sample containing a heterogenous population. Cycling conditions were optimized for each sample. The product of this reaction was then purified using either gel extraction or PCR clean-up.

In the second PCR reaction, a short region (~150 nucleotides) surrounding the mutation site was amplified. In addition to overlapping the template DNA, the forward
primer began with the M13 sequence and the reverse primer began with the reverse illumina sequence: M13 sequence: 5’ – GTAAAACGACGGCCAGT… Illumina Reverse: 5’ – CAAGCAGAAGACGGCATA. Cycling conditions were optimized for each sample. The product of this reaction was then purified using PCR clean up using 96-well plates, taking care to wash wells with buffer PB as recommended by manufacturer.

The third PCR attached the template described above to the previously purified adapter primer (described in section 2.7). This reaction required three priming components, and required the same reaction mixture and cycling conditions for all samples. The reaction used the following mixture: 10 µL Buffer (Phusion), 2.5 µL template (PCR product described in proceeding paragraph), 2.5 µL adapter DNA (described in section 2.7), 1 µL Polymerase (Phusion), 1 µL dNTP (10 mM), 0.5 µL Reverse Primer (Illumina Reverse, 25 µM), 0.5 µL Forward Primer (Illumina Forward, 25 µM), and 32 µL ddH₂O. Cycling conditions did not require optimization as templating regions do not vary from sample to sample; optimum cycling conditions are shown below, in Table 2.2.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
<th>Repetitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>98 C</td>
<td>120 s</td>
<td>1x</td>
</tr>
<tr>
<td>98 C</td>
<td>30 s</td>
<td></td>
</tr>
<tr>
<td>53 C</td>
<td>30 s</td>
<td>32x</td>
</tr>
<tr>
<td>72 C</td>
<td>60 s</td>
<td></td>
</tr>
<tr>
<td>72 C</td>
<td>10 min</td>
<td>1x</td>
</tr>
<tr>
<td>10 C</td>
<td>Hold</td>
<td>1x</td>
</tr>
</tbody>
</table>

Table 2.2. Cycling conditions for FREQ-Seq construct PCR#3.
Products of this reaction were purified using PCR purification. Following purification, concentration of samples was measured by nanodrop, and samples were then pooled and sent out for Illumina sequencing. Construct assembly process is shown below in Figure 2.1.

**Figure 2.1 FREQ-Seq Barcode Assembly Scheme.** Section (A) shows a PCR amplified product, with a mutated section highlighted with a lightning bolt. The region around the mutation site is then amplified in a second PCR reaction, with an M13 tag on the forward
primer, and an Illumina R tag on the reverse primer. The product of this reaction is then used in a third PCR reaction, containing an Illumina R primer, an illumina F primer, and a barcode primer that serves as an identifier for the sample.

2.9. Analysis of FREQ-Seq data. A total of 23 barcodes were used, one for each day, and 13 loci were sequenced per barcode. There was considerable variation in read coverage among days. The read coverage for two barcodes (13 and 21) was too low to give reliable estimates. Therefore, these barcodes were re-sequenced as well as any sample with a read coverage below 100 reads. The overall average read coverage was 19,321 reads per loci and day (median = 979), with an average coverage of 1,115 reads per loci and day for the first sequencing set (N=245) and an average coverage of 162,035 for the second set (N=31).

To assess the frequencies of the different alleles in the populations, we identified all the reads that matched a search query for the wildtype allele (same bases as the reference sequence S613 at the mutation site) or a mutant allele at the mutation site perfectly. The queries consisted of an A, C, G or T at the SNP site and the 10 flanking bases up- and downstream. The frequency of a mutant allele was calculated as the number of reads with that particular mutations divided by the total number of alleles with perfect matches (for A, C, G or T at the SNP site). Indels were approached in a similar way, however instead of searching for possible SNPs, the search query consisted of the indel sequence and the flanking ten bases upstream and downstream. To assess the frequencies of mutations, the number of reads with the wildtype sequence were determined, as well as the number of reads with two or one deletions and two or one
insertions of the same indel sequence, while the flanking regions of ten nucleotides upstream and downstream of the mutated region were held constant. The frequency of a mutant allele was calculated as the number of reads with a mutation divided by the total number of reads for this site (which included wildtype, two and one insertions, and two and one deletions).

Error rates for single nucleotide substitutions were also assessed for the four bases (positions) surrounding the SNP. The average error rate at a given position was calculated as (1 - frequency of non-wildtype allele)/3 and the average rate for a mutation site was calculated as the average across the four positions surrounding the mutation site. The error rate for a deletion was calculated as the (frequency of (2 deletions+1 insertion+2 insertions))/3 and similarly for insertions as (frequency of (2 deletions+1 deletions+2 insertions))/3. As liaF position 168 had both insertion and deletion mutant alleles, these two mutations were omitted from the error rate calculations. The overall average error rate across sites and days was 0.0018 with a 95% CI confidence interval ranging from 0.0016 to 0.002. The average error rate for SNPs was 0.0028 (± 0.0002 95% CI) and 0.0003 (± 0.00014) for indels. Based on this error rate, true mutations could be readily distinguished from sequencing errors at a level of 0.5%. Any mutant allele frequencies below 0.5% were treated as missing values.

2.10 Biofilm formation assays. Biofilm formation was measured for strains R712, S613, TDR4, TDR7, TDR8, TDR13, TDR19.1, TDR19.2, TDR22, and TDR28 using a crystal-violet staining procedure adapted from Mohamed et al. The assay was begun by streaking frozen strains onto bile esculin agar, and incubating twenty-four hours at 37°C. The next day, a swab of cells was restreaked on
LBHI agar and incubated another twenty-four hours. On the third day, a swab of cells was inoculated into freshly filter-sterilized TSBG media (tryptic soy broth + 0.25% glucose) and the liquid culture was incubated at 37°C for twenty-four hours.

On the fourth day, TSBG cultures were vigorously vortexed to homogenize the culture. Cultures were then diluted 100 fold in fresh TSBG. Aliquots of 200 µl of diluted culture were pipetted into a 96-well flat-bottomed plate, which was placed in a low-disturbance incubator for exactly 24 hours at 37°C.

On the fifth day, the plate was removed from the incubator and an OD$_{600}$ was taken using a microplate reader (noting if any wells have abnormal OD$_{600}$ values). Media was gently dumped out by inverting the plate and gently swirling and flicking. Using a multi-channel pipettor, each well was gently rinsed three times with phosphate buffered saline (PBS), gently dripping PBS into each well and removing it by inverting a flicking the plate. After the final wash, the plate was inverted and dried over a stack of paper towels. Washing steps produced biohazard waste, and required consistency moreso than the remainder of the protocol.

Cells were fixed to the plate by adding 200 µl Bouin’s fixative to each well under a fume hood and incubating 30 minutes at room temperature in the dark. Fixative was discarded by inverting the plate, noting that there was no longer a need to be gentle. The plate was washed once more with PBS.

Fixed cells were stained by adding 200 µl of 100% crystal violet to each well, and incubating for 30 minutes at room temperature. Crystal violet was removed by inverting the plate, then washing the plate with distilled water by submerging the plate into a
container filled with distilled water and swishing it. Water was removed by inverting the plate, and the distilled water rinse was repeated twice more.

The stained plate was blotted again with paper towels, and allowed to air-dry for 15 minutes at room temperature in the dark. Crystal violet was solubilized by adding 200 µl of ethanol:acetone (80:20) mixture to each well and incubating for 20 minutes at room temperature in the dark. Using the microplate reader, \( \text{OD}_{570} \) was measured to quantify biofilm formation via crystal violet absorption. \( \text{OD}_{570} \) was measured three times to ensure consistency, vigorously shaking the plate for 10 minutes inside the plate reader before each reading.

During set-up of 96-well plates growing biofilm-forming culture, no less than 4 wells were used for each strain. Each assay also included at least one column with only media (no bacterial inoculum). Due to the intrinsic variability of the assay, meticulous technique was crucial, and several replicates were required. Crystal violet solutions were made fresh after two weeks due to precipitation.

Turbidostat population mixtures were also assayed at eight time points during turbidostat adaptation (Days 2, 4, 3, 18, 20, 22, 23, and 24). The turbidostat population mixtures on Day 1 were plated on BEA using an aliquot of the frozen stock, and were propagated on subsequent plates and into liquid broth using loops full of cells to capture the diversity of the evolving population.

2.11. Growth rate assays. Growth rates were measured as a function of DAP concentration using a method previously described by Walkiewicz et al. The assay was begun by first plating strains in bile esculin agar (BEA) and incubating twenty-four hours at 37°C. The next day, a swab of cells was restreaked on LBHI agar and incubated
another twenty-four hours. On the third day, a fresh tube of LBHI liquid media with 0.25% glucose (w/v) was freshly prepared by filter sterilizing, and a swab of bacteria was used to inoculate the broth. The tubes were placed in a shaking 37°C incubator and incubated for twelve to sixteen hours.

On the fourth day, 200 µl of culture was used to inoculate fresh tubes containing 10 mL LBHI + 0.25% glucose (a 1:50 fold dilution). The newly inoculated tubes were grown to OD$_{600}$ = 0.3 – 0.4 (preferably closer to 0.3), recording the OD600. To each well of a 96-deep-well plate, 990 µl of filter-sterilized LBHI broth + 0.25% glucose was added, plus antibiotic if applicable. Next, 10 µl of cell culture was added at appropriate OD$_{600}$ to each corresponding well. Following addition of culture, a multichannel pipettor was used to gently mix each well, transferring 150µl of culture from each well to a normal 96-well plate. The 96-well plate was placed onto the microplate reader at 37°C, and shaken for 24 hours at a medium speed, taking reading at an OD$_{600}$ every 5 minutes.

Data was exported into Microsoft Excel, and RStudio was used to generate growth curves. Plates were set-up to always include one column without culture, and when using drug, one column containing susceptible control, and at least one well per column without drug. Templates of 96-well plates were printed and used to record information for each well.

Growth rates at each concentration DAP were plotted against biofilm formation for each strain, and correlation coefficients were calculated using Microsoft Excel.

2.12 RNA isolation, cDNA synthesis, and operon mapping. To examine liaFSR and yvlB operon structures, RNA from *E. faecalis* strain S613 was purified by first growing
10 ml of cells overnight in liquid LBHI culture at 37°C. Cells were diluted 100-fold, and allowed to grow to mid-log (OD$_{600}$ = 0.2). Cells were pelleted, and enzymatically lysed using lysozyme and mutanolysin (see Genomic DNA isolation and Whole Genome Sequencing section).

Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Valencia, CA). RNA integrity was confirmed using a denaturing agarose gel. RNA was treated with DNA-free Kit (Applied Biosystems, Carlsbad, CA) to remove residual genomic DNA. Removal of DNA was confirmed by the inability to amplify liaF, using RNA to template a PCR reaction. Complementary DNA (cDNA) was synthesized using SuperScript III First-strand synthesis kit (Invitrogen Life Technologies, Grand Island, NY).

Strain S613 cDNA was used as template in a series of PCR reactions to confirm the structure of the liaFSR operon. Using a series of primers each overlapping two genes, we confirmed that gene 1795 (putative transcription elongation factor GreA) is not on the same RNA strand as gene 1796 (liaF) (see Table 4.3). The three known liaFSR components were in turn confirmed to be on the same RNA strand (genes 1796, 1797, and 1798). Finally, gene 1799 (potassium uptake protein TrkA) was found to not be on the same RNA strand as gene 1798 (liaR). Using a similar approach, yvlB (gene 2118) was confirmed to be on the same RNA strand as gene 2117, a putative pspC homologue.

2.13 Construction of reporter plasmids. Reporter constructs were assembled using circular polymerase extension cloning (CPEC), as previously described by Quan et al. Components of the liaFSR system were amplified from E. faecalis S613 genomic DNA, while Lactococcus lactis repA was amplified from the genome of E. faecalis strain.
CK111. T7 RNA polymerase DNA was amplified from plasmid pTara, and GFP DNA was amplified from the genome of *E. coli* strain RU1-GFP.

**2.14 Preparation of competent *Escherichia coli* EC1000 cells.** Chemically competent *E. coli* strain EC1000 cells were prepared by first inoculating a single colony into 5-10 ml of Luria-Bertani broth (LB). After incubating overnight at 37°C, cells were diluted in LB in an Erlenmeyer flask until OD$_{600}$ = 0.1 (approximately 10 fold dilution). Cells were incubated with shaking at 37°C until OD$_{600}$ = 0.3-0.4, then cells were transferred into a falcon tube and chilled in ice water for approximately 15 minutes (Note: do not add salt to the ice water or the cells will freeze).

Cells were pelleted by spinning at 5,000 RPM for 5 minutes, maintaining 4°C temperature. The pellet was drained, then resuspended in 8 ml of ice-cold Transformation Buffer 1 (described below) for each 25 ml of cell culture used. Cells were incubated on ice 15 minutes, and then cells were pelleted again. The pellet was drained and resuspended again, using 2.2 ml of ice-cold Transformation Buffer 2 (described below) per 25 ml culture used. Cells were aliquoted into sterile eppendorf tubes, and stored at -80°C for future use.

To prepare 1L of Transformation Buffer 1, the following were added: 12g RbCl, 9.9 g MnCl$_2$*4H$_2$O, 30 ml of 1 M potassium acetate (pH 7.5), 1.6 g CaCl$_2$*2H$_2$O, and 119 ml glycerol. H$_2$O was added to 1L volume and adjusted to pH 5.8 using 200 mM acetic acid. Buffer was sterilized by filtration, and stored at 4°C.

To prepare 500 ml of Transformation Buffer 2, the following were added: 10 ml of 0.5 M 3-(N-morpholino)propanesulfonic acid (MOPS) at pH 6.8, 0.6 g RbCl, 5.5 g
CaCl$_2$$\cdot$H$_2$O, and 60 ml glycerol. H$_2$O was added to 500 ml volume, buffer was sterilized by filtration, and stored at 4°C.

2.15. Transformation of *E. coli* EC1000. Transforming EC1000 cells by heat-shock was begun by thawing 100 µl of cells on ice. An addition of 1-3 µl of DNA was made (generally plasmid prep) and incubated on ice ~30 minutes. Cells were heat shocked at 42°C for 45 seconds. The cells were placed back on ice, and 350 µl LB was added. Cells were incubated with shaking 30-45 minutes at 37°C. 100-200 µl of cells were plated on an LB plate containing appropriate antibiotic, and incubated overnight at 37°C.

2.16. Preparation of electrocompetent *E. faecalis* CK111 cells. Preparation of electrocompetent *Enterococcus faecalis* strain CK111 cells was begun by inoculating a single colony into 10 ml of brain-heart infusion (BHI) broth and incubating overnight at 37°C with shaking.

On the second day, 2g glycine and 40 ml of 1 M sucrose were added to 40 ml of BHI broth. The solution was filter sterilized, and 2 ml of the overnight CK111 culture were added to it. Cells were incubated at 37°C for 24 hours with gentle agitation.

On the third day, the culture was pelleted by centrifuging 10 minutes at 10,000 rpm at 4°C. The pellet was gently resuspended in 20 ml of chilled Buffer PSM (described below). The cells were pelleted and resuspended two more times, keeping the cells at 4°C. After the third time, cells were resuspended in PSM with 15% (v/v) glycerol until the culture reached OD$_{600}$ = 7 (cells measured after being diluted 10-fold, at which point OD$_{600}$ = 0.7). Cells were aliquoted into pre-chilled eppendorf tubes and stored at -80°C.
To prepare Buffer PSM, a solution containing 7 mM potassium phosphate (pH 7.4) and 1 mM MgCl$_2$ was prepared. After autoclaving the solution, and filter-sterilized sucrose was added to a final concentration of 300 mM.

2.17. Transformation of plasmid DNA into *Enterococcus faecalis* CK111.

Transformation was begun by thawing electrocompetent CK111 cells on ice. After adding 1 µg DNA to 50-100 µl cells, they were incubated ~30 minutes on ice. Next cells were pipetted cells into an electroporation cuvette (0.1 cm), and shocked at 1.25 K (Volts) and 25uF.

After electroporation, 1 ml BHI broth was added, and cells were incubated at 37°C for 1-2 hours with shaking. Next, 100-300 µl were plated on BHI agar with appropriate antibiotic and incubated 24 hours at 37°C. If using chloramphenicol, growth often took more than 24 hours. Thus, to prevent plate from drying edges were wrapped in parafilm during prolonged incubation. When plates had incubated more than 5 days with no transformants, the transformation had failed.

2.18. Conjugation of *E. faecalis* CK111 with *E. faecalis* S613 substrains.

Conjugation was begun by plating both donor strain CK111 (transformed with plasmid, as described in section 2.17) and recipient strain S613 (or S613 substrain) and incubating overnight at 37°C. On the second day, both strains were resuspended in liquid BHI media to OD$_{600}$ = 0.08 – 0.10. In a falcon tube, 5 ml of recipient were mixed with 500 µl of donor (10:1 recipient to donor ratio).
A cellulose nitrate membrane filter (Whatman 0.45µm cat 7184-001) was placed into a device for filter mating (0.45µm Ref Millipore Swinex Cat SX0001300) and autoclaved to sterilize. After placing the bacterial mixture into a 10 ml syringe connected to filter mating device, the transconjugant mix was gently pushed through the filter. The filter was removed with sterile tweezers, and placed onto a BHI plate (with no antibiotics) overnight at 37°C.

On the third day, the filter was removed from the plate and placed into a falcon tube with 1 ml BHI broth. Cells were resuspended thoroughly by vortexing. Cells were diluted 10 fold, and 50 µl were plated on BHI agar containing appropriate antibiotics.

For plasmids containing chloramphenicol acetyltransferase and S613 recipient substrains, BHI containing chloramphenicol and vancomycin was used, as S613 is vancomycin resistant. Optimum concentrations for antibiotics are 10 µg/ml chloramphenicol and 48 µg/ml vancomycin. As noted in section 2.17, *E. faecalis* was allowed multiple days to grow in the presence of chloramphenicol.

**2.19 Fluorescent reporter assays.** Strains containing GFP were assayed for fluorescence by first plating cells overnight. In the case of cells containing a reporter plasmid potentially activated by cell envelope stressors, cells were plated both without DAP and at a concentration of DAP roughly equal to half of the MIC value of the strain in question.

