Title

Engineering a High-Yield Platform for Protein Secretion in Bacteria

Author

Burdette, Lisa

Publication Date

2019

Peer reviewed|Thesis/dissertation

Engineering a High-Yield Platform for Protein Secretion in Bacteria

By

Lisa A Burdette

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jay D. Keasling, Chair
Professor Danielle Tullman-Ercek
Professor Wenjun Zhang
Professor John Dueber

Fall 2019

Abstract

Engineering a High-Yield Platform for Protein Secretion in Bacteria

By

Lisa A Burdette

Doctor of Philosophy in Chemical Engineering

University of California, Berkeley

Professor Jay D. Keasling, Chair

Recombinant proteins comprise billion-dollar industries, from therapeutics to enzymes used as catalysts in industrial processes. As the application space of recombinant proteins expands as a result of advancements in protein design and engineering, so does the demand for production systems that minimize the cost of developing and manufacturing the new protein products. Protein secretion in bacteria has the potential to combine the inexpensive cultivation, genetic tractability, and high productivity of traditional intracellular bacteria expression with the simplified downstream processing provided by eukaryotic systems that secrete the product into the extracellular space. The SPI-1 type III secretion system (T3SS) of *Salmonella enterica* Typhimurium is an ideal engineering target for heterologous protein secretion in bacteria—it is well-characterized, not essential for cell viability, and has demonstrated the capability to secrete heterologous proteins at titers up to hundreds of milligrams per liter.

This work expands upon initial engineering achievements in adapting the SPI-1 T3SS for heterologous protein production. **Chapters 2 and 3** describe design rules for growth media promote high secretion titers. A simple combination of glycolytic carbon source, buffering agent, high osmolarity, and essential nutrients was sufficient to increase secretion titer at least fourfold for multiple secreted proteins. **Chapter 4** describes the advantages and disadvantages of several methods to measure secretion titer. **Chapter 5** combines the knowledge acquired in Chapters 2-4 to create a library construction and screening workflow to identify hypersecreting variants of the T3SS needle protein, PrgI, via systematic, comprehensive mutagenesis. In addition to identifying many hypersecreting variants, systematic mutagenesis of PrgI revealed design rules for functional needle assembly and high secretion titers. Finally, **Chapter 6** outlines strain modifications that increase the utility of *S. enterica* Typhimurium as a recombinant protein production strain. Deletion of several pathogenic and non-essential genetic loci had a minimal effect on secretion titer, while deletion of the T3SS tip complex increased secretion titer. In combination with an optimized growth medium, genomic modifications that increased expression of the master regulator *hilD* produced higher secretion titers than observed in Chapter 2 and, importantly, simplified T3SS system expression. Collectively, the work described here advances the SPI-1 T3SS as a heterologous protein production platform and provides insight into the native function of the system.

Dedication

For my parents, Barb and Steve

Table of Contents

Abstract **1**
Dedication ..... **i**
Table of Contents ..... **ii**
List of Figures ..... **vi**
List of Tables ..... **ix**
List of Abbreviations ..... **x**
Acknowledgments ..... **xi**

Chapter 1: Introduction **1**

1.1 Industrial Protein Production in Gram-negative Bacteria 1
1.1.1 Characteristics that Differentiate Secretion Systems 2
1.1.2 Overview of Engineering Strategies 4

1.2 Engineering the T3SS for Heterologous Protein Production 7
1.2.1 System Structure and Overview 7
1.2.2 Flagellar T3SS 8
1.2.3 Injectisome 8

1.3 SPI-1 T3SS Biology: Considerations for Developing Engineering Strategies 10
1.3.1 SPI-1 T3SS Regulation 11
1.3.2 SPI-1 T3SS Structure 13
1.3.3 SPI-1 T3SS Protein Secretion 14

1.4 Other Virulence Factors of *Salmonella enterica* Typhimurium 15
1.5 Conclusions and Engineering Goals 16

Chapter 2: Medium Composition Affects Secretion Titer **18**

2.1 Introduction 18
2.2 Methods 19
2.2.1 Strains and Growth Conditions 19
2.2.2 Medium Formulations 20
2.2.3 Plasmid Construction 20
2.2.4 Recombineering 20
2.2.5 Protein Separation and Western Blotting 22
2.2.6 Protein Quantification 22
2.2.7 Flow Cytometry 22

2.3 Results 23
2.3.1 Growth medium affects secretion titer 23
2.3.2 Non-ionic carbon sources decrease secretion titer 25
2.3.3 Secretion titer increases with ionic content 27

2.3.4 Carbon sources and buffers have a synergistic effect on expression and secretion titer 29
2.3.5 Carbon sources and buffers affect T3SS transcriptional activity 31
2.3.6 The phosphate-sensing system PhoBR has no effect on expression or secretion titer with the engineered T3SS 34
2.3.7 An ideal medium formulation includes glycerol, phosphate, and sodium chloride 35
2.3.8 Increases in secretion titer from optimized growth media and strain improvements are additive and general for diverse secreted proteins 37
2.4 Discussion 39