The next day cells were resuspended in liquid LBHI and a droplet was placed onto a microscope slide. The slide was placed into a fluorescent microscope under 60x amplification, and was exposed at 395 nm for 500ms to detect GFP fluorescence.
Captured images were quantified for fluorescence by measuring intensity of individual cells using Imagej software.
Chapter 3. Experimental Evolution adapting *E. faecalis* to DAP

3.1 Introduction. Although humans have been influencing evolution for centuries through agriculture and domestication of animals, the study of evolution within a controlled laboratory environment is thought to have begun with a study published by William Dallinger in 1878. After culturing a “minute and intensely active organism, which, on closer and more careful examination, [he] found to be of a form entirely new to [him]”, Dallinger used an incubator of his own design to gradually adapt cells to higher temperature, growing the microbes on “an infusion of animal matter… dilute with water, [that] had vegetable substances placed in it”. Initially growing at 60°F, Dallinger at the outset observed that aliquots of cells grew with “distress” at 73°F, and could not grow at 158°F. However, after seven years of gradually increasing temperature, Dallinger was able to culture cells with ease at 158°F.

Although use of experimental evolution was sporadic in the century following Dallinger’s experiments, today laboratory evolution has become a more widely used technique for the study of evolution. Researchers have used experimental evolution to select for adaptation across a wide variety of organisms and selection conditions, evolving faster development in fruit flies, running endurance in mice, and cold temperature tolerance in fish. Most experimental evolution studies, however, have focused on microorganisms. As Elena and Lenski succinctly detail in their review on the topic, use of microorganisms has numerous advantages, as microbes propagate easily and rapidly, allow for large populations in small spaces, are easily frozen and stored, frequently reproduce asexually, and frequently have genetic tools readily available.
Perhaps the most well known example of microbial experimental evolution was begun in 1988, in the lab of Dr. Richard Lenski. In this experiment, a clonal population of *Escherichia coli* was first inoculated into liquid Davis minimal (DM) broth media supplemented with 25 mg/L glucose\(^{75}\). Each day, an aliquot of cells were transferred to a fresh flask of media, and another aliquot frozen away for future study. Designed as a long-term evolution experiment, serial transfers are still ongoing into the present day, with the experiment having currently spanned more than 25 years and an estimated 50,000 bacterial generations\(^{76}\). Analysis of frozen samples from this single experiment has garnered tremendous insight into evolution, including a finding that adaptation decelerated in the population despite constant genomic evolution\(^{77}\), and explaining the prevalence of a mutator phenotypes through a study of mutation supply rate\(^{54}\). One particularly striking observation, the identification of a subpopulation that evolved the capacity to metabolize citrate\(^{78}\), provided unique insight into development of gain-of-function mutations, demonstrating the significance of historical contingency in their development\(^{79}\). Serial flask-transfer approaches have been applied across other microbes as well, having been used to monitor the success of beneficial mutations across numerous parallel yeast populations\(^{53}\) and explore epistasis among beneficial mutations in *Methylobacterium extorquen*\(^{52}\). While flask-transfer approaches are used in this thesis to study adaptation of *E. faecalis* to Daptomycin, this work focuses predominantly on adaptation within a turbidostat.

Unlike flask-transfer experiments, in which populations are subjected to a series of batch cultures and undergo 5-10 doublings per day over a range of population sizes and varying drug concentrations, a turbidostat can maintain a single large population
(\textgreater 10^{11} \text{ total cells}) in continuous culture throughout selection (see section 2.4). By constantly titrating in new media as a function of population size, the turbidostat allows precise control of antibiotic concentration and maintains the population in exponential growth at a constant density, producing over 1,500 generations over 24 days. Maintenance of exponential growth establishes an important distinction from previous chemostat experiments, which maintain a much slower growth rate by diluting cells in stationary phase \textsuperscript{80}. Using the same growth vessel throughout the course of the experiment, turbidostats maintain both planktonic and non-planktonic (e.g. biofilm) populations in co-culture while increasing selective pressure. The objective of this design was to establish a method that can apply accurate and controllable selection conditions to a sufficiently large population to rapidly survey accessible adaptive mutations available to the evolving genome.

Within the turbidostat, we estimate that the entire genome of \textit{E. faecalis} will experience sufficient single nucleotide mutations to provide a thorough exploration of genomic sequence space during a single day (even under non-stressful conditions). Indeed, if we assume a “non-stressful” basal mutation rate comparable to that of gram-positive bacteria \textit{B. subtilis} (4.7 \texttimes 10^{-10} mutations per nucleotide per generation) \textsuperscript{81}, we estimate that 1.4 \texttimes 10^8 random mutations should occur each generation in the turbidostat (considering the constant population of roughly 10^{11} cells and the \textasciitilde 3 Mb genome). Mutation rate plays an important role in adaptation, and under the stress of selection it is possible that mutation rates may be significantly higher than basal levels measured for \textit{B. subtilis} or \textit{E. coli} \textsuperscript{49,60,81}.
As discussed in Chapter 1, evolution of antibiotic resistance is a primary focus of this work. Antibiotics not only provide a powerful selective tool for experimental evolution, but insight gleaned from laboratory adaptation using antibiotics can provide data of medical as well as evolutionary relevance. In one such study in 2006, Friedman et al. used a flask-based approach in _S. aureus_ to identify Daptomycin resistance determinants, providing greater insight into a phenotype that had become a problem in the clinic. In another study in 2012, Toprak et al. 58 engineered an experimental evolution system with a novel selection regime (dubbed the “morbidostat”) to follow evolutionary trajectories of resistance to multiple antibiotics in _E. coli_. By increasing drug concentration as cells grew more rapidly, Toprak et al. were able to maintain a constant growth rate while maximizing tolerable stress, leading to a uniquely rapid selection regime 58. In this thesis, a turbidostat growth system was used with an antibiotic selection regime where concentration was adjusted each day to the maximal dosage that did not significantly reduce the doubling time of the population (see section 2.4). Although the primary focus of this thesis lies on analysis of evolutionary trajectories throughout adaptation within a turbidostat, flask-based adaptation experiments were first utilized to identify parameters for subsequent experiments using a turbidostat. Study of flask-adapted strains also provides a unique opportunity to observe how adaptation differs across the two environments.

### 3.2 Flask based adaptation

To evaluate the timescale involved in evolution of DAP resistance in _E. faecalis_, flask adaptation experiments were performed using two different strains, OG1RF and S613 (as described in section 2.2). Minimum inhibitory
concentrations (MICs) of each strain on each day of the experiment are shown below in
*Figure 3.1.*

*Figure 3.1. Serial Flask Adaptation of E. faecalis strains S613 and OG1RF to DAP.*

Strains were serially propagated twice daily through incrementally increasing concentrations of Daptomycin, as described in Section 2.2. MIC values (shown on Y-axis) were determined using broth macrodilution. *By day 14, strain S613 was able to tolerate over 100 µg/ml of Daptomycin.*

Although the two strains had different DAP tolerance on each day of the experiment, by day 14 both strains could tolerate over 100 µg/ml of Daptomycin. This demonstrates both that *E. faecalis* can become DAP resistant in a relatively short timescale, despite differences in selections regimes the two strains are able to tolerate.

To confirm the above results, the above experiment with strain S613 was replicated twice: once by myself, and another independent replicate in the lab of Dr. Caesar Arias. Following adaptation, multiple endpoint colonies from each S613
adaptation were isolated, and genes *liaF*, *gdpD*, and *cls* were sequenced, with results shown below in *Table 3.1*.

**Table 3.1. Genotypes of Serial Flask Adapted DAP Resistant S613 Variants.**

<table>
<thead>
<tr>
<th>Sample</th>
<th>liaF genotype</th>
<th>gdpD genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adapt #1, colonies 1-17</td>
<td>Ins I177</td>
<td>WT</td>
</tr>
<tr>
<td>Adapt #1, colony 18</td>
<td>ΔI177</td>
<td>ΔI170</td>
</tr>
<tr>
<td>Adapt #2, colonies 1-6</td>
<td>Ins I177</td>
<td>WT</td>
</tr>
<tr>
<td>Arias Adapt, colonies 1-6</td>
<td>Ins I177</td>
<td>WT</td>
</tr>
</tbody>
</table>

Colonies were isolated and sequenced from each of three serial flask adaptation experiments. While cardiolipin synthase (CLS) was also sequenced in each case, no mutations were identified. Samples from “Adapt #1” and “Adapt #2” were derived from adaptation experiments performed in the Shamoo lab, while samples from “Arias Adapt” were derived from experiments performed in the Arias lab. In the above table, “Ins” denotes an insertion mutation, “Δ” denotes a deletion mutation, and “WT” denotes wild-type cells.

These three genes were chosen for sequencing after mutations were observed in each within clinically isolated resistant strain R712. In experiments “Adapt #2” and “Arias adapt”, the endpoint colonies suggest a homogenous population at the end of the experiment, as each colony isolated has only mutation *liaF*^{ins I177}. Experiment “Adapt #1”, however, shows one anomalous colony with mutant alleles *liaF*^{ΔI177} *gdpD*^{ΔI170}. This colony not only demonstrates heterogeneity within the endpoint population of the experiment, but also possesses a genotype similar to that of strain R712, which contains
mutant alleles $liaF^{\Delta I_{177}}$, $gdpD^{\Delta I_{170}}$, and $cls^{AK_{61}}$. Despite this discrepancy, results from each flask based adaptation seem largely reproducible, predominantly resulting in only mutation $liaF^{\text{ins I}_{177}}$. Significantly, $cls$ mutations were not recovered in any flask-based adaptations described above, which later studies would suggest play a role in alterations of the cell envelope and production of biofilm in DAP resistant strains (See chapter 6)\(^{25}\).

As Daptomycin is only efficacious against gram-positive bacteria, inclusion of a broader spectrum antibiotic should inhibit growth of a wider variety of contaminating bacteria. While contamination may not typically be an issue during a single overnight growth, in a continuous evolution experiment a contaminating organism has a long period of time to grow and compete with the organism of interest, allowing even a small number of contaminating cells to interfere with the experiment. In an attempt to address issues with contamination during serial transfer experiments, a fourth serial transfer experiment was conducted as described above, in the presence of a constant quantity of tetracycline (25 µg/ml). Tetracycline was chosen as *E. faecalis* S613 can tolerate high amounts of the drug without a significant decrease in growth rate. Upon completion of this serial adaptation, however, no mutations were observed in $liaF$, $gdpD$, or $cls$. This suggests that despite the capacity of strain S613 to tolerate tetracycline, presence of the drug seems to significantly alter the evolutionary landscape as cells evolve DAP resistance.

To examine the difference between serially propagating planktonic and stationary (biofilm) cultures, a final flask adaptation was performed (by graduate student Yanru Dou) wherein a metallic surface was serially transferred between cultures rather than a liquid aliquot (see Section 2.2). During the course of the experiment, a colorless biofilm accrued on the surface of the razor blade, seeming to increase in mass day by day.
Following adaptation, six colonies were isolated and sequenced for liaF, cls, gdpD, and drmA, with genotypes shown below in *Table 3.2*.

**Table 3.2. Genotypes of Solid Surface Transferred DAP Resistant E. faecalis**

<table>
<thead>
<tr>
<th>Sample</th>
<th>liaF genotype</th>
<th>cls genotype</th>
<th>drmA genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colony 1</td>
<td>ΔI177</td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td>Colony 2</td>
<td>Ins I177</td>
<td>-</td>
<td>WT</td>
</tr>
<tr>
<td>Colony 3</td>
<td>ΔI177</td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td>Colony 4</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td>Colony 5</td>
<td>Ins I177</td>
<td>WT</td>
<td>WT</td>
</tr>
<tr>
<td>Colony 6</td>
<td>Ins I177</td>
<td>ΔNFQ(74-76)</td>
<td>L4Stop</td>
</tr>
</tbody>
</table>

*Six colonies were isolated from planktonic culture from the endpoint of the solid surface serial adaptation experiment. Samples were sequenced for liaF, cls, gdpD, and drmA, although no mutations were observed in gdpD. In the above table, Dash denotes sequencing failure, “Ins” denotes an insertion mutation, “Δ” denotes a deletion mutation, and “WT” denotes wild-type cells.*

Surprisingly, serial propagation across a solid surface revealed very different genotypic results in isolated colonies. Across the six colonies sequenced, genotype appears to have been far more variable than previously observed during planktonic transfers. As the environment within each flask becomes more complex with the addition of a solid surface, it is plausible that cells can adapt to different niches, thereby leading to greater diversity.
Among the colonies shown in Table 3.2, Colony 6 is unique in that it possesses mutations both in \( cls \) and \( drmA \), marking the first instance in which a mutation in this clinically relevant DAP resistance gene. Interestingly, the genotype of this colony (\( liaF^{\text{ins1177}} \), \( cls^{\text{ANFQ(74-76)}} \), \( drmA^{L4\text{Stop}} \)) bears remarkably similarity to that of turbidostat isolated strain TDR13 (\( liaF^{\Delta1177} \), \( cls^{\text{ANFQ(74-76)}} \), \( drmA^{L4\text{Stop}} \)), the isolation of which is discussed in chapter 4 (see Table 4.2). As strain TDR13 has the highest biofilm-forming propensity of any strain studied in this thesis (see Figure 6.1), it is likely that propagation of biofilm coating the solid surface played a key role in facilitating the success of this genotype within the flask.

3.3 Turbidostat based adaptation.

After ascertaining that \( E.\ faecalis \) strain S613 can evolve DAP resistance within a two-week time scale (see section 3.2), a continuous culture turbidostat was used to run an adaptation experiment (see sections 2.3, 2.4). The turbidostat was run for a period of approximately 24 days, increasing concentration of DAP incrementally, as shown below in Figure 3.2.
Figure 3.2. Turbidostat Adaptation Drug Ramp and Media Consumption. As described in Section 2.4, the turbidostat was run for a period of 24 days total, increasing from 0 µg/ml DAP to a maximum of 20 µg/ml. Prior to the run shown above, cells were grown in the turbidostat for seven days, after which a single colony was isolated, and used to reinoculate the turbidostat for the DAP adaptation experiment. Carboy icons shown above represent media usage each day. Around day 12 (2 µg/ml DAP) media usage appeared to increase, as did the quantity of biofilm visible inside the growth chamber.

Cells were grown at 8 sequentially higher concentrations of DAP: 0 µg/ml, 0.5 µg/ml, 0.75 µg/ml, 2 µg/ml, 3 µg/ml, 4 µg/ml, 6 µg/ml, and 20 µg/ml. Following growth at 2 µg/ml, the quantity of biofilm appeared to visibility increase inside the vessel, and the media consumption rate increased as well. On each day of the experiment, multiple 10 ml population samples were taken and stored at -80°C for future study. MIC tests
performed on frozen isolates following the conclusion of the adaptation experiment
demonstrate that DAP resistance gradually increased during adaptation, as shown below
in Table 3.3.

Table 3.3. DAP MICs of Intermediary Turbidostat Populations

<table>
<thead>
<tr>
<th>Days after DAP exposure</th>
<th>MIC (µg/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.75</td>
</tr>
<tr>
<td>4</td>
<td>0.25</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td>24</td>
<td>24</td>
</tr>
</tbody>
</table>

Minimum inhibitory concentrations of DAP were determined for mixed population
samples during the course of adaptation using Etest.

As Table 3.3 demonstrates, DAP resistance within the population increased over
time, eventually yielding a highly resistant population on the final day of the experiment.
After day 24, a population sample was plated on LBHI and a number of colonies were
selected for further study (see sections 2.4-2.5). To accommodate the additional
complexity of the analysis of the turbidostat experiment, Chapter 4 is consigned to its
description.
Chapter 4. Isolation and Sequence analysis of DAP resistant strains

4.1 Introduction. As an organism enters a new environment and grows into a population, it naturally begins to acquire variability. While selective pressure trims some of this variability, frequently an amount of diversity is maintained as organisms evolve to occupy specific niches. In the case of experimental evolution, while the laboratory environment may lack the complexity of an ecological setting, previous experiments have found diverse lineages within the same laboratory population.

In a long-term experimental evolution study performed in the lab of Dr. Richard Lenski, an initially clonal population of *E. coli* was grown for over 40,000 generations in flasks using glucose-limited minimal media. Although citrate was present in the media, the bacteria were unable to efficiently utilize the nutrient until after approximately 31,000 generations, at which time a citrate-using variant (Cit+) was isolated. Interestingly, despite the clear advantage of Cit+ mutants in this experiment, Cit− variants persisted in the population as well, with both variants being isolated from the same time-point. Using whole genome sequencing, the authors compared these variants and identified the determinants of citrate usage, as relatively few mutations separated the two lineages. However, it is important to note that while Cit+ and Cit− mutants isolated after 31,000 generations do not have many mutations relative to one another, they both contain numerous mutations relative to their ancestral founding strain, which in turn suggests that historical contingency plays a key role in evolution of citrate usage. As this study demonstrates, strain isolation coupled with whole genome sequencing not only can be used to follow the evolution of a single lineage, but also to examine the evolutionary dynamics of a population.
In a similar study by Herring et al. using flask-based transfer of *E. coli*, a population was serially transferred across minimal media supplemented with glycerol, a poorly utilized nutrient. After ~223 generations (44 days), the population had evolved the ability to metabolize glycerol, and five strains were isolated and compared to their ancestor using whole genome sequencing. Although each strain contained multiple mutations (many of them different), it was observed that all five strains had a mutation in glycerol kinase *glpK*, and two strains had a mutation in RNA polymerase subunit *rpoC*. This observation shows that conserved evolutionary trends can be observed using whole genome sequencing by comparing genotypes across polymorphic isolates.

As discussed in Chapter 3, clinically isolated *E. faecalis* strain S613 was continuously cultured in a turbidostat under increasing concentrations of DAP, evolving resistance across a total of 23.8 days of drug exposure. In previous studies using turbidostats, Couñago et al. found that as a mesophilic bacteria adapted to higher temperatures, a polymorphic population arose within the turbidostat as numerous subpopulations competed. However, by the final day of the experiment, a single subpopulation (mutant allele *Adk*Q119R/Q16L) had outcompeted other mutant lines, sweeping the population. Given the shorter generation time and larger population size within the turbidostat relative to flask-based studies, it is plausible that at the end of a turbidostat adaptation experiment a single “most fit” strain will have swept the population. Conversely, in experiments using enterococci the prevalence of biofilm inside the turbidostat growth vessel may either facilitate persistence of less fit strains or allow multiple subpopulations to flourish in disparate niches, with both situations leading to increased polymorphism.
By isolating and sequencing strains following evolution of DAP resistance in \textit{E. faecalis} strain S613 in the turbidostat, not only can the degree of polymorphism present within the population be examined, but specific DAP-resistant lineages can also be studied in greater detail. Using whole genome sequencing and comparing isolates to ancestral strain S613, resistant genotypes and conserved evolutionary trajectories can be identified between strains. Additionally, isolating individual resistant strains allows additional study of resistance-associated phenotypes.

\textbf{4.2 TDR strain isolation and screening.} At the end of the turbidostat experiment (discussed in section 3.3), a population sample was taken and plated on non-selective LBHI media. From this plate, 48 strains were isolated, stored, and given names beginning with TDR1-48, with TDR standing for Turbidostat-derived DAP Resistant, and the number standing for the order in which they were isolated (for instance, the fourth strain isolated is TDR4). Previously identified DAP resistance associated genes \textit{liaF}, \textit{CLS}, and \textit{gdpD} were then sequenced in each TDR strain\textsuperscript{10}. Minimum inhibitory concentrations (MICs) for DAP were also determined for each strain using agar microdilution method\textsuperscript{63}. The results of these experiments are shown on the following page, in Table 4.1. Note that following later acquisition of whole genome sequence data (detailed in section 4.3), \textit{liaFSR} genes \textit{liaR} and \textit{yvlB} were sequenced in several strains as well, with this data also shown in Table 4.1
**Table 4.1 Genetic Screen and MIC testing of Randomly Isolated TDR strains**

<table>
<thead>
<tr>
<th>Strain</th>
<th>liaF</th>
<th>cls</th>
<th>liaR</th>
<th>yvlB</th>
<th>MIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>S613</td>
<td>WT</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>0.25</td>
</tr>
<tr>
<td>TDR6</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>1</td>
</tr>
<tr>
<td>TDR15</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>TDR23</td>
<td>insI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>TDR2</td>
<td>ΔI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>TDR1</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR12</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR17</td>
<td>ΔI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR18</td>
<td>ΔI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR20</td>
<td>T194I</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR21</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>2</td>
</tr>
<tr>
<td>TDR25</td>
<td>WT</td>
<td>ΔNFQ74-76</td>
<td>WT</td>
<td>V289fs</td>
<td>2</td>
</tr>
<tr>
<td>TDR34</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR3</td>
<td>T194I</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR11</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR14</td>
<td>ΔI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR24</td>
<td>WT</td>
<td>ΔNFQ74-76</td>
<td>WT</td>
<td>V289fs</td>
<td>2</td>
</tr>
<tr>
<td>TDR39</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR41</td>
<td>insI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR48</td>
<td>insI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>R712</td>
<td>ΔI177</td>
<td>ΔK61</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR29</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR31</td>
<td>insI177</td>
<td>V37A</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR37</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR46</td>
<td>ΔI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>TDR9</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>TDR</td>
<td>Genotype</td>
<td>Genotype</td>
<td>Genotype</td>
<td>Genotype</td>
<td>Genotype</td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>10</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>4</td>
</tr>
<tr>
<td>35</td>
<td>ΔI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>36</td>
<td>T194I</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>32</td>
<td>WT</td>
<td>ΔNFQ74-76</td>
<td>WT</td>
<td>V289fs</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>WT</td>
<td>ΔNFQ74-76</td>
<td>D191N</td>
<td>WT</td>
<td>4</td>
</tr>
<tr>
<td>33</td>
<td>insI177</td>
<td>V37A</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>38</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>40</td>
<td>ΔI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>43</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>45</td>
<td>ΔI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>26</td>
<td>ΔI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>30</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>V289fs</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>WT</td>
<td>WT</td>
<td>WT</td>
<td>V289fs</td>
<td>8</td>
</tr>
<tr>
<td>47</td>
<td>ΔI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>ΔI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>ΔI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>insI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>27</td>
<td>insI177</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>42</td>
<td>ΔI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>22</td>
<td>T194I</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>44</td>
<td>T194I</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>19.2</td>
<td>WT</td>
<td>ΔNFQ74-76</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>28</td>
<td>insI177</td>
<td>WT</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>19.1</td>
<td>WT</td>
<td>R218Q</td>
<td>WT</td>
<td>V289fs</td>
<td>8</td>
</tr>
</tbody>
</table>

_Turbidostat derived strains (TDR) were randomly isolated as single colonies from LBHI agar plates without DAP at 37°C at the end of the turbidostat experiment (see Section 2.5). Genes liaF and CLS (cardiolipin synthase) were sequenced in all 50 strains to_
screen for distinct genotypes for use in whole genome sequencing. Following identification of mutant liaR and yvlB alleles using whole genome sequencing (see Section 4.3), these genes were also sequenced in TDR strains lacking mutations in liaF. Three strains (TDR6, TDR10, and TDR21) did not have any mutations in liaF, liaR or yvlB and sequencing of pspC and liaS showed WT alleles at these loci, which suggest that these strains either had no mutations within the LiaFSR pathway or had new mutations in unknown downstream targets. All 50 strains isolated had MIC\(_{\text{DAP}}\) values measured in triplicate using agar microdilution assay, with average value shown. In the above table, “Ins” denotes an insertion mutation, “∆” denotes a deletion mutation, “fs” denotes a frameshift mutation, “WT” denotes wild-type cells, and dash denotes no data was collected.

Comparing the variable genotypes and MIC results shown in Table 4.1 across strains, it became rapidly apparent that the endpoint population of the turbidostat was highly polymorphic. While some TDR strains are highly resistant (MIC ≥ 8µg/ml) and others less so (MIC of 1 µg/ml), every TDR strain is significantly more resistant than clinical strain S613.

Looking at the genotypic data in Table 4.1 in more detail, some interesting trends become apparent. Among the 50 randomly selected end-point strains, 94% have a mutation in the liaFSR pathway, including yvlB (for linkage between yvlB and liaFSR, see Appendix B). The majority of mutations occurred in liaF (78%), with changes in yvlB (12%) and liaR (4%) comprising the remainder. Strikingly, 27% of liaF mutants had the same amino acid deletion observed in clinical strain R712 (liaF\(^{\text{Al177}}\)). Although 6% of
strains in *Table 4.1* have no observed mutation in the *liaFSR* system, note that there are likely multiple unknown downstream targets of the *liaFSR* system that may contain mutations in these strains (for additional discussion of this idea, see Appendix B).

While 46% of strains shown in *Table 4.1* carried a mutation in *cls*, surprisingly the mutant allele present in clinical strain R712 (*cls*\(^{AK61}\)) was not observed in any TDR strain. Mutations in gene *gdpD*, the third mutated gene in strain R712, were also not observed in any TDR strains.

Overall, this data demonstrates that *liaFSR* mutations are critical for DAP resistance. *cls* mutations also appear to play a significant role, though they are not nearly as prevalent as *liaFSR* mutations. *GdpD*, conversely, appears to not be a significant DAP resistance determinant.

Considering the variable evolutionary trajectories followed by many TDR strains, it is likely that additional genes carry mutations elsewhere in the genome. To explore this possibility, seven genetically distinct strains were chosen for whole genome sequencing, each with MIC\(_{\text{DAP}} \geq 4\)µg/ml. The strains chosen were TDR4, TDR7, TDR8, TDR13, TDR19, TDR22, and TDR28.

### 4.3 Comparative whole genome sequence analysis and putative functions of resistance associated alleles

Whole genome sequencing reactions were run by SeqWright, as discussed in section 2.6. Also discussed in section 2.6, sequencing data revealed the strain designated “TDR19” in *Table 4.1* was in fact a mixture of three strains, all of which were isolated following this finding. Two of these strains were genetically distinct, and were hence renamed “TDR19.1” and “TDR19.2”, while the third strain was
genetically identical to TDR8. Coding mutations identified in each of these strains are shown below, in Table 4.2.

**Table 4.2. Coding mutations identified in DAP resistant E. faecalis S613 strains derived from a convergent phenotype (DAP resistance)**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Coding changes identified during DAP adaptation</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDR4</td>
<td>liaK&lt;sup&gt;D191N&lt;/sup&gt; cls&lt;sup&gt;ΔNFQ74-76, S104L&lt;/sup&gt; gshF&lt;sup&gt;E554K&lt;/sup&gt; yyb&lt;sup&gt;T1440S&lt;/sup&gt;</td>
</tr>
<tr>
<td>TDR7</td>
<td>yvlB&lt;sup&gt;V289fs&lt;/sup&gt;</td>
</tr>
<tr>
<td>TDR8</td>
<td>lia&lt;sup&gt;ΔI177&lt;/sup&gt; cls&lt;sup&gt;ΔNFQ74-76&lt;/sup&gt;</td>
</tr>
<tr>
<td>TDR13</td>
<td>lia&lt;sup&gt;ΔI177&lt;/sup&gt; cls&lt;sup&gt;ΔNFQ74-76&lt;/sup&gt; drmA&lt;sup&gt;L4stop&lt;/sup&gt;</td>
</tr>
<tr>
<td>TDR22</td>
<td>lia&lt;sup&gt;T194I&lt;/sup&gt; drmA&lt;sup&gt;N150fs&lt;/sup&gt;</td>
</tr>
<tr>
<td>TDR28</td>
<td>lia&lt;sup&gt;ΔI177&lt;/sup&gt; drmA&lt;sup&gt;N150fs&lt;/sup&gt; mdpA&lt;sup&gt;F185fs&lt;/sup&gt;</td>
</tr>
<tr>
<td>TDR19.1</td>
<td>yvlB&lt;sup&gt;V289fs&lt;/sup&gt; cls&lt;sup&gt;R218Q&lt;/sup&gt;</td>
</tr>
<tr>
<td>TDR19.2</td>
<td>lia&lt;sup&gt;D191N&lt;/sup&gt; cls&lt;sup&gt;ΔNFQ74-76, ins A114-16&lt;/sup&gt; gshF&lt;sup&gt;E554K&lt;/sup&gt; yyb&lt;sup&gt;T1440S&lt;/sup&gt;</td>
</tr>
<tr>
<td>R712</td>
<td>lia&lt;sup&gt;ΔI177&lt;/sup&gt; cls&lt;sup&gt;AK61&lt;/sup&gt; gdp&lt;sup&gt;ΔI170&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

*Turbidostat Derived Strains (TDR) are end-point isolates from day 23 of experimental evolution. Comparative whole-genome sequencing of the end-point strains relative to E. faecalis S613 revealed adaptive changes associated with adaptation to DAP. All mutations in coding regions shown above (for non-coding changes, see Table 4.4). Above “Δ” denotes a deletion, “ins” an insertion, and “fs” a frameshift.*
Among the genotypes outlined in Table 4.2, coding mutations were identified in a total of 8 genes. Gene ID information for each mutated allele is shown below in Table 4.3, based on the GFF file for strain S613 deposited in PATRIC (for genome assembly details see section 2.6).

**Table 4.3. Names and reference numbers of DAP resistance genes.**

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1268</td>
<td>Putative c-di-Amp phosphodiesterase YybT</td>
</tr>
<tr>
<td>1796</td>
<td>LiaF</td>
</tr>
<tr>
<td>1798</td>
<td>LiaR</td>
</tr>
<tr>
<td>2118</td>
<td>YvlB</td>
</tr>
<tr>
<td>2158</td>
<td>Daptomycin resistance membrane protein A (drmA)</td>
</tr>
<tr>
<td>2231</td>
<td>Glutathion Synthesis Protein F (GshF)</td>
</tr>
<tr>
<td>2385</td>
<td>Cardiolipin synthase (cls)</td>
</tr>
<tr>
<td>2690</td>
<td>Multidrug resistance pump A (mdpA)</td>
</tr>
</tbody>
</table>

Gene ID numbers annotated using gff3 file dated 11/12/ 2010, downloaded from PATRIC at [http://www.patricbrc.org](http://www.patricbrc.org) for E. faecalis S613. For liaF, liaR, gshF, and cls, common names were also derived from gff3 reference file. Gene yybT is named for a homologue in B. subtilis (98% coverage, 39% identity), which has been previously characterized\(^84\), though its closest homologue in Staphylococcus aureus is denoted as gdpP\(^85\). Gene yvlB is also named for a homologue in B. subtilis (59% coverage, 32% identity) previously identified by whole genome sequencing\(^86\), though its function remains unknown. As drmA and mdpA did not have clear homologues of known function, gene names were coined for this study.
Looking at the collection of DAP resistant genotypes shown above in Table 4.2, several trends become apparent. Perhaps most notably, in every TDR strain sequenced, a mutation was observed in the liaFSR signaling pathway. LiaFSR is a component of the CESR regulon, and responds to changes in cell envelope integrity by regulating downstream genes to counteract damage. LiaFSR mediated responses that lead to remodeling of the cell envelope are consistent with studies suggesting that DAP inserts and subsequently oligomerizes in the membrane in a calcium and phosphatidylglycerol (PG) dependent manner, which disrupts the functional integrity of the membrane. In B. subtilis, LiaF strongly inhibits signaling through LiaS and its cognate response regulator LiaR (EMBL entries AEP92336, CAA11744, and CAA11745, respectively), which in turn regulates its own operon as well as yvlB and others. In E. faecalis S613, the genes yvlB and pspC are likely in the same operon, and mutations in either of these genes have previously been associated with DAP resistance in enterococci (see Appendix B). Currently knowledge of this pathway is summarized graphically below in Figure 4.1.
Figure 4.1. LiaFSR signaling pathway in Enterococcus faecalis. Proposed functions of liaFSR components are based on homologous proteins previously studied in Streptococcus mutans$^{89}$. Linkage between yvlB, pspC, and liaFSR signaling system was elucidated in this study, and is discussed further in Appendix B. It is also likely that liaR regulates additional operons not shown above.

Strain TDR7 was observed to have only a single mutant allele, $yvlB^{V289fs}$, yet was DAP resistant nonetheless, suggesting that truncation of the C-terminal domain of YvlB is sufficient for resistance. Palmer et al. also observed a frameshift mutation at position
Ile-390 in YvlB (EF1753) of a DAP resistant *E. faecalis* variant of strain V583, demonstrating that the importance of this protein is not confined to a single strain or experiment.

As none of the adaptive mutations in LiaFSR signaling components introduced stop codons or frameshifts that would decouple the pathway response, LiaFSR response in enterococci is likely activated by antibiotics that induces cell wall stress, as observed in *B. subtilis, Streptococcus mutans,* and the comparable VraTSR (vancomycin-resistance associated sensor/regulator) in *Staphylococcus aureus.* The prevalence of mutations to liaFSR during DAP adaptation suggests that alteration of this signaling pathway is a general and effective response to membrane insult. However, as no TDR strain studied carries more that one mutation in this pathway, it is likely that while a single mutation is necessary for DAP resistance, additional mutations do not lead to additional increases in fitness.

As discussed briefly in section 4.2, mutations to gene *CLS* were among the most common in the turbidostat population (as well as in clinical strain R712), suggesting an important role for changes to Cls activity in response to DAP. A link between Cls activity and DAP adaptation is supported by the observation that Cls catalyzes the formation of cardiolipin (CL) from two molecules of phosphatidylglycerol (PG), and that DAP activity is dependent on the presence of PG and calcium. Clinical isolates *E. faecalis* R712 and *E. faecium* R446 both have mutations in *cls* and a decrease in membrane PG content, as well as distinctive changes to the cell envelope following DAP exposure. *In vitro* biochemical analysis of adaptive mutants Cls$^{R218Q}$ and Cls$^{H215R}$ from *E. faecium* showed
that each mutation increased Cls activity. As $\textit{cls}^{R218Q}$ was observed in \textit{E. faecalis} strain TDR19.1, it is likely that mutations to \textit{cls} in these populations are doing likewise.

Strains TDR22, TDR13, and TDR28 strains also carried a mutation in gene \textit{drmA}, in each case either a frameshift or a SNP leading to a stop codon, likely leading to inactivation of the protein. A previously uncharacterized gene, \textit{drmA} was first identified in this study, its name signifying “Daptomycin resistance-associated membrane protein A”. Bioinformatics analysis suggests that \textit{drmA} encodes a six-pass integral membrane protein of unknown function, leaving open speculation as to the effects of this mutation on protein function. Interestingly, a different frameshift mutation in \textit{drmA} was observed in a separate study in \textit{E. faecalis} V583 after adaptation to DAP highlighting the importance of this gene in resistance.

In strain TDR28, a mutation was observed in gene \textit{mdpA}, another previously uncharacterized gene. Labeled “Multidrug resistance pump A” in this study, bioinformatic analysis suggests \textit{mdpA} is a putative ABC transporter. Curiously, the mutation observed in this gene, F185fs, results in a frameshift mutation, likely disabling protein function. Currently, the role of this protein in DAP resistance is unclear.

In strains TDR4 and TDR19.2, the same mutation was observed in glutathione synthesis protein GshF, a SNP causing mutation E554K. In most organisms glutathione is synthesized by two independent proteins, glutamylcysteine ligase and glutathione synthetase. However, GshF is capable of carrying out the functions of both proteins, allowing it to synthesize glutathione in isolation. GshF has previously been implicated in the oxidative stress response across multiple species. GshF is commonly found among
mammalian pathogens and could also have a role in mitigating DNA damage caused by general oxidative stress.

A second mutant allele common to strains TDR4 and TDR19.2 is \textit{yybT}^{I440S}. YybT homologues in both \textit{B. subtilis} (EMBL entry CAB16088) and \textit{S. aureus} (denoted as GdpP in \textit{S. aureus}, EMBL entry AEW64051) have c-di-AMP phosphodiesterase activity, with potential roles in stress signaling and response. As a global second messenger, cyclic dinucleotide signaling has been implicated in quorum sensing, adhesion, and biofilm formation. While cyclic-dinucleotide signaling has not been previously characterized in enterococci, the identification of a YybT homolog and the subsequent identification of adaptive mutations within this gene is consistent with a role for c-di-nucleotide signaling in stress response.

Although relatively few intergenic mutations were identified, they may also affect oxidative stress response. We were surprised at the paucity of intergenic and silent mutations in the population, as only three putative intergenic mutations and one silent mutation were identified, shown below in Table 4.4.

\textit{TABLE 4.4. Noncoding and silent mutations in TDR strains.}

<table>
<thead>
<tr>
<th>Strains</th>
<th>Ref. Base</th>
<th>Called Base</th>
<th>Contig</th>
<th>Position</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDR28</td>
<td>A</td>
<td>-</td>
<td>772</td>
<td>26895</td>
<td>None</td>
</tr>
<tr>
<td>TDR13</td>
<td>-</td>
<td>A</td>
<td>775</td>
<td>11338</td>
<td>None</td>
</tr>
<tr>
<td>TDR28</td>
<td>G</td>
<td>A</td>
<td>778</td>
<td>82121</td>
<td>None*</td>
</tr>
<tr>
<td>TDR4, TDR19.2</td>
<td>A</td>
<td>G</td>
<td>718</td>
<td>18256</td>
<td>Hyp. Protein</td>
</tr>
</tbody>
</table>

Mutations annotated using gff3 file dated 11/12/2010, downloaded from PATRIC at http://www.patricbrc.org for \textit{E. faecalis} strain S613. *Although designated as intergenic
in PATRIC file, further analysis reveals contig 778 mutation may lie in a gene of a putative SelA homologue. This mutation also displays unusual mutagenic frequency.

The first intergenic change identified, observed only in strain TDR13, occurred in contig 775 as insertion of a single ‘T’ within a putative 14,225bp plasmid. While it is possible this mutation falls in a terminator or promoter region, its potential role in adaptation remains unclear.

Cellular response to oxidative stress may also be affected by two additional mutations in noncoding regions of strain TDR28. As shown in Table 4.4, a mutation within contig 772 produced a deletion of a single ‘T’ within a run of seven ‘T’s. A strong stemloop structure followed by a T-stretch is readily predicted using the program m-fold for this region. The predicted secondary structure is consistent with a Rho-independent termination site, and thus the mutation is within a potential transcriptional termination site that is 3’ to a putative operon containing a pyridine nucleotide-disulfide family oxidoreductase (PATRIC gene 1570), which may help reduce oxidative stress. The deletion of a single ‘T’ within the T-stretch of the terminator may alter the termination efficiency though the effects of this mutation remain unknown.

A second mutation observed in this strain, though initially designated as intergenic based on the PATRIC reference genome, may in fact lie in a putative selenocysteine synthase homologue of selA in contig 778 (Table 4.4), a gene frequently associated with oxidative stress defense, resulting in mutation A178T (99% amino acid identity with selenocysteine synthase from E. faecalis strain ATCC 4200, accession NZ_GG670377.1). It is possible that this gene may not have been previously identified.
since its putative start codon lies inside gene 1753 (Metallo-dependent hydrolase C subgroup B), and thus the two genes overlap. In other Gram positive organisms, *selA* encodes a Sec synthase that is part of a biosynthetic pathway for the formation of selenocysteine\(^9\). Selenoprotein synthesis in Gram-positive bacteria is found in anaerobes and is often associated with antioxidant defense. While a potential role for selenoproteins in enterococci remains unclear, this mutation may also be relevant to cellular redox and oxidative stress. Modulation of the oxidative stress response pathways could be a potentially important and general response to DAP-induced damage in enterococci that follows changes to *liaFSR* signaling.

The mutation in contig 778 also seems to be present at an unusual and fluctuating frequency. In all TDR strains listed in Table 2, 50% of sequencing reads showed the mutation (G), while the remaining 50% showed the WT sequence (A). Sanger sequencing revealed that strain S613 also had both an A and a G at nearly equal frequency at this position. As sequenced samples were derived from single colonies, this would suggest an extraordinarily high mutation rate at this position. WGS reads from TDR28, however, show 83% ‘G’ at this position, and only 17% ‘A’. This at first led us to conclude that the high mutation rate at this position was biased towards ‘G’ in this colony.

To investigate further, TDR28 was streaked across a plate, and six single colonies were isolated and Sanger sequenced at this position. Curiously, three colonies showed a ‘G’ at this position, while the remaining three showed a mixture of ‘G’ and ‘A’. Passage and resequencing of these six colonies across a subsequent plate showed these genotypes to remain stable. In summary, S613 and other TDR strains contain a 50/50 mixture of ‘G’ and ‘A’ at the mutation site in contig 778, while strain TDR28 seems to contain two
genotypic subtypes: the first contains a 50/50 mixture of ‘G’ and ‘A’, while the second subtype has a ‘G’ at the mutation site. The potential reasons for the high mutation rate at this locus remain obscure but since the founder S613 has comparable heterogeneity it is not likely to be associated with DAP resistance.

4.4 Whole genome sequencing of mixed populations from days 3, 10 and 20.

To identify intermediary mutations that may have gone extinct by the end of the turbidostat experiment, three population mixtures from days 3, 10, and 20 were outgrown overnight in LBHI broth and sequenced. While no mutations occurring above 5% population frequency were observed, discrepancies in population frequencies of known mutant alleles between population mixtures and FREQ-Seq data leads us to believe that outgrowth of mixed populations in the absence of antibiotics has enriched WT alleles in these samples. Our observation that outgrowths of frozen intermediates from our experiments can lead to misleading results is an important caveat to experimental evolution studies, particularly those using strong selective conditions. It is clear that outgrowth selection conditions (e.g. DAP concentration) will favor particular strains, and therefore may not accurately represent the population frequencies at the time the sample was taken. Additionally, it is possible that exposure and adaptation to strongly bactericidal agents associated with membrane stress may lead to substantially altered long-term storage stability and sensitivity to freeze-thaw. For these reasons, all our genomic DNA preparations for FREQ-Seq experiments (discussed in Chapter 5) were prepared directly from the population and not from outgrowths of stored samples. While no novel mutant alleles were identified at these intermediary time-points, the possibility
of their presence in the population at low frequency cannot be excluded conclusively. Nonetheless, our observations are consistent with initial transient responses to DAP in the population (such as transcriptional regulation) rather than extinct genotypes.

4.5 **Observation of non-canonical contingency loci.** Unexpectedly, we found that several mutations in *cls* and *liaF* evolved through insertions or deletions of entire codons in repeat sequences, suggesting that these sites may be contingency loci. Facilitated by such short nucleotide repeats, contingency loci are localized regions of hyper mutability, which mediate “high frequency, reversible, genotypic switching” \(^{100}\). In one such case in *cls*, a direct repeat encoding NFQNFQ\(_{74-79}\) in the ancestor S613 was mutated to NFQ in 42% of the final population (*Table 4.1*), suggesting that slipping along this repeat may be quite frequent. In addition to the NFQ deletion, we also observed a duplication of AII\(_{14-16}\) in *cls* in TDR19.2 (*Table 4.2*). Like *cls*, adaptive changes to *liaF* occurred from an insertion or deletion of an entire codon in a repeat of four Ile codons from positions 177-182.

While short contingency loci that can slip by codons within direct repeats appear to be quite common to *E. faecalis*, they do not appear to be present in *E. faecium*, as shown below in Figure 4.2 \(^{10,11}\), suggesting mechanistic differences to adaptation in these two closely related species.
Figure 4.2. Contingency loci identified in genes cls and liaF in E. faecalis strains.

Mutations of R712, TDR8 and other turbidostat-evolved strains identified these loci as potential contingency loci, with either in-frame deletions or insertions of direct nucleotide repeats. Sequences of strains other than TDR8 were retrieved from PATRIC (http://www.patricbrc.org/portal/portal/patric/Home).

Similarly, the adaptive mutation $gdpD^{AT170}$ in the clinical strain R712 was also a single codon deletion within an isoleucine repeat. Canonical contingency loci such as...
those in *Haemophilus influenzae* produce rapid and reversible phenotypic switching through frameshift mutation-mediate protein deactivation\textsuperscript{100}, but the codon slippage observed in *E. faecalis* is a highly specialized modulation of activity that may produce more nuanced changes in activity. Several genes responsible for CESR in *E. faecalis* have these repetitive and slip prone sequences, suggesting that these “slippery” sequences may have been selected by frequent encounters with the stressors that require facile changes to the cell envelope.
Chapter 5. Allelic Frequency Analysis of Intermediate Populations

5.1 Introduction. As next generation sequencing technology has expanded in the last decade, sequencing developments have enabled researchers to examine microbial communities in increasingly greater detail. In addition to making the ability to sequencing genomes of individual isolates cheaper and more widespread, next generation sequencing has made it possible to sequence mixed populations using an approach known as deep sequencing. Relying on an extremely high degree of sequencing coverage, deep sequencing has been applied in numerous areas, including identification of viruses, tracing bacterial disease determinants, characterizing microbial communities in extreme environments, and facilitating novel transcriptomic analyses.

Recognizing the revolutionary potential of deep sequencing techniques, the United States National Institute of Health (NIH) launched an initiative in 2008 with the goal of characterizing the microorganisms on and within the human body, called the Human Microbiome Project (HMP). Largely through deep sequencing of the 16S ribosomal subunit (commonly used as a taxonomic marker gene), the HMP has worked to establish a “core microbiome” in five major areas (airways, skin, oral cavity, gastrointestinal tract, and vagina), finding a surprising degree of diversity at each site. Although the project is still underway, there is already data linking microbiological abnormalities identified through frequency analysis to diseases states, including autism, colitis, and inflammatory bowel disease.

While most deep sequencing studies have focused on quantifying microbial species, a recent paper by Chubiz et al. details a novel approach for measuring frequencies of specific alleles within a mixed population. Using Illumina sequencing...
technology, this approach (known as FREQ-Seq) is capable of generating hundreds of thousands of sequencing reads across a relatively short genetic space (see section 2.9). By comparing the number of sequencing reads of mutant and wild-type alleles, this approach allows much more precise measurements of allelic frequencies in a polymorphic population compared to other next generation sequencing approaches \cite{111,112}, facilitating investigation beyond the species level.

As discussed in Chapter 4, the end-point population within the turbidostat following adaptation to DAP was highly polymorphic. By measuring allelic frequencies of resistance-associated alleles (listed in Table 4.2), I was able to generate a quantitative metric of their evolutionary success. Additionally, mapping allelic frequencies throughout the course of adaptation has allowed monitoring of the evolutionary dynamics of DAP resistance as genes rise and fall in frequency, providing insight into the process of adaptation.

**5.2 FREQ-Seq analysis and relative success of mutant alleles.** To deconstruct the adaptive network of genomic changes responsible for DAP resistance, we determined the intermediate allelic frequencies of the mutations identified by whole genome sequencing using a high-throughput DNA barcode-based method (FREQ-Seq) \cite{62} (for details see Sections 2.7, 2.8, and 2.9). In all, thirteen mutant alleles were quantified on each of 23 days of turbidostat adaptation. While these 13 alleles encompass the majority of alleles listed in Table 4.2, alleles cls$^{R218Q}$ and cls$^{\text{insAI114-16}}$ were not included, as sequence analysis of strains TDR19.1 and TDR19.2 was not yet completed when FREQ-Seq studies were undertaken. Allele $\text{drmA}^{L.4\text{Stop}}$ was included in the sequencing run,
however sequencing failed for this allele. Frequency measurement results are shown below, in *Figure 5.1.*

**Figure 5.1.** *Evolutionary dynamics of eleven alleles associated with DAP resistance during adaptation.* The frequency of eleven alleles associated with DAP resistance identified from the end of the adaptation (see Section 4.3) were quantitated using DNA barcoding technique FREQ-Seq $^{62}$ (See Sections 5.2, 2.7, 2.8, and 2.9) and plotted as a function of time during the experiment (x-axis) (A) The scale bar on the left indicates a frequency of 50% for a single allele in the population. Note that frequencies correspond to individual alleles, and thus population percentages need not total to 100%. Dashed red lines depict changes in DAP concentration, labeled on upper X-axis. Alleles appearing below 0.5% frequency are not shown. The last frequencies shown are for Day 23 but experiment continued until Day 24.
Using this technique, we are able to follow the rise and fall of mutant alleles at frequencies as low as 0.5%, accurately tracking the evolutionary dynamics of the polymorphic population during adaptation. This, in turn, highlights a number of interesting trends during adaptation.

As shown in Figure 5.2, the first mutations observed within the population fall in the liaFSR pathway. Although prevalent at early stages of adaptation, mutant alleles $liaR^{D191N}$ and $yvlB^{V289fs}$ comprise less than half of the population between days 1 and 13, suggesting that early adaptation was either facilitated through non-genomic changes, such as transcriptional regulation, or early adaptive genomic changes that had little success later in the population and moved to extinction as DAP concentration increased.

Sequencing of population mixtures isolated from days 3, 10, and 20 did not identify any new adaptive alleles at a minimum frequency of 5% (see Section 4.4). This observation is consistent with the hypothesis that transient responses, such as transcriptional regulation, can convey resistance at low DAP concentrations rather than extinct genotypes; however, outgrowth of mixed population sequencing samples makes this observation inconclusive (see Section 4.4). Following day 13, additional liaFSR pathway mutations arise in the population establishing alteration of this pathway as not only an early response to DAP, but also the source of the most prevalent mutant alleles within the turbidostat.

Several liaFSR alleles appear to be negatively correlated with one another, suggesting competition within the population between mutant alleles in this pathway. In Figure 5.1, this is visible between alleles $liaR^{D191N}$ and $yvlB^{V289fs}$ on days 11 and 14; between $yvlB^{V289fs}$ and $liaF^{ΔI177}$ on day 14; and between $liaF^{ΔI177}$ and $liaR^{T194I}$ on days 14-21. Competition between these alleles is supported by the observation that out of the
50 randomly isolated TDR strains from the final day of the experiment (day 24), no strain had more than one mutation in the \textit{liaFSR} pathway (see \textit{Table 4.1}).

Another instance of correlated variation within the population can be observed between alleles $yybT^{I440S}$ and $gshF^{E554K}$, with both alleles rising simultaneously in the population (\textit{Figure 5.1}). We hypothesize these alleles are epistatically linked in strains TDR4 and TDR19.2, and may also be linked to allele $liaR^{D191N}$ (\textit{Table 4.2, Figure 5.1}).

Rigorously, it is possible that the success of alleles $yybT^{I440S}$ and $gshF^{E554K}$ could be attributed to evolutionary hitch-hiking, as linked allele $liaR^{D191N}$ arose first in the population and reached a high frequency in isolation (33\% of population on day 2). However, by day 14 $liaR^{D191N}$ had fallen to a significantly lower frequency (0.6\%) as DAP concentration increased. Upon the appearance of alleles $yybT^{I440S}$ and $gshF^{E554K}$, allele $liaR^{D191N}$ experienced renewed success, rising up to 26.9\% of the population by day 18. Thus, the evolutionary success of $liaR^{D191N}$ in the presence of alleles $yybT^{I440S}$ and $gshF^{E554K}$ suggests that they mediate a fitness benefit.

Proteins YybT and GshF share the common distinction of being responsive to stress induced damage and earlier studies have clearly shown that insertion of the DAP-Ca$^{+2}$ complex into the membrane leads to a loss of ions (such as K$^{+}$) and functional integrity, leading to a significant increase in oxidative stress on the cell.$^{26}$

While many mutations in cardiolipin synthase were observed in the turbidostat, only allele $cls^{ANFQ(74-76)}$ appeared to reach success within the population, rising to a frequency of 44.7\% on day 22 of the experiment. As discussed in section 4.3, this mutation likely increases activity of CLS; the anomalous success of this allele can likely be attributed to an increased localized mutation rate arising from a contingency loci (as
discussed in section 4.5) or a greater increase in activity than other mutant \textit{cls} alleles. This observation highlights the utility of frequency measurements, as they clearly demonstrate the success of this allele over other \textit{cls} mutants.

We noted that as adaptation proceeded, mutations of increasing variety but smaller overall frequency were observed. Mutants \textit{cls}^{R218Q}, \textit{cls}^{ins AII14-16}, \textit{drmA}^{L4stop}, \textit{drmA}^{N150fs}, \textit{mdpA}^{F185fs}, \textit{cls}^{S104L}, and \textit{cls}^{V37A} were all present at less than 5\% of the endpoint population. Generally rising in frequency on the final few days of the adaptation experiment, these mutants may be epistatic with earlier changes or increase fitness very modestly. Thus, these mutations likely only experience success when arising the background of much more significant changes in fitness brought about through alteration of the \textit{liaFSR} pathway.

\textbf{5.3. Inference of strain phylogenies using endpoint genotypes.} By observing when each mutation first appeared in the population (\textit{Figure 5.1}), we compared their order of appearance with the endpoint genotype of each TDR strain (\textit{Table 4.2}) to infer a phylogeny for each lineage, shown below in \textit{Figure 5.2}. Note that lineages shown in \textit{Figure 5.2} are vertically ranked according to biofilm forming propensities of individual strains. Biofilm formation data, and its importance to DAP resistance, is discussed in detail in Chapter 6.
Figure 5.2. Phylogeny inferred from allelic frequency measurements and linkage in endpoint strains. All the successful evolutionary trajectories involve initial changes directly in the liaFSR pathway including a newly identified target operon containing yvlB (see Appendix B). Changes in cls, gshF, or yybT, while potentially important, occur later and are likely epistatic with earlier changes. Line thickness corresponds to success of a specific genotype by the end of the experiment. Strains are vertically ordered based on biofilm forming propensities of TDR strains (see Chapter 6).

Figure 5.2 demonstrates that all successful evolutionary lineages included an initial mutation in the liaFSR signaling pathway. When considered in conjunction with the observation that 94% of endpoint strains carry a mutation in one of these genes (discussed in Section 4.2), it is clear that liaFSR mutations are the most significant DAP resistance determinants identified in this study. These two pieces of data also demonstrate
that the eight whole-genome sequenced TDR strains are representative of the total endpoint population.

Cls mutations were observed across five separate evolutionary trajectories, and always followed mutations that altered LiaFSR signaling (Figure 5.2), suggesting that cls mutations alone were insufficient to confer clinical levels of resistance, and instead may compensate for earlier changes made through the LiaFSR pathway. Indeed, replacement of a wild-type cls allele for a mutated one belonging to a DAP-resistant strain did not affect DAP susceptibility in *E. faecium*\(^{25,90}\). Expression ofCls\(^{R218Q}\) from a multi-copy plasmid in *E. faecalis* OG1RF, however, increased the MIC\(_{DAP}\) suggesting that in the right genomic context, and with potentially higher expression levels,Cls\(^{R218Q}\) could increase resistance\(^{46}\). The same adaptive mutation,Cls\(^{R218Q}\), was also observed in DAP resistant clinical isolates of *E. faecium*\(^10\). The most frequent mutant cls allele observed in the population (Figure 5.1), cls\(^{ΔNFQ(74-76)}\), appears to have arisen independently in the population in at least three different genetic backgrounds (assuming asexuality). This observation suggests either that this particular allele provides a much more significant evolutionary benefit than other cls mutants, or alternatively this position could be a contingency locus thereby increasing its mutation frequency (as discussed in section 4.5).

Although asexuality was assumed during the construction of phylogenies shown in Figure 5.2, there is a possibility that strains in the turbidostat are capable of horizontal gene transfer through conjugation. While capacity for horizontal transfer would require altering the analysis of strain phylogenies in this thesis, such an observation would have both evolutionary and medical relevance considering the history of strains S613 and R712\(^{44}\). Typically conjugation would be expected to be dependent upon an origin of
transfer (oriT) sequence, however an oriT-independent conjugative transfer of genomic DNA has observed in this thesis between *E. faecalis* strains S613 and CK111 (see Appendix E). If DAP-resistant S613 mutants are capable of conjugatively transferring genomic DNA between one another inside the turbidostat, it is possible that mutations such as *cls*\(^ {\Delta NFQ74-76}\) and *drmA*\(^ {N150fs}\) could have each originated in a single strain and spread into other populations through horizontal gene transfer. Although mechanistic details behind oriT-independent transfer of genomic DNA remain unclear, it is possible that high frequency recombination could help mediate such a transfer.

On the last few days of adaptation, a relatively diverse collection of mutations arose across most lineages. As discussed in section 5.2, the low frequencies and late appearance of these mutant alleles suggests they are not significant resistance determinants.

### 5.4 Cellular hierarchy of adaptive changes.

Allelic frequency analysis clearly identified a hierarchy of change that reflects the *in vivo* organizing principles of the cell as it undergoes natural selection. *Figure 5.3*, shown below, illustrates the hierarchy of organizing principles underlying adaptation to DAP applied to the liaFSR system.
**Figure 5.3.** Evolution during strong selection proceeds through a conserved hierarchy of change. In biology, phenotype is the characteristic upon which natural selection is acting (e.g. DAP resistance). Changes bringing about increased fitness are largely changes to cellular ‘systems’, (e.g. liaFSR signaling network). The liaFSR network can, in principle, be activated by any of the proteins that regulate the critical operons, such as LiaF, LiaR, or YvlB. In turn, there are numerous mutations that can increase fitness, and though they may vary in position or substitution, their conserved effect is to alter protein function to modulate liaFSR signaling. Thus the hierarchy of change moves from highly conserved to less conserved during adaptation.

While the phenotype (DAP resistance) is the object of selection, the changes required to bring about adaptation through the LiaFSR system are less constrained and can proceed through a number of genes (liaF, liaR and yvlB). In turn, the number of mutations within a specific gene that can produce the required change in protein function...
can be quite numerous. Though mutations in the liaFSR system are used as an example in Figure 5.3, this idea also applies to cls and drmA in the context of this experiment, as multiple mutations were observed in these loci across different strains. Thus, while eight DAP resistant genotypes were monitored during adaptation, they all followed the same broader evolutionary trajectory. Our results suggest that while specific mutations may vary, the pathways they affect and the phenotypes they confer are well conserved, a finding in good agreement with temperature adaptation studies by Tenaillon et al. in E. coli. Recognition of such general trends can provide an important context for the rational selection of potential drug targets and the design of new strategies for antibiotic development.
Chapter 6. Phenotypic Characterization of TDR strains

6.1 Introduction. Since the discovery of penicillin in 1928, hundreds of antibiotics have been approved for medical use, across several chemical classes. Perhaps not surprisingly, bacteria have evolved in numerous ways to become resistant to the wide array of available antibiotics. While there may be numerous genetic routes to evolving resistance, only four chemical mechanisms of resistance have been identified. One mechanism entails deactivating or modifying the antibiotic itself, with the best example being degradation of β-lactam antibiotics by the enzyme β-lactamase. In a second mechanism, the antibiotic target site is modified in the bacteria; an example of this mechanism is the substitution of D-alanine-D-alanine peptides (the target of vancomycin) with D-alanine-D-lactate, thereby rendering vancomycin ineffective\textsuperscript{113}. A third mechanism of resistance entails modification of metabolic pathways, with the best characterized example being modification of folic acid metabolism to avoid the bacteriostatic effect of sulfonamide antibiotics\textsuperscript{114}. In the fourth mechanism of antibiotic resistance, drug accumulation inside the cell is reduced, typically by decreasing cell permeability or by active efflux of the antibiotic by protein pumps\textsuperscript{115}.

As a recently approved antibiotic, the mechanism of Daptomycin (DAP) action has only recently been characterized (see section 1.3), and while genes associated with DAP resistance have been identified\textsuperscript{11,45,46} (see Chapter 4), the mechanism of drug resistance remains unknown. By studying how resistant cells have phenotypically differentiated from susceptible strains, insight can be garnered as to how resistance is mediated at a cellular level. In turn, a better understanding the cellular mechanism of
DAP resistance would assist in designing a modified drug or an adjuvant to restore its antimicrobial efficacy.

Cellular phenotypes associated with DAP resistance were first reported in *Staphylococcus aureus* in 2009, using a laboratory-derived DAP resistant isolate (evolved using a process similar to the flask adaptation procedure described in section 2.2)\(^{116}\). Membrane analysis of resistant isolates compared to their susceptible ancestor revealed a decrease in membrane fluidity, an increase in membrane lysyl-phosphatidylglycerol (LPG), increased flipping of LPG to the outer membrane, and increased expression of staphylococcal resistance-associated gene *mprF*. As discussed in Chapter 1, phosphatidylglycerol (PG) is the binding target of Daptomycin, facilitating membrane insertion and subsequent cell death. Thus, increased membrane LPG may be associated with a decrease in PG content, as LPG is synthesized via esterification of PG. Synthesis of LPG is mediated in *S. aureus* by *mprF*\(^{117}\), and would be predicted to lead to a decrease in membrane fluidity as LPG is a longer chain lipid. Thus, it is plausible that many membrane alterations in *S. aureus* can be explained by increased *mprF* activity.

A similar study reported in 2009 identified DAP resistance-associated membrane abnormalities in *Bacillus subtilis*, also using a laboratory-derived DAP resistant isolate designated Dap(R)1\(^{23}\). In this isolate, reduced PG content was observed, with genetic studies indicating that mutant phosphatidylglycerol synthase allele *pgsA\(^{A64V}\) leads to this alteration. Morphologically, Dap(R)1 also exhibited aberrant septum placement and thickened peptidoglycan at cell poles, however as numerous mutations were identified in this strain, the genetic determinants of these abnormalities remain unclear.
Most recently, DAP resistance-associated membrane abnormalities have been identified in enterococci as well. In 2011, Arias et al. reported a thickened cell envelope in DAP resistant isolates *E. faecium* R446 and *E. faecalis* R712 compared to their susceptible ancestors, with cell division abnormalities also observed in strain R712, a phenotype similar to that observed in *B. subtilis*. Subsequent lipid analysis using two DAP-susceptible and DAP-resistant clinical strain pairs (*E. faecalis* S613 and R712; *E. faecium* S447 and R446) found a significant increase in glycerolphosphodiglycodiacylglycerol (GP-DGDAG) content, a decrease in membrane fluidity, and a decrease PG in DAP resistant isolates, a finding similar to that observed in *S. aureus*.

In a detailed study using *E. faecalis* strains S613, R712, and a series of allelic replacement mutants (S613$_{\text{liaFI177}}$, S613$_{\text{gdpDI170}}$, and S613$_{\text{liaFI177gdpDI170}}$), Tran et al. found that cardiolipin synthase mutation cls$_{\Delta K61}$ is responsible for reduction of PG content in strain R712, likely linked to increases in cls activity as discussed in section 4.3. Surprisingly, levels of cardiolipin (CL) were unchanged between strains S613 and R712, highlighting the complexity of membrane lipid maintenance. Using fluorescently tagged cardiolipin and DAP molecules, the authors also observed both co-localization of CL and DAP following exposure of cells to the antibiotic, and found that mutant allele liaF$_{\text{All177}}$ was responsible for altered placement of membrane domains containing CL, preventing DAP from binding at the cell division septum.

Together, these observations across three different species suggest a convergent mechanism of DAP resistance mediated through alterations in the cell envelope and membrane. One model put forth by Tran et al. proposes that mutations in the liaFSR system increase DAP resistance by altering lipid domains such that the drug is no longer
able to bind at the division septum; \textit{cls} mutations further increase resistance by reducing the quantity of PG, and possibly through changes in other lipid quantities\textsuperscript{118}. However, this explanation does not account for other phenotypic observations in resistant strains, including thickened cell envelopes and increased biofilm within the turbidostat vessel (see section 2.4). While these irregularities may be unrelated to DAP resistance, it is also possible that they represent additional mechanisms of resistance.

As discussed in Chapter 5, genotypes DAP resistant \textit{E. faecalis} isolates appear to be convergent, with a limited subset of alleles and pathways reaching evolutionary success. This suggests that DAP resistant strains are evolving resistance through a conserved biochemical mechanism leading to convergent phenotypes. To explore this hypothesis, assays quantifying biofilm formation, measuring growth rates as a function of DAP concentration, and examining cell envelope morphology were used.

### 6.2 Increased biofilm production in DAP resistant strains.

During the course of the turbidostat adaptation experiment, a progressive build up of flocculent biofilm matter was observed within the growth chamber. As the association between biofilm formation and antibiotic resistance has been previously established\textsuperscript{119}, increased production of biofilm may represent a general evolutionary route to resistance to DAP. To examine this possibility, biofilm-forming propensity was measured in strains S613, R712, and in sequenced TDR strains (listed in Table 4.2) using an assay adapted from Mohamed \textit{et al.}\textsuperscript{68} (described in section 2.10). The results across five replicate assays are shown below in \textit{Figure 6.1}. 

94
Figure 6.1. Biofilm forming propensity of DAP resistant Enterococcus faecalis strains.

Box plots depict first and third quartile of biofilm formation, with median value shown in thick black line. Colour indicates specific TDR strain. Error bars extending from each box show minimum and maximum range of data. Dotted horizontal lines show the first and third quartiles of biofilm formation for R712. Biofilm formation was normalized to biofilm formation of S613, shown in red horizontal line. Biofilm measurements were conducted and figure assembled by lab member Jiayi Kong.

As shown in Figure 6.1, biofilm formation increased substantially in most DAP resistant strains studied. Despite this clear trend, we could not identify a single allele responsible for the phenotype, but rather combinations of alleles whose net effect is to increase biofilm formation. Turbidostat-derived DAP Resistant (TDR) Strains TDR13
and TDR28 had the most aggressive biofilm forming phenotype and the highest DAP resistance. These strains included changes to uncharacterized proteins *drmA* and *mdpA*, respectively, in addition to *liaF* and *cls* (see Table 4.2). Strains TDR7 and TDR19.1 (containing mutant allele *yvlB*^V289fs) are poor biofilm formers, but are still clinically resistant. Thus, increased biofilm formation itself is not required for attaining clinical levels of resistance *in vitro*, but it is clearly correlated with the majority of the successful evolutionary trajectories.

6.3 Effects of DAP on growth rates. While MIC tests are capable of identifying high levels of resistance, most CLSI approved protocols designed for clinical and hospital settings have a level of variability, as speed of set-up and interpretation are key prioritizes. In a more detailed analysis such as the work described in this thesis, however, use of clinical protocols leads to difficulty in distinguishing differences in drug tolerance among resistant isolates. To more precisely measure the effects of DAP on growth of resistant *E. faecalis* isolates, growth rates of strains S613, R712, and sequenced TDR strains were measured as a function of DAP concentration using a 96-well plate protocol previously described by Walkiewicz et al. ⁶⁷ (see section 2.11 for details). During the course of the experiment, it was observed that strain TDR13 displayed a high degree of variability during attempts to measure cell density within the microtiter plate, which we hypothesize to stem from this strains anomalously high capacity for biofilm formation, and thus this strain has been excluded from growth rate analyses. Growth rate data is shown below, in *Figure 6.2*. 
Figure 6.2. Growth rates of DAP resistant E. faecalis strains as a function of DAP concentration. Growth rates are normalized based on growth at no drug. As described previously, strain TDR13 was excluded, as its growth rate could not be measured due to the high amount of biofilm formation. Average standard deviation was 0.12. Growth rate measurements performed by lab member Jiayi Kong.

Although all strains (aside from S613) shown in Figure 6.2 are clinically DAP resistant, as DAP concentration increased, the capacity of each strain to tolerate the drug varied widely. Analyzing Figure 6.2, we can see at first that, as expected, strain S613 has the lowest tolerance, displaying an inability to grow at 4 µg/ml DAP. Strains TDR19.1 and R712 next show an inability to grow at 16 µg/ml or higher, with TDR19.1 showing a severely impacted growth rate at lower concentrations. At 32 µg/ml, three more strains (TDR7, TDR4, and TDR8) show an inability to grow, with strain TDR7 displaying the greatest inhibition of growth rate at lower concentrations. Surprisingly,
strains TDR22, TDR28, and TDR19.2 were able to tolerate every drug concentration assayed, growing at 32 µg/ml with varying degrees of growth rate inhibition.

6.4 Correlation between biofilm formation and DAP resistance. Comparing growth rates among DAP resistance strains as a function of DAP concentration in section 6.3, it soon became clear that strains producing more biofilm exhibited greater tolerance of DAP. To examine this relationship, biofilm forming propensity was correlated to maximum growth rate for each strain at each concentration of DAP, and is shown below in Figure 6.3.
Figure 6.3. Changes in biofilm formation are correlated to growth rate at increasing concentrations of DAP. Correlation between data from Figure 6.1 and Figure 6.2 are shown. Overall, strains that produce more biofilm have a higher growth rate (e.g. fitness) as DAP concentration increases. Biofilm formation was so strong in strain TDR13 that growth rates could not be measured. TDR13 aggregated so severely that accurate
measurement of cell numbers from optical density in 96-well plates was not possible, as mentioned in text of section 6.3. Correlation performed by lab member Jiayi Kong.

As DAP concentration increases in Figure 6.3, the slope of relationship between growth rate and biofilm formation grows progressively steeper, demonstrating a direct correlation between formation of biofilm and drug tolerance. At 16µg/ml, poor biofilm formers either do not grow at all or have severely inhibited growth. Although there does not appear to be a significant correlation between biofilm formation and growth rate in the absence of DAP (upper left panel of Figure 6.3), this data is likely not of sufficient precision to gauge fitness differences between mutant lines.

To summarize the varying drug tolerance across TDR strains, Figure 5.2 and Figure 6.1 have been combined below in Figure 6.4, ranking lineages according to their degree of biofilm formation (and thereby drug tolerance).
Figure 6.4. Phylogeny inferred from allelic frequency measurements and linkage in endpoint strains highlights biofilm associated lineages. Details of (A) and (B) discussed in Figures 5.2 and 6.1, respectively. Considering the positive correlation between biofilm formation and DAP tolerance (as shown in Figure 6.3), the combined figure shown spatially separates more drug-tolerant lineages (top) from less tolerant lineages (bottom). Combining these two figures allows visualization of both each genotype and its associated level of DAP resistance.

As Figure 6.4 illustrates, specific genotypes (such as that of strain TDR13) are more prone to biofilm formation, whereas other lineages (such as TDR7 and TDR19.1) are significantly less so. Clinically, biofilm formation is strongly linked to pathogenesis via adherence and resistance to phagocytosis in *E. faecalis*-mediated endocarditis and urinary tract infections. While *yvlB* mutations were present in weak biofilm forming strains TDR7 and TDR19.1, they are absent thus far in clinically identified DAP resistant strains of *E. faecalis*, which is consistent with a model wherein biofilm formation results in a combination of resistance and pathogenicity.
6.5 TEM of Membrane alterations observed in DAP resistant enterococci.

As mentioned in section 6.1, previous analysis of clinical isolate *E. faecalis* R712 observed aberrant membrane structures, including septal defects and increased cell wall size using transmission electron microscopy \(^{45}\). To probe for analogous abnormalities in TDR isolates, transmission electron microscopy (TEM) was utilized in a similar manner to collect a series of images, shown below in *Figure 6.5*.

![Figure 6.5](image)

*Figure 6.5. Transmission electron microscopy of strains TDR13, TDR8 and TDR28 show changes to cell wall features.* Relative to ancestral strain S613 (A and B), TEM shows abnormal septation in strain TDR13 (C and D) comparable to clinical strain R712\(^{45}\). By contrast, strains TDR8 (E and F) and TDR28 (G and H) show normal cell division. All TDR strains however showed increased cell wall thickness, another characteristic of strain R712. TEM images were captured by Dr. Truc T. Tran at University of Texas Medical School.
While only a representative subset of TEM images collected are shown in Figure 6.5, in fact all eight end-point strains showed aberrant membrane structures, with septal defects and increased cell wall thickness. These changes in membrane structure varied among the eight strains, with strain TDR13 showing severe septal defects comparable to R712 (Figure 6.5). Compared to S613 (Fig 6.5A-B), representative strains TDR13 (Figure. 6.5C-D), TDR28 (Figure 6.5E-F), and TDR8 (Fig. 6.5G-H) show both aberrant septal structures and multiple septal formations prior to complete cell separation. In TDR28 and TDR8, septal structures form along the cylindrical portion of the cell, with identifiable poles. The most striking changes to septation and cell division are observed in TDR13, with cells lacking a specific polarity. Several septal structures are identified within the same cell and separation of cells is rare.

As mentioned in the caption of Figure 6.5, the mean (±SD) of cell wall thickness was determined for TDR28, TDR8, and TDR13, shown below in Figure 6.6.
**Figure 6.6. Cell wall thickness of DAP resistant E. faecalis strains.** Cell wall thickness increased significantly in all three TDR strains compared to S613 (10.43 ± 1.34 nm in S613 vs 14.8 ± 1.87 nm, 13.7 ± 2.65 nm and 17.3 ± 2.16 nm in TD13, TDR8 and TDR28, respectively, all p < 0.001 in pairwise Student’s t-tests). Changes in cell wall thickness are comparable to those observed in strain R712\(^{45}\). Thickness measurements were performed and figure prepared by Dr. Truc T. Tran at University of Texas Medical School.

The observed changes in morphology are similar to those previously observed in DAP resistant clinical isolate R712\(^{45}\). Together these results confirm the important alterations of cell division associated with development of DAP resistance in enterococci, and demonstrate that a variety of genotypes converge on a common morphological phenotype in DAP resistant *E. faecalis*. 
Chapter 7. Conclusions and Future Studies

7.1. Discussion. Despite the clear ecological differences between bacterial infection in a patient and growth inside a turbidostat, this thesis has demonstrated that evolution of DAP resistance follows the same broader evolutionary trajectory across both environments. Examining the phylogenies of the DAP resistant strains shown in Figure 5.2, evolutionary routes to DAP resistance appear highly constrained. Out of over 3,000 possible genes in E. faecalis strain S613, mutations were only observed in 8 genes inside the turbidostat, representing less than 0.3% of available sequencing space. Clinical isolate R712 demonstrates that mutations in these genes are conserved in a clinical environment as well, as this strain also exhibits mutations in cls and liaF (see Table 4.2).

Broadening our scope across other strains and species, Palmer et al. also observed mutations in cls, yvlB (denoted EF1753), and drmA (denoted EF1797) in E. faecalis strain V583 using a flask-based adaptation procedure. In E. faecium, Munita et al. demonstrated the importance of mutations to liaFSR across multiple clinical isolates, while Humphries et al. identified mutations in genes pspC (a member of the yvlB operon) and cls in two E. faecium strains.

Perhaps more strikingly, mutations in the same codons (liaF\textsuperscript{All\textsuperscript{77}} and liaF\textsuperscript{ins\textsuperscript{177}}) were observed not only in turbidostat strains (see Table 4.2), but also in clinical isolate R712 and multiple independent flask adaptation experiments (see Table 3.1). Similarly, the same mutant allele cls\textsuperscript{ANFQ(74-76)} was also observed in three different genetic backgrounds in the turbidostat (see Table 4.2), a solid-surface flask transfer experiment (see Table 3.2), and in a liquid flask transfer experiment in strain V583. As discussed in section 4.5, the probability of observing common mutations across multiple cell lines
may be increased by contingency loci, which provide localized increases in mutation rate.

While the number of permutations open to a genome is vast, it is clear from our work and others that adaptation to strong selection can follow reproducible paths despite diverse environmental conditions. However, in contrast to other studies the turbidostat-based quantitative experimental evolution approach not only allows identification of adaptive mutations, but can also elucidate their order and potential epistatic relationships.

Through allelic frequency studies of intermediary populations, it is clear that the path to resistance primarily clusters into two steps; changes to liaFSR signaling followed by alterations to cls. As shown in Figure 5.2, across each TDR strain a mutation in the liaFSR system was the first to occur, followed in most cases by a mutation in cls. Across the population as a whole, mutant alleles liaF^{A177}, yvlB^{V289fs}, liaR^{D191N}, and cls^{ANFQ74-76} reached the highest frequencies (see Figure 5.1), quantitatively demonstrating their evolutionary success. Clinical DAP resistance in strain TDR7 (carrying only mutant allele yvlB^{V289fs}) as well as allelic replacement strain S613 liaF^{A177} (carrying only mutant allele liaF^{A177}) show that alterations in the liaFSR pathway are sufficient for DAP resistance; conversely allelic replacement studies in E. faecium have shown that resistance-associated cls mutations do not increase resistance in isolation. These observations are consistent with a model wherein altered liaFSR signaling not only increases DAP resistance, but epistatically facilitates fitness gains from cls mutations.

Although not observed in other organisms or strains of enterococci, mutant alleles in genes yybT and gshF achieved surprising success in the turbidostat population, reaching frequencies higher than 30% (see Figure 5.1). Arising the background of alleles liaR^{D191N} and cls^{ANFQ74-76} (see Figure 5.2), the frequencies of alleles yybT^{4440S} and...
$gshF^{E554K}$ appear linked to one another in the population, suggesting epistasis (see Figure 5.1). As discussed in section 4.3, mutations in these genes may represent a novel route to DAP resistance through modulation of oxidative stress response. For each DAP resistance associated gene identified in this study, the order in which they were observed as well as their putative cellular mechanisms is summarized below, in Figure 7.1.

![Diagram](image)

**Figure 7.1. Timeline of adaptation to DAP within the cell of E. faecalis.** Order and cellular context of adaptive alleles is shown above. Putative localization and functions were ascribed based on homologues as discussed in text and from bioinformatic analysis. Question marks denote unknown function. Proposed functions of liaFSR components are based on homologous proteins previously studied in Streptococcus mutans \(^{89}\). PG: phosphatidylglycerol, CL: cardiolipin, GSH: glutathione.

As Figure 7.1 depicts, mutant alleles tend to cluster into three groups based on their order of occurrence. Although putative mechanisms are proposed here for many
resistance-associated proteins, several still have an unknown function or role in resistance, leaving room for future studies as discussed in section 7.2.

Observations of a reproducible, ordered path to resistance are in good agreement with previous studies of evolutionary constraint by Weinreich et al.\textsuperscript{51}. Beginning with five point mutations in the gene \textit{TEM} previously shown\textsuperscript{124} to mediate a high degree of resistance to antibiotic cefotaxime, the authors engineered every possible permutation of these mutations, resulting in 32 total mutants. After measuring cefotaxime resistance for each mutant, Weinreich et al. determined the selective accessibility of each route to resistance, assuming that neutral or deleterious mutations have no chance of population fixation. Surprisingly, among 120 total potential routes to the resistance, the authors found 102 to be selectively inaccessible\textsuperscript{51}, demonstrating the limited number of evolutionary trajectories available during adaptation. As this work has identified relatively few alleles associated with DAP resistance, it is likely that adaptation to DAP can similarly be reached through a limited number of evolutionary paths.

In addition to identifying a conserved order of mutations leading to DAP resistance, this work also identified a hierarchy of adaptation reflecting \textit{in vivo} organizing principles of the cell as it undergoes selection. By comparing protein function across mutant alleles, this study identified convergent evolutionary trajectories across numerous genes, relatively few pathways, and in nearly every case involving cell membrane chemistry. As depicted in \textit{Figure 5.3}, the most striking example of this observation was in the \textit{liaFSR} system: across five mutant alleles, only three genes were affected, falling in a single pathway regulating cell envelope stress response. Extending this observation
across other studies of DAP resistance, a total of 15 mutations have been identified in the liaFSR signaling system in enterococci, shown below in Table 7.1.

Table 7.1. **DAP-resistance associated liaFSR mutations identified in enterococci.**

<table>
<thead>
<tr>
<th>Allele</th>
<th>Ancestral Strain</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>liaF&lt;sup&gt;<code>A1177</code>&lt;/sup&gt;</td>
<td>E. faecalis S613</td>
<td>This work, Arias et al</td>
</tr>
<tr>
<td>liaF&lt;sup&gt;<code>T194I</code>&lt;/sup&gt;</td>
<td>E. faecalis S613</td>
<td>This work</td>
</tr>
<tr>
<td>liaF&lt;sup&gt;<code>L39F</code>&lt;/sup&gt;</td>
<td>E. faecium clinical isolate</td>
<td>Arias et al</td>
</tr>
<tr>
<td>liaF&lt;sup&gt;<code>L27I</code>&lt;/sup&gt;</td>
<td>E. faecium clinical isolate</td>
<td>Arias et al</td>
</tr>
<tr>
<td>liaF&lt;sup&gt;<code>A180T</code>&lt;/sup&gt;</td>
<td>E. faecalis S613</td>
<td>Arias et al</td>
</tr>
<tr>
<td>liaF&lt;sup&gt;<code>E192G</code>&lt;/sup&gt;</td>
<td>E. faecium clinical isolate</td>
<td>Munita et al</td>
</tr>
<tr>
<td>liaF&lt;sup&gt;<code>T120A</code>&lt;/sup&gt;</td>
<td>E. faecium clinical isolates</td>
<td>Munita et al, Arias et al</td>
</tr>
<tr>
<td>liaF&lt;sup&gt;<code>H264Q</code>&lt;/sup&gt;</td>
<td>E. faecium clinical isolate</td>
<td>Munita et al</td>
</tr>
<tr>
<td>liaR&lt;sup&gt;<code>D191N</code>&lt;/sup&gt;</td>
<td>E. faecalis S613</td>
<td>This work</td>
</tr>
<tr>
<td>liaR&lt;sup&gt;<code>W73C</code>&lt;/sup&gt;</td>
<td>E. faecium clinical isolates</td>
<td>Munita et al, Arias et al</td>
</tr>
<tr>
<td>yvlB&lt;sup&gt;<code>V298fs</code>&lt;/sup&gt;</td>
<td>E. faecalis S613</td>
<td>This work</td>
</tr>
<tr>
<td>yvlB&lt;sup&gt;<code>I390fs</code>&lt;/sup&gt;</td>
<td>E. faecalis V583</td>
<td>Palmer et al</td>
</tr>
<tr>
<td>pspC&lt;sup&gt;<code>229fs</code>&lt;/sup&gt;</td>
<td>E. faecium clinical isolate</td>
<td>Humphries et al</td>
</tr>
</tbody>
</table>

Mutations listed above were identified either in this work, or in previous studies by Arias et al. 45, Munita et al. 47, Palmer et al. 46, or Humphries et al11

The number of independent mutations in Table 7.1 shows that evolution through this pathway has remained constant across a majority of DAP resistant strains, likely rendering the same function on a cellular level.

Mutations in gene cls also demonstrate this conserved hierarchy of mutations. Across this study and others in E. faecium, a total of eight distinct mutations in cardiolipin synthase have been associated with DAP resistance; identifying alleles cls<sup>`ANFQ74-76`</sup>, cls<sup>`R218Q`</sup>, cls<sup>`S104L`</sup>, and cls<sup>`ins AIH14-16`</sup> in this work, and alleles cls<sup>`H215R`</sup>, cls<sup>`N12S`</sup> 11.
$cls^{N131}$, and $cls^{ins110NPL}$. Kinetics studies of two purified mutant proteins from *E. faecium* (H215R and R218Q) have found increased activity in both cases, a finding that is likely the case for other Cls mutants as well. As discussed in Chapter 6, increased Cls activity would lead to reduced concentrations of lipid substrate PG, reduction of which has been linked to DAP resistance in strains of *S. aureus*, *B. subtilis*, and enterococci. Thus, on a cellular level reduced PG content is conserved across many DAP resistant strains, demonstrating a highly conserved evolutionary response on a cellular level.

Compared to mutations in the *liaFSR* system and *cls*, the relatively scarce frequencies of mutant alleles arising in the turbidostat after day 20 of the experiment suggest they mediate less of a fitness benefit (see *Figure 5.1*). However, there does appear to be one example of evolutionary convergence in the case of allele *drmA*. As shown in *Figure 5.2*, two separate evolutionary trajectories carry mutations likely disabling the function of this protein (alleles $drmA^{L4Stop}$ and $drmA^{N150fs}$); in a separate evolution experiment by Palmer *et al.* another frameshift mutation was also observed (allele $drmA^{D30fs}$). These separate observations of frameshift and stop-codon mutations suggest convergent inactivation of DrmA protein function, however as the function of this protein is currently unknown, the effects on a cellular level are unclear.

Independent of genetic data, evolutionary trajectories appeared to be highly conserved on a phenotypic level in DAP resistant strains. As shown in *Figure 6.1*, propensity for biofilm formation was increased in a majority of TDR strains, as well as clinical strain R712, with growth rate data indicating a strong correlation between biofilm formation and DAP tolerance (*Figure 6.3*). Microscopy studies also found increased cell
wall thickness in strain R712 \(^{45}\) and all sequenced TDR strains, with several resistant strains also exhibiting cell division abnormalities (Figure 6.5). Together, these results demonstrate that a variety of DAP resistance associated genotypes converge on a small number of common morphological and phenotypic traits in \(E. \text{faecalis}\), which likely carry a fitness cost in the absence of DAP.

Observations of hierarchical evolutionary convergence in this thesis agree well with previous studies by Tenaillon \textit{et al.} \(^{56}\) on populations of \(E. \text{coli}\) adapting to higher temperature. Using laboratory-adapted \(E. \text{coli}\) strain B REL1206, the authors adapted 115 independent lines to grow at 42.2ºC each for ~2,000 generations, sequencing the genome of one colony from each adapted line. Although a total of 1258 mutations were identified, 668 of these mutations fall into 12 functional units, among which the most mutations were observed within the extended RNA polymerase complex \(^{56}\). Had Tenaillon \textit{et al.} confined their focus to looking at individual genes, this degree of convergence would not be apparent, as the largest number of mutations in an individual gene was only 87, observed in \(rpoB\). When considering gene function and grouping them within the context of their cellular mechanisms, however, evolutionary convergence rapidly becomes evident. The findings of Tenaillon \textit{et al.} in addition to those of this thesis demonstrate that while the number of mutations within a specific gene that can produce the required change in protein function can be quite numerous, the number of proteins, pathways, and cellular mechanisms involved are relatively few in comparison.
7.2. Future work

As the work described in this thesis demonstrates, quantitative experimental evolution has been used to not only elucidate genetic determinants of resistance, but also to quantitatively assess their significance both through frequency analysis and convergence across subpopulations. By combining the strengths of experimental evolution and genomic analysis of intermediary populations, this approach provides important insights into drug resistance that are hard to glean from genomic or clinical data alone, allowing researchers to move from a reactive approach to a strategy that can anticipate resistance in a predictive manner. Although used here to study adaptation of *E. faecalis* strain S613 to daptomycin, quantitative experimental evolution could be used for any strain or medical isolate, including those not easily amenable to traditional genetic techniques.

Currently, quantitative experimental evolution is being applied in the Shamoo lab in two instances: Kathryn Beabout is using this procedure to explore adaptation of *E. faecalis* to Tigecycline, another frontline antibiotic, while Dr. Troy Hammerstrom is studying adaptation of gram-negative pathogen *Acinetobacter baumannii* to Tigecycline. Taken together with the work described in this thesis, these experiments will provide a powerful demonstration of the utility of quantitative experimental evolution across different antibiotic stressors and organisms.

One of the primary strengths of quantitative experimental evolution lies in its capacity to prioritize the significance of resistance-associated proteins. This prioritization allows further study to be focused upon more significant proteins, providing new targets for drug design, crystallographic analysis, and biochemical studies. Among
the proteins linked to DAP resistance in this thesis, Cls has already been the subject of biochemical study, however multiple proteins in the enterococcal liaFSR pathway are yet to be fully characterized. As of the time of this writing, Shamoo lab member Dr. Milya Davlieva has successfully expressed and purified proteins LiaR and YvlB, and crystallographic studies of each are currently underway. DNA binding studies using LiaR, along with protein binding studies using YvlB, are also being employed to better understand the function of each protein and identify novel components of the liaFSR regulon. A better understanding of proteins YvlB and LiaR will in turn give more insight into how DAP resistance is mediated on a molecular level, and identification of new components of the liaFSR regulon will provide additional targets for proteomic study. Thus, identification of the liaFSR system in this thesis as the primary DAP resistance determinant in *E. faecalis* helps to underscore the significance of Dr. Davlieva’s current line of research.

Although genes *gshF* and *yybT* have not previously been associated with DAP resistance, frequency analysis within the turbidostat found that mutant alleles in these genes were present at high frequency during adaptation to DAP (see *Figure 5.1*). Though bioinformatic analysis has identified *E. faecalis yybT* as a putative homologue of *yybT* in *B. subtilis* and *S. aureus*, as a protein involved in signaling the details of YybT’s function in *E. faecalis* remain unclear. Currently Shamoo lab member Xu Wang is working to both elucidate the function of each domain of this protein as well as crystallize it to determine its structure. Conversely, GshF has known function that has been previously characterized, but the genetic linkage between alleles *gshF^{E554K}* and *yybT^{I440S}* raises the possibility of an interaction between GshF and YybT in evolution of DAP resistance.
Characterization of these two proteins has clinical as well as biochemical significance, as DAP-resistance associated mutations in YybT and GshF could represent a novel route to resistance.

Proteomic analysis could also be expanded to other proteins identified as being associated with DAP resistance, such as DrmA. Despite the low frequency of occurrence of mutant alleles in this gene (>5% of the population, see Figure 5.1), multiple convergent mutations disabling function of this gene were observed in the turbidostat, as well as an independent flask-based adaptation performed by Palmer et al. DrmA mutations were observed in the two TDR strains most strongly associated with biofilm formation (TDR13 and TDR28), suggesting that mutations through this gene could play a role in mediating high levels of resistance. As a gene of unknown function with no characterized homologues, information gleaned about the role of DrmA would likely prove novel, however identifying the function of this protein might prove difficult.

Previously studied in streptococci and bacilli, modifications to the liaFSR pathway were identified in this work as the primary DAP resistance determinant in E. faecalis. While data from other organisms suggests that increased liaFSR activity increases DAP resistance, it has not yet been demonstrated that DAP resistance-associated liaFSR mutations lead to such an increase in E. faecalis despite the work described in Appendices C and D. As discussed in Appendix D, the lack of fluorescent signal from both strain CK111 and TDR strains carrying pCAM10 suggests that GFP is not a sufficiently sensitive reporter to detect signal expression from the liaFSR promoter. One approach to increasing signaling would be to clone T7 RNA polymerase followed by the T7 RNA polymerase promoter sequence in between the liaFSR promoter and GFP on
the plasmid. Under this design, activation of the liaFSR promoter would drive expression of T7 RNA polymerase, a small amount of which would in turn drive higher expression of GFP, thereby increasing signal.

Another possible approach would be to use the liaFSR promoter to drive expression of lacZ, encoding β-galactosidase, as this enzyme is a more sensitive reporter than GFP. However, as lacZ is currently encoded on plasmids pCAM10 and pCAM11 to ensure proper plasmid incorporation, lacZ would have to be removed from its current position on each plasmid and a different technique would have to be used to verify plasmid presence (such as PCR).

In addition to quantifying activity of the liaFSR system, a reporter plasmid would also allow differentiation between liaFSR activating and inactivating conditions in E. faecalis. This would in turn enable transcriptomic studies of the liaFSR system in enterococci, using techniques such as RNA-Seq or microarray analysis to detect changes in regulation when the system is active. Identifying genes up-regulated during periods of liaFSR activity could in turn be used to compile a more comprehensive liaFSR regulon, which would not only allow a more thorough understanding of a crucial cell envelope stress response system, but would generate additional targets for proteomic analysis and antibiotic drug design.

Although studies using a reporter plasmid have been focused on derivatives of strain E. faecalis S613 in this thesis, genetic rearrangements observed in this species may render this organism intractable for such an approach. As discussed in Appendix D, TDR strains appear both to not replicate plasmids following conjugation and undergo genetic rearrangement in the liaFSR operon. Future studies using genetic techniques would best
be advised to focus on more readily manipulated organisms such as *E. faecalis* strains derived from OG1 (such as strains OG1RF and CK111).

Also discussed in Appendix E is the observation that strain S613 can act as a donor of genomic DNA when conjugated with strain CK111. When constructing phylogenies of TDR strains shown in *Figure 5.2*, cells were assumed to be asexual during adaptation. If this assumption is incorrect, mutant alleles (such as $cls^{ANF?4-76}$) assumed to have arisen in multiple independent backgrounds through separate mutational events may have instead arisen in a single background and subsequently spread between strains through conjugation. Elucidating the capacity of cells of strain S613 to conjugate with one another would not only remove assumptions from the work detailed in this thesis, but also provide useful information for future adaptation experiments using this strain.
Bibliography


34. CDC. Centers for Disease Control and Prevention. (2009).


117. Staubitz, P., Neumann, H., Schneider, T., Wiedemann, I. & Peschel, A. MprF-mediated biosynthesis of lysylphosphatidylglycerol, an important...


Appendices

Appendix A. liaFSR signaling in *Enterococcus faecalis*. As discussed in Chapter 5, allelic frequency measurements of turbidostat populations found that mutant alleles within the liaFSR pathway were the first to be observed as DAP resistance evolved, and were maintained at the highest frequency relative to other mutant alleles. This finding highlights the significance of the liaFSR pathway to resistance, a pathway that remains largely uncharacterized in enterococci.

Predominantly studied in *Bacillus subtilis*, the liaFSR signaling system regulates cell envelope stress response in gram-positive bacteria, analogous to the phage-shock protein (Psp) system in gram-negative bacteria. In *B. subtilis*, six proteins have been identified in the liaFSR system, depicted below in Figure A.1.
Figure A.1. LiaFSR signaling system components in Bacillus subtilis. Signaling system depicted as described by Wolf et al. 87. LiaFSR operon protein functions are depicted where known. Currently, the role of LiaG and the interaction between LiaI and LiaH remain uncharacterized. LiaH is depicted in its active oligomeric form.

Although the function of LiaI and LiaG remain unknown, details have been elucidated for each other component of the liaFSR pathway. Cell envelope stress is
initially detected by transmembrane sensor kinase LiaS, leading to phosphorylation of the response regulator LiaR. Although the mechanism of LiaS activation is unknown, it becomes activated by a variety of cell wall stressors, including cell-wall active antibiotics (such as vancomycin, bacitracin, and ramoplanin), lysozyme treatment, and detergent treatment\textsuperscript{125}. In the absence of stress, protein LiaF inhibits kinase activity of LiaS, preventing activation of the system. LiaF knock-out mutants in \textit{B. subtilis} show that in the absence of this inhibitor protein, the \textit{liaFSR} system remains constitutively active ("Lia ON mutants"\textsuperscript{87}), demonstrating the importance of LiaF in regulating stress response. Following LiaR phosphorylation by LiaS, LiaR becomes active as a response regulator, leading to the transcription of numerous downstream target operons. Although LiaR has been shown to regulate numerous targets, including the \textit{liaFSR} operon itself, proteomic analysis of Lia ON mutants suggests the primary target of this system is \textit{liaH}\textsuperscript{87}. Shown to form a 36-mer in vitro, LiaH becomes greatly upregulated in response to LiaR activity, analogous to protein PspA in \textit{E. coli}. How this 36-mer might interact with the membrane or other proteins to neutralize cell envelope stress, however, remains unclear.

While bioinformatic analysis has identified clear homologues for \textit{liaF}, \textit{liaS}, and \textit{liaR} in sequenced strains of enterococci and streptococci, there does not appear to be a \textit{liaH} homologue, leaving the primary target of the \textit{liaFSR} regulon unclear in these genii. In a paper studying \textit{liaFSR} signaling in \textit{Streptococcus pneumoniae}, strain TMSP008 was grown in the presence of \textit{liaFSR}-inducing antibiotic bacitracin, and total RNA was then collected for microarray analysis\textsuperscript{125}. Surprisingly, transcriptional data did not indicate a
single primary liaFSR target protein, but rather showed moderate up-regulation of numerous proteins (see Eldholm et al. \textsuperscript{125}, Table 2).

Similar to streptococci, currently the only identified liaFSR components in enterococci are liaF, liaS, and liaR. Although evidence presented in this thesis also implicates the operon containing genes yvlD, pspC and yvlB in the liaFSR system, the functions of these proteins, along with the identity of any other putative effector enzyme, remains unknown. Considering the importance of this signaling system to DAP resistance, identification of the liaFSR regulon in \textit{E. faecalis} is key to understanding both the mechanism of resistance, as well as the functions of mutations identified in this pathway.

**Appendix B. Linkage between yvlB and liaFSR.** As detailed in Table 4.2, a DAP-resistance associated mutation was identified in yvlB (yvlB\textsuperscript{V289fs}), a gene of unknown function. As the only mutation observed in DAP resistant isolate TDR7, mutant allele yvlB\textsuperscript{V289fs} appears sufficient for DAP resistance. Gene yvlB is named for a homologue in \textit{B. subtilis} (59\% coverage, 32\% identity) previously identified by whole genome sequencing \textsuperscript{86}, however the structure of the operon containing this gene appears to differ between \textit{B. subtilis} and \textit{E. faecalis} S613. As described in section 2.12, operon analysis using cDNA found that yvlB shares an operon with genes yvlD and pspC in \textit{E. faecalis} strain S613, as depicted on the operon map shown below in Figure B.1.
Figure B.1. Putative yvl operon in *E. faecalis* S613. Genes yvlD and yvlB are named for homologues in *B. subtilis*, whereas pspC is named for homology to *E. coli* protein Phage Shock Protein C. Details of operon assignment are discussed in Section 2.12.

This operon has been linked to DAP resistance in other strains of enterococci as well. In a 2011 study using a laboratory-adapted DAP resistant variant of *E. faecalis* strain V583, Palmer *et al.* also identified a frameshift mutation in yvlB at position Ile-390 (yvlB designated EF1753)\(^{46}\). In a separate study using a clinically isolated pair of *Enterococcus faecium* strains, Humphries *et al.* identified a DAP resistance associated mutation in pspC\(^{11}\).

Cross-referencing the yvl operon with the liaFSR regulon identified by microarray in related organism *S. pneumoniae*\(^{125}\) (discussed above), it would appear that pspC (locus spr0810 in *S. pneumoniae*) is upregulated in the presence of liaFSR stress, suggesting that the yvl operon is regulated by the liaFSR system. This connection suggests that mutations in the liaFSR system and yvl operon, both shown to be sufficient for DAP resistance, likely function through the same mechanism.

Appendix C. Design and cloning of fluorescent reporter plasmid to measure liaFSR activity. As previous studies have shown increased liaFSR activity leads to greater DAP resistance in *B. subtilis*\(^{126}\), we hypothesize that DAP resistant enterococci also exhibit increased liaFSR activity. To test this hypothesis, a fluorescent reporter
A plasmid system was designed to measure liaFSR activity in DAP resistant TDR strains. Although *E. faecalis* strain S613 has been found to not be transformable, bacterial conjugation has been successfully applied to introduce new genetic elements. Using plasmid pHOU3 as a template, plasmids pCAM10 and pCAM11 were designed as shown below in Figure C.1 and cloned using circular polymerase extension cloning (CPEC), a sequence-independent cloning technique.

![Plasmid Maps](image)

**Figure C.1. Plasmid maps of pCAM10 and pCAM11.** Though the two plasmids are largely similar, pCAM10 contains *E. faecalis* strain S613 liaFSR putative promoter sequence upstream of GFP, whereas pCAM11 has no promoter or ribosome-binding site prior to GFP. Genes pheS, lacZ, and chloramphenicol acetyltransferase (CAT) stem from pHOU3 backbone. Gene repA originates from *Lactococcus lactis*, and was amplified from the genome of *E. faecalis* strain CK111 for cloning. Plasmids were
cloning using sequence-independent circular polymerase extension cloning (CPEC) procedure (see section 2.13).

A conserved feature of the liaFSR system is its capacity to drive its own expression through LiaR binding of the liaFSR promoter\(^{87,125}\). As shown in Figure C.1, we have cloned the putative liaFSR promoter from the genome strain S613 upstream of green fluorescent protein (GFP) in plasmid pCAM10 (GFP cloned from genomic DNA of lab strain *E. coli* RU1-GFP). Thus, greater LiaR activity should lead to greater expression of GFP on the plasmid, enabling liaFSR activity to be quantified by measuring fluorescence. By placing this reporter system on a plasmid, it can be conjugated into strain S613 and each TDR strain to dynamically monitor liaFSR activity under a range of growth conditions to characterize how various resistance-linked liaFSR mutations effect signaling. To control for background expression of GFP, a plasmid without a liaFSR promoter was also designed (pCAM11).

Genes *lacZ*, chloramphenicol acetyltransferase (*CAT*), and *pheS* were also present on template plasmid pHOU3\(^{45}\). Chloramphenicol resistance gene *CAT* allows selection for strains with the plasmid during transformation and conjugation (see sections 2.15 through 2.18); similarly, gene *lacZ* allows screening for cells containing plasmid using X-gal as a secondary confirmation of plasmid presence. While gene *pheS* was used as a counterselection marker in previous studies\(^{45}\), the function of this gene was not utilized in this work.

As Arias *et al.* detail\(^ {45}\), plasmid pHOU3 is itself constructed from plasmid pCJK47, designed by Kristich *et al.* to function as a conjugative donor plasmid in *E.*
faecalis strain CK111. While conjugation of this plasmid is required for insertion into strain S613 and TDR strains, Kristich et al. designed their system to require gene repA from Lactococcus lactis in trans for plasmid replication, as the plasmid contains an origin of replication originally from lactococcal plasmid pWV01. In their conjugative system, repA is present in the chromosome of strain CK111, enabling plasmid replication in donor strain CK111 but not in recipient strains. Thus, to enable plasmid replication in S613 and TDR strains, repA was cloned from the genome of strain CK111 into the backbone of the plasmids, as shown in Figure C.1.

Appendix D. Conjugation and genetic rearrangements of TDR strains.
Constructs described in Appendix C were chemically transformed into E. coli strain EC1000, an E. coli host with a chromosomal copy of L. lactis repA, and purified (see sections 2.14 and 2.15). Purified plasmids pCAM10 and pCAM11 were then transformed into competent cells of strain CK111 through electroporation (see sections 2.16 and 2.17). Following electroporation, cells were plated on chloramphenicol to select for transformants, and were subsequently screened for β-galactosidase activity to ensure retention of the plasmid.

Strain CK111 cells containing plasmid were then conjugated with strains S613, R712, and each TDR strain (see section 2.18). Cells were selected on both chloramphenicol (to ensure presence of plasmid) and vancomycin (to select against donor strain CK111), and were subsequently screened for β-galactosidase activity. Several strains were observed to have no β-galactosidase activity despite their chloramphenicol resistance, suggesting that a portion of the plasmid had been lost during conjugation.
However, chloramphenicol resistance and β-galactosidase activity were observed after conjugation with both plasmids in strains S613, R712, TDR7, TDR13, TDR22, and TDR28, suggesting presence of plasmid in these strains.

To measure activity of the liaFSR system in the presence of DAP, the strains described above (in addition to strain CK111 carrying both plasmids) were grown on both solid LBHI media and LBHI media containing DAP at half the MIC value of each strain. Cells were then resuspended, and a fluorescent microscope was used to measure GFP fluorescence (see section 2.19), using E. coli strain RU1-GFP as a positive control and S613 without plasmid as a negative control. However, no fluorescence was observed in any strains apart from the positive control.

To better understand this absence of signal, plasmid replication was examined by first purifying genomic DNA from each of the conjugated strains listed above, as well as strains of CK111 carrying pCAM10 and pCAM11. Primers binding within GFP designed to amplify the entire plasmid were then used in PCR reactions with two samples of CK111 containing pCAM10, S613 containing pCAM10, and TDR13 containing pCAM10. Products were then run on a 1% agarose gel, shown below in Figure D.1.
Figure D.1. PCR Amplification of pCAM10 DNA in E. faecalis. All template DNA from strains either transformed or conjugated with plasmid pCAM10. Lanes from right to left: lane 1: ladder; lane 2: CK111 gDNA, lane 3: CK111 gDNA (duplicate sample); lane 4: S613 gDNA; lane 5: TDR13 gDNA.

While a band at the expected weight of pCAM10 was observed in both gDNA samples from strain CK111, surprisingly no bands were observed for strains S613 and TDR13. This result suggests that plasmid DNA is not present in strains S613 and TDR13, despite their chloramphenicol resistance and β-galactosidase activity. To confirm this result across other strains, this PCR was repeated across additional samples, shown below in Figure D.2.
**Figure D.2. PCR Amplification of pCAM10 DNA in E. faecalis and E. coli samples.**

All template DNA from strains either transformed or conjugated with plasmid pCAM10. Lanes from right to left: lane 1: ladder; lane 2: TDR13 miniprep DNA; lane 3: S613 miniprep DNA; lane 4: TDR22 gDNA; lane 5: TDR13 gDNA; lane 6: TDR7 gDNA; lane 7: E. coli strain EC1000 miniprep DNA.

As the results shown in *Figure D.2*, demonstrate, it would appear that while plasmid pCAM10 can be amplified in strains *E. faecalis* CK111 and *E. coli* EC1000, plasmid does not appear to be present in strains S613, TDR13, TDR22, and TDR7. As all the aforementioned strains exhibit chloramphenicol resistance and β-galactosidase activity, incorporation of plasmid DNA into the genomic DNA (gDNA) of S613-derived strains would be consistent with the data above. As the *liaFSR* promoter site is the only region of plasmid pCAM10 that has homology to genomic DNA in S613-derived strains,
this site is the most likely position for recombination incorporating the plasmid into the genome.

To look for recombination at this position, a PCR reaction was designed using a forward primer located within the *liaFSR* promoter (Plia_F) and a reverse primer located within *liaF* (liaF_R). In a wild-type background, the expected product of this reaction is approximately 1 kilobase (kb), however should the plasmid have recombined into the genome at this position, the PCR product would be longer. As plasmid pCAM11 lacks the *liaFSR* promoter sequence, strains containing pCAM11 would not be expected to recombine at this position. This reaction was run using genomic DNA from TDR13 containing pCAM11, TDR13 containing pCAM10, S613 containing pCAM10, and wild-type S613, shown below in *Figure D.3*. 

![PCR gel image](image-url)
Figure D.3. PCR product of region between liaFSR promoter and liaF. Template gDNA from strains conjugated with either pCAM11 or pCAM10, while all PCR reactions were run with primers Plia_F and liaF_R. Lanes from left to right used the following template DNA: lane 1: ladder; lane 2: TDR13 containing pCAM11; lane 3: TDR13 containing pCAM10; lane 4: S613 containing pCAM10; lane 5: wild-type S613 (no plasmid).

While gDNA from TDR13 with pCAM11, S613 with pCAM10, and wild-type S613 showed the expected product at 1 kb in Figure D.3, product amplified from TDR13 containing pCAM10 showed a ~2.5 kb product. This would suggest an insertion of ~1.5 kb of DNA at this position, which would be consistent with insertion of a fragment of pCAM10 plasmid. As no insertion was observed in strain S613 containing pCAM10, this PCR reaction was repeated across additional samples, shown below in Figure D.4.
Figure D.4. **PCR product of region between liaFSR promoter and liaF across additional samples.** All template gDNA from strains conjugated with pCAM10, while all PCR reactions were run with primers Plia_F and liaF_R. Lanes, from left to right: lane 1: ladder; lane 2: TDR7; lane 3: TDR22; lane 4: TDR28; lane 5: R712.

As Figure D.4 shows, PCR product using template gDNA from TDR7 and TDR22 carrying pCAM10 again showed the larger-size (~2.5 kb) product. Unexpectedly, these samples also showed the wild-type band at ~1 kb, as well as a fainter band of ~1.5 kb.

To elucidate the identity of the inserted segment between the liaFSR promoter and liaF, the PCR product amplified from TDR13 carrying pCAM10 (shown in Figure D.3 lane 3) was Sangar sequenced using primer Plia_F. Although an inserted segment of DNA was identified in this PCR product, surprisingly it was inserted in the intergenic space after liaF (upstream of liaS), and did not match DNA from plasmid pCAM10. BLASTN results show that a 238 nucleotide portion of this insert matches sequences previously observed on four *E. faecium* plasmids (pF856, pNB2354, pS177, and p5753cA), as well as a putative *E. faecium* transposon. When BLASTN searches are limited to the genome of either *E. faecalis* strain S613 or strain OG1RF, however, no results are returned, suggesting a foreign origin of this region (sequencing results shown in Appendix F). While the origin of this inserted DNA segment remains unclear, it is highly likely that such a large insertion disrupts function of the liaFSR operon thereby interfering with any results the reporter system might have gleaned.
Appendix E. Conjugation of genomic DNA between *E. faecalis* strains

**CK111 and S613.** As discussed in Chapter 5, knowledge of a bacterial strain’s capacity for sexual recombination is integral when inferring an evolutionary phylogeny. Assembling phylogenies as strain S613 adapts to DAP, this strain was assumed to be asexual throughout adaptation. To test this assumption, strain S613 was conjugated with strain CK111 (see section 2.18) in the absence of a plasmid to test the capacity of these strains to undergo sexual recombination. Following conjugation, cells were selected on both spectinomycin and vancomycin, and four strains resistant to both antibiotics were isolated, labeled “conjugant #4”, “conjugant #5”, “conjugant #14”, and “conjugant #16”.

As the determinants for spectinomycin resistance are thought to be chromosomally encoded in strain CK111\textsuperscript{127} and PATRIC genome files show the cassette encoding vancomycin resistance lies on the chromosome of strain S613, we hypothesize that during conjugation, either strain CK111 donated chromosomal spectinomycin resistance determinants to strain S613, or strain S613 donated vancomycin resistance determinants to strain CK111. As *L. lactis* gene *repA* has been previously inserted into the genome of strain CK111\textsuperscript{127}, presence or absence of this gene can identify the origin of conjugant strains listed above. Thus, *repA* was amplified using PCR from gDNA isolated from each strain and products were run on an agarose gel, shown below in *Figure E.1*. 


---

\textsuperscript{127} Liao et al. (2017)
Figure E.1. Amplification of *L. lactis* repA in *E. faecalis* strains. Genomic DNA template was used for each PCR reaction with primers amplifying gene repA. Lanes, from left to right: lane 1: ladder; lane 2: CK111; lane 3: S613; lane 4: conjugant #4; lane 5: conjugant #5; lane 6: conjugant #14; lane 7: conjugant #16.

As Figure E.1 shows, there is a band for *L. lactis* repA in strain CK111, as well as all conjugated strands, that is absent in strain S613. This data shows that all conjugated strains possess chromosomal DNA from strain CK111 unrelated to the selective event, suggesting that strain S613 acted as a donor strain, transferring chromosomal vancomycin resistance determinants to recipient strain CK111. As the vancomycin resistance operon has been identified in *E. faecalis* strain S613 using whole genome sequencing, PCR primers were designed to amplify a portion of this cassette (VAN cassette). These primers were used in a PCR reaction with gDNA from each conjugated strain as well as S613 and CK111, shown below in Figure E.2.
**Figure E.2. PCR amplification of VAN cassette in E. faecalis strains.** Genomic DNA template was used for each PCR reaction. Lanes, from left to right, used gDNA template from the following strains: lane 1: ladder; lane 2: CK111; lane 3: S613; lane 4: conjugant #4; lane 5: conjugant #5; lane 6: conjugant #14; lane 7: conjugant #16.

Although Figure E.2 shows non-specific amplification products from each genomic DNA sample, there is a clear band at ~5 kb that matches the size of the expected product from the VAN cassette in strain S613. While this band is absent in strain CK111, is present in each of the conjugated strains, suggesting presence of the VAN cassette in each of them.

Taken together, the data from Figures E.1 & E.2 show that strain S613 acts as conjugative donor when grown with strain CK111, donating a chromosomal copy of the VAN cassette which becomes incorporated into the genome of strain CK111, thereby
leading to vancomycin resistance. Although there is currently no evidence that strain S613 is able to conjugate with other cells of the same strain during the course of adaptation, this finding does raise the possibility as it demonstrates the capacity of strain S613 to function as a conjugative donor of chromosomal DNA.

Appendix F. Sequences of genetic insertion in conjugated *E. faecalis* strain TDR13.

Sequence of PCR product yielded when amplifying TDR13 carrying pCAM10 gDNA using primers Plia_F and LiaF_R. Sangar sequencing performed using primer Plia_F. The region that does not match with genomic S613 DNA sequence is highlighted in yellow.

Appendix G. Thermostability analysis of *B. subtilis* adenylate kinase mutants.

Although the focus of this thesis project relates to analysis of evolutionary trajectories to antibiotic adaptation in *Enterococcus faecalis*, here I discuss an independent project examining thermostable adaptation of *B. subtilis* adenylate kinase. Although tangential
antibiotic resistance, this project is also a study of evolution on a molecular level, published by *Biophysical Journal* in 2010 \(^{129}\).

This project began as a deeper examination into thermostable adenylate kinase mutants previous identified by Couñago *et al.* \(^{60}\). In this study, the gene encoding adenylate kinase (ADK) from mesophilic *Bacillus subtilis* was inserted into *Geobacillus stearothermophilus*, a thermophile, by homologous recombination, replacing the native adenylate kinase gene. This bacterial culture was then grown in a turbidostat, and the temperature was slowly raised. As adenylate kinase is both the only protein in the genome unadapted to high temperatures and is essential to cellular function (regulating ATP homeostasis), selective pressure was focused on ADK, requiring this protein to evolve thermostability for the cell to survive. After sequencing ADK within the *G. stearothermophilus* population on each day of the experiment, several mutations were observed.

The first mutation to arise in the population was Q199R, which rapidly reached fixation. Arising in the background of Q199R were the following five mutations competing within the turbidostat: G213E, T179I, G214R, A193V, and Q16L. Whether Q199R epistatically facilitated these secondary mutations, however, was initially unclear. To address this question, epistasis of Q199R with other mutations was investigated by using site-directed mutagenesis to study mutants G213E, T179I, G214R, A193V, and Q16L in isolation (see Appendix H). Following mutagenesis, each mutant protein was purified (see Appendix H) and thermostability was measuring using circular dichroism melting analysis (see Appendix H). Melting temperatures for each mutant are show
below in Table G.1, with wild-type ADK, mutant Q199R, and Q199R double mutants included as controls.

**Table G.1. Thermostability of ADK<sub>BSUB</sub> Mutants determined by CD with and without the initial Q199R Mutation.**

<table>
<thead>
<tr>
<th>Protein</th>
<th>T&lt;sub&gt;M&lt;/sub&gt; (°C)</th>
<th>Calculated T&lt;sub&gt;M&lt;/sub&gt; if mutations had been additive versus actual stability of the double mutants (Calc(T&lt;sub&gt;M&lt;/sub&gt;)-Actual (T&lt;sub&gt;M&lt;/sub&gt;)) °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt;</td>
<td>48.4 ± 0.2</td>
<td></td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R</td>
<td>49.5 ± 1.2</td>
<td></td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/G213E</td>
<td>51.4 ± 0.3</td>
<td>&lt;30</td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/G213E</td>
<td>&lt;30</td>
<td>&lt;31.1 vs 51.4 (-17.3)</td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/A193V</td>
<td>54.6 ± 0.2</td>
<td>50.0 vs 54.6 (-4.6)</td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; A193V</td>
<td>48.9 ± 0.7</td>
<td>48.9 vs 53.3 (-4.9)</td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/T179I</td>
<td>53.3 ± 0.5</td>
<td>48.9 vs 53.3 (-4.9)</td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; T179I</td>
<td>47.8 ± 0.2</td>
<td>50.9 vs 53.4 (-2.4)</td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/G214R</td>
<td>53.4 ± 0.6</td>
<td>50.9 vs 53.4 (-2.4)</td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; G214R</td>
<td>49.8 ± 2.3</td>
<td>56.5 vs 63.3 (-6.8)</td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/Q16L</td>
<td>63.3 ± 1.5</td>
<td>56.5 vs 63.3 (-6.8)</td>
</tr>
<tr>
<td>ADK&lt;sub&gt;BSUB&lt;/sub&gt; Q16L</td>
<td>55.4 ± 0.4</td>
<td>56.5 vs 63.3 (-6.8)</td>
</tr>
</tbody>
</table>

In the column on the right, Calculated T<sub>M</sub> was determined by adding the difference in stability between Wild-type ADK<sub>BSUB</sub> and ADK<sub>BSUB</sub> Q199R (1.1°C) to the melting temperature of each single mutant. (Figure from Miller et al., 2010<sup>129</sup>. Melting temperatures of double mutants determined by lab member Milya Davlieva)

Comparing the stabilities of wild type AK<sub>BSUB</sub> and Q199R, a T<sub>M</sub> increase of 1.1°C is observed in Q199R. Thus, if the thermostability benefits resulting from each secondary mutation were additive with respect to Q199R, we would anticipate a T<sub>M</sub> decrease of 1.1°C upon removal of Q199R. However, this is not the case for any of the double mutants studied, demonstrating that Q199R epistatically enables thermostability gains in each of the five double mutants.
In the far right column of Table G.1, epistasis has been quantified for each mutant studied by subtracting observed double mutant T\textsubscript{m} from the T\textsubscript{m} expected if Q199R bestowed an additive benefit. As mutation Q199R increases the melting point of wild type AK\textsubscript{BSUB} by 1.1ºC, additive T\textsubscript{m} values were calculated by adding 1.1ºC to the T\textsubscript{m} of each single mutant. Comparing (Calc T\textsubscript{m} - Actual T\textsubscript{m}) values given in Table 1, Q199R bestows a varying degree of epistatic benefit among the five mutants studied, however the most striking mutant in this regard is Q199R/G213E. While a stabilizing mutation in the presence of Q199R, in its absence G213E is greatly destabilizing, an effect not observed in any of the other mutants. Remarkably, attempts to determine the melting temperature of the G213E single mutant proved problematic, as the protein produced negligible elliptical signal throughout the melting experiment from 20-80ºC. Thus, to examine the secondary structure of AK\textsubscript{BSUB}G213E at the lowest temperature in the
melting experiment, a wavelength scan was performed at low temperature, shown below in Figure G.1.

*Figure G.1. Far-UV Wavelength scans of AK_{BSUB} Q199R and G213E at 20°C.* Figure from Miller et al., 2010.\(^{129}\) AK_{BSUB}Q199R exhibits minima at approximately 209 and 220 nm, indicating alpha helical secondary structure, while AK_{BSUB}G213E exhibits minimum at 204 nm, indicating random coil.

As shown in Figure G.1, AK_{BSUB}Q199R had a predominantly alpha helical structure, as expected, whereas AK_{BSUB}G213E exhibited a random coil secondary structure. This illustrates that in the absence of Q199R, presence of G213E mutation results in a largely unfolded protein structure, emphasizing the importance of epistasis during adaptive evolution.
To examine the mechanisms behind mutants Q199R/G213E and Q199R/G214R, a number of mutants affecting positions 213 and 214 were made through mutagenesis, purified, and their melting temperatures determined (see Appendix H), shown below, in *Table G.2.*

**Table G.2. Thermostability of AK<sub>BSUB</sub> Mutants at Gly-213 and Gly-214 as determined by CD.**

<table>
<thead>
<tr>
<th>Protein</th>
<th>T&lt;sub&gt;M&lt;/sub&gt; (ºC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK&lt;sub&gt;BSUB&lt;/sub&gt;</td>
<td>48.4 ± 0.2</td>
</tr>
<tr>
<td>AK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R</td>
<td>49.5 ± 1.2</td>
</tr>
<tr>
<td>AK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/G213E</td>
<td>51.4 ± 0.3</td>
</tr>
<tr>
<td>AK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/G214R</td>
<td>53.4 ± 0.6</td>
</tr>
<tr>
<td>AK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/G213A</td>
<td>37.9 ± 0.4</td>
</tr>
<tr>
<td>AK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/G213Q</td>
<td>39.3 ± 0.9</td>
</tr>
<tr>
<td>AK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/G214A</td>
<td>49.1 ± 0.1</td>
</tr>
<tr>
<td>AK&lt;sub&gt;BSUB&lt;/sub&gt; Q199R/G214K</td>
<td>48.4 ± 0.6</td>
</tr>
</tbody>
</table>

*Figure from Miller et al., 2010*<sup>129</sup>.

As shown in *Table G.2,* both mutations affecting position 213 decreased the thermostability of AK<sub>BSUB</sub> relative to Q199R/G213E. As mutant protein Q199R/G213A was designed to test the theory that the mechanism behind stabilizing mutation G213E in Q199R/G213E was increased helical rigidity, this hypothesis can be ruled out based on the decreased stability of this mutant. Similarly, mutant Q199R/G213Q, designed to test whether the glutamic acid in Q199R/G213E was forming a hydrogen bond or ion pair, was similarly destabilized, suggesting that an ion pair is being formed.

Looking at position 214, the two mutations affecting this position were also destabilized relative to Q199R/G214R. Analogous to Q199R/G213A, destabilization of
Q199R/G214A serves to rule out increased helical rigidity as the mechanism of mutation G214R. Mutant protein Q199R/G214K, designed to test if ion pairing was occurring as position 214, was also destabilized, thereby ruling out ion pairing as a likely cause for the mechanism of G214R and leaving hydrogen bond formation as the most likely explanation.

From a structural standpoint, the explanations put forward for Q199R/G213E and Q199R/G214R are consistent with the crystal structure, as shown below in Figure G.2.

Figure G.2. Crystallographic analysis of (a) AK

- SUB mutant Q199R/G213E and (b) Q199R/G214R. Bonding interactions depicted with dotted lines. Figure from Miller et al., 2010. Crystal structures solved by other lab members.

Figure G.2(a) above demonstrates that mutation G213E within Q199R/G213E has the capacity to fit into an electrostatic network, as it fits into a pattern of nearby alternately charged residues, consistent with the previous description of G213E functioning through ion pairing. In Figure G.2(b), the arginine in position 214 of Q199R/G214R is shown to be in close enough proximity with Y109 on an adjacent beta-
sheet to form a hydrogen bond, also consistent with data gleaned from mutagenesis experiments.

Appendix H. Methods applied for study of adenylate kinase protein

Site Directed Mutagenesis of Adenylate Kinase. Mutagenesis reactions were begun by using plasmid pET28 containing subcloned Bacillus subtilis adenylate kinase, using a modified Quikchange protocol. Each primers was designed to introduce a single nucleotide polymorphism (SNP) by placing the desired mutant nucleotide at the center of a primer pair of around 30 nucleotides in length. Note that while the mutant nucleotide will not pair with the template, the extended length of the primer is designed to increase the likelihood of a mismatch, thereby introducing the mutation. The reverse primer used in the reaction is the reverse complement of the forward primer. (Note: website PrimerX (http://www.bioinformatics.org/primerox/) has an automated mutagenesis primer design program).

For each reaction, the following reaction mixture was used: 5 µl Pfu Buffer (10X), 2 µl plasmid template (~100 ng/µl), 1 µl dNTPs (10 mM), 1 µl forward primer (25 µM), 1 µl reverse primer (25 µM), 40.5 µl H2O, and 0.5 µl Pfu Turbo polymerase. Cycling conditions are listed below, in Table 2.3.

Table H.1. Cycling conditions for Quikchange Mutagenesis of Adk

<table>
<thead>
<tr>
<th>Segment</th>
<th>Cycles</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>95°C</td>
<td>2 min</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>95°C</td>
<td>1 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>65°C</td>
<td>1 min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68°C</td>
<td>8 min</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>68°C</td>
<td>10 min</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4°C</td>
<td>forever</td>
</tr>
</tbody>
</table>
The conditions shown in Table 2.3 are optimized to amplify adenylate kinase. The 68°C extension time can be modified depending on the length of your DNA product.

PCR conditions can be optimized in multiple ways. Annealing temperature can be varied, between 55-65°C. An MgCl₂ gradient can also be used, adding 1-4 µl of 50 mM MgCl₂ in a reaction. A dNTP gradient can also be used to optimize the reaction. If none of the above optimizations yield success, a new primer pair can be designed with fewer or additional nucleotides.

For successful PCR reaction, products were run on an agarose gel. After producing a product of the expected plasmid size, products were PCR purified and digested with DpnI to remove template plasmid. Next 0.5 µL of 20 U/ml DpnI was added to the PCR product, centrifuging briefly at 10,000 RPM, then the sample was incubated at 37°C for one hour. When wild-type product was observed following DpnI digest, process was repeated and digestion time was increased to up to 2 hours.

Following DpnI digest, products were transformed into chemically competent E. coli XL1 blue cells. Cells were thawed on ice, and between 2.5 and 4.0 µl of DpnI digested PCR product was added to 50 µl of cells. When no colonies were observed, process was repeated and amount of DNA was increased up to 10 µl. Tubes were lightly flicked after adding DNA, and allowed to sit on ice for ~30 minutes. Cells were heat shocked at 42°C for 35 seconds, and then 0.9 ml Super-Optimum broth with Catabolite inhibition (S.O.C.) or LB was added. Tubes were shaken at 37°C for 1 hour, then 100-200 µl of cells were plated onto ampicillin plates and incubated at 37°C overnight.

After incubation, colonies were picked and grown in 5-10 ml of LB overnight. Plasmid was purified and sequenced to confirm mutation.
Expression and Purification of Adenylate Kinase. Expression and purification of mutant adenylate kinase protein was begun by transforming pET28 plasmid with subcloned Adk mutant of interest into *E. coli* BL21 DE3 cells using the transformation protocol described above.

Protein expression was then begun by inoculating *E. coli* BL21 DE3 into 50 of LB media containing and growing overnight at 37°C with shaking. On the next day, the 50 ml culture is added to 2L of LB containing 100 µg/ml ampicillin, and grown until OD$_{600}$ = 0.70. Next, 0.5 ml of 0.5 M IPTG is added, and cells are grown for another 4 hours. Cells are then pelleted using rotor JSP-F500, spinning at 8,000 rpm for 15 minutes at 4°C. At this point cells can be frozen and stored.

Purification was begun using a DE52 anion exchange gravity column. For this column, two buffers were made: low salt Buffer A, and high salt Buffer B. Buffer A contained 200 ml of 1 M Tris, 40 ml of 5 M NaCl, 4 ml of 1 M MgCl$_2$, 2.4 ml of 0.5 M EDTA, 0.043g DTT, and 3753.6 ml H$_2$O. To make Buffer B, 58.44g NaCl was added to a 1L aliquot of Buffer A.

Prior to running the DE52 column, the pelleted cells were resuspended in ~70 ml of Buffer A and poured into a beaker on ice. The cells were sonicated at 70% duty cycle, 7.5 output control on pulse mode for 2 minutes, followed by a two minute break, four times (16 minutes total time). Sonicated cells were then centrifuged at 15,000 RPM for 30 minutes at 4°C using rotor JA-20. The supernatant was then removed and run through a 0.22 µm syringe filter. Filtered supernatant was then diluted in ~500 ml Buffer A, and applied to an equilibrated DE52 column (see DE52 protocol for equilibration). The column was then washed with Buffer A until no protein was detected in the flow-through
(A$_{280}$ < 0.05). A gradient maker was then prepared with 100 ml Buffer A and 100 ml Buffer B, and was run through the column. Several 5 ml fractions were collected until all buffer was run through the column. A 15% sodium dodecylsulphate polyacrylamide gel (SDS-PAGE) was used to see which fractions contained Adk (MW = 24 kDa). Adk-containing fractions were pooled, and dialyzed overnight against Buffer A at 4°C.

The second column used in protein purification was a Q Sepharose anion exchange column within the Fast Protein Liquid Chromatography (FPLC) machine. The column was washed with 10 volumes of deionized (DI) water, and was then equilibrated with 5 volumes of Buffer A. Dialyzed protein was then injected onto the column, and a 45 ml gradient from 0% to 35% Buffer B was used to elute the protein. A fraction collected was used to collect 5 ml fractions. The column was then cleaned with 100% Buffer B. A 15% SDS-PAGE gel was again run to see which fractions contained Adk, which were then pooled and dialyzed against 2L of Buffer A at 4°C overnight.

The third column used during purification was a blue-sepharose column containing Cibacron Blue F3G dye (Bio-Rad Laboratories). This column was run as a gravity column at room temperature. Stored in 20% ethanol, the column was first washed with 300 ml water, and then equilibrated with 200 ml of Buffer A. Dialyzed protein was then applied to the column, and washed with 5 column volumes (200 ml) Buffer A. A higher-salt Buffer B2 solution was then made by adding 146.1 g of NaCl to 1L of Buffer A, and a 200 ml gradient of Buffer A and Buffer B2 was used to elute protein from the column, collecting 5 ml fractions. A 15% SDS-PAGE gel was again run to see which fractions contained Adk, which were then pooled and dialyzed against Buffer A overnight.
The fourth and final column used in purification of Adk was a GE Healthcare HiLoad 16/60 Superdex 200 pg Size Exclusion column, run inside the FPLC. Prior to preparing the column, the protein sample was concentration to a volume of 2 ml using a Sartorius Stedim 10,000 Molecular Weight Cut-off filter centrifuging at 3,400 RPM at 4ºC. The column was equilibrated with 1 volume of 25 mM Tris-HCl pH 7.4 with 150 mM NaCl. The sample was then applied to the column, and the column was run at 0.5 ml/minute. Fractions of 3 ml were collected, and were checked for protein content using 15% SDS-PAGE. Sample containing Adk were pooled, and dialyzed against 2L of 20 mM 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid (HEPES) buffer (pH 7.0) overnight at 4ºC.

Dialyzed protein was then concentrated to 30 mg/ml concentration (molar absorptivity = 11760 M⁻¹cm⁻¹). Protein was aliquoted into 200 µL tubes, flash-frozen in liquid nitrogen, and stored at -80ºC for future use.

**Circular Dichroism Melting Curves of Thermostable Adenylate Kinase Mutants.** To analyze secondary structures of proteins purified as described above, protein samples were thawed on ice and diluted to a concentration of 20 µM in 10 mM K₂HPO₄ buffer (pH 7.2). A 0.1 cm quartz cuvette was filled with protein, and placed into a circular dichroism (CD) spectropolarimeter. A spectral scan was performed from 300-200 nm to ensure protein folding. As Adk is predominantly alpha helical, the spectropolarimeter was then set to monitor 220 nm, and the cuvette was heated from 20ºC to 80ºC at a rate of 35ºC per hour. Temperature was then plotted against CD signal, and
the point of maximum first derivative was used to determine melting temperature for each protein.