Chapter 3: A Defined Medium for Increased Protein Secretion Titer 42

3.1 Introduction 42
3.2 Methods 42
3.2.1 Strains and Growth Conditions 42
3.2.2 DNA Manipulations 43
3.2.3 Protein Separation and Western Blotting 44
3.2.4 Protein Purification 44
3.2.5 Fragmentation Analysis and Zymograms 44
3.3 Results 45
3.3.1 An amino acid supplement in minimal media provides similar growth characteristics as LB-L 45
3.3.2 Supplemented defined media produces secretion titers equivalent to LB-L 47
3.3.3 Glycerol is an optimal carbon source in NCE 49
3.3.4 A high phosphate concentration is critical for increased secretion titer in defined media 50
3.3.5 Fragments appear in the secreted fraction after eight hours in the presence of glucose and glycerol 53
3.4 Discussion 57

Chapter 4: Methods for Measuring Protein Secretion Titer 59

4.1 Introduction 59
4.2 Methods 59
4.2.1 Strains and Growth Conditions 59
4.2.2 DNA Manipulations 60
4.2.3 Protein Separation and Western Blotting 62
4.2.4 Protein Quantification via Stained PAGE Gels 62
4.2.5 Protein Quantification via Western Blotting 62
4.2.6 Protein Purification 62
4.2.7 Split GFP Assay 63

4.2.8 Tetracysteine Tag Assay 63
4.2.9 Alkaline Phosphatase Activity Assay 63
4.3 Results 64
4.3.1 Western blots and directly stained SDS-PAGE gels are universal analysis techniques that require careful optimization 64
4.3.2 The tetracysteine tag is not ideal for T3SS secretion titer measurement 69
4.3.3 The split GFP assay has low sensitivity and is affected by medium composition 75
4.3.4 An alkaline phosphatase activity assay increases throughput of secretion titer analysis 81
4.4 Discussion 84
4.5 Acknowledgments 85

Chapter 5: Engineering the T3SS Needle Protein, PrgI, for Increased Secretion Titer 86

5.1 Introduction 86
5.2 Methods 88
5.2.1 Strains and Growth Conditions for Secretion Experiments 88
5.2.2 PCR and Cloning 90
5.2.3 Strain Construction for Single Genomic Modifications 93
5.2.4 Library Construction 93
5.2.5 Library Screening 94
5.2.6 Alkaline Phosphatase Activity 94
5.2.7 Sample Preparation for High Throughput Sequencing 94
5.2.8 High-Throughput Sequencing Data Processing 96
5.3 Results 96
5.3.1 Saturation mutagenesis at a single residue of PrgI reveals variable secretion titer 96
5.3.2 Optimizing an alkaline phosphatase activity assay for high-throughput screening of secretion titer 98
5.3.3 Generating a genomically-encoded CCM library 106
5.3.4 A “secretion fitness landscape” predicts secretion titer resulting from a single amino acid change in PrgI 109
5.3.5 The secretion fitness landscape predicted high-secreting variants 115
5.3.6 Double mutants confer increased secretion titer 117
5.4 Discussion 117
5.4.1 Secretion fitness of substitutions varies by helix 117
5.4.2 Outstanding questions 119
5.5 Acknowledgments 120

Chapter 6: Strain Optimization for Use of the SPI-1 T3SS as a Protein Production Platform 121

6.1 Introduction 121
6.2 Methods 122
6.2.1 Strains and Growth Conditions 122
6.2.2 Recombineering 123
6.2.3 Protein Separation and Western Blotting 124
6.2.4 Flow Cytometry 124
6.3 Results 125
6.3.1 SPI-1 effector knockouts do not alter secretion titer 125
6.3.2 A T3SS tip complex knockout increases secretion titer twofold 126
6.3.3 Non-essential pathogenic gene knockouts do not alter secretion titer 127
6.3.4 Methods to overexpress *hilD* confer increased secretion 128
6.3.5 Secretion titer is maximized by overexpressing *hilD* in media optimized for heterologous protein secretion 130
6.4 Discussion 133
6.5 Acknowledgments 133

Chapter 7: Conclusion 134

References 135
Appendix A: Chapter 2 Supplementary Figures 148
Appendix B: Defined Medium Recipes 153
Appendix C: Weighted-Average Secretion Score Calculations 155

List of Figures

Figure 1.1 One-step secretion systems. 3
Figure 1.2 Two-step secretion systems. 3
Figure 1.3 *Salmonella* pathogenicity island 1 regulation. 11
Figure 1.4 Structure of the T3SS. 13
Figure 2.1 Medium composition affects secretion titer. 24
Figure 2.2 Non-ionic carbon sources decrease secretion titer in LB-L. 26
Figure 2.3 Expression and secretion increase with ionic content. 28
Figure 2.4 Sugars and buffers have a synergistic effect on expression and secretion titer. 30
Figure 2.5 Expression and secretion per cell vary among combinations of buffers, salts and carbon sources. 31
Figure 2.6 Carbon sources and buffers affect T3SS transcriptional activity. 32
Figure 2.7 Carbon sources and buffers affect T3SS transcriptional activity. 33
Figure 2.8 Expression and secretion titer are not affected by PhoB. 34
Figure 2.9 An ideal medium formulation includes glycerol, phosphate, and sodium chloride. 36
Figure 2.10 Increases in secretion titer from optimized media and strain modifications are additive. 38
Figure 2.11 Titer of secreted proteins. 39
Figure 2.12 Undefined media components affects expression and secretion titer through mechanisms other than conductivity. 41
Figure 3.1 Growth curves in minimal media and LB-L with 0.4% glycerol or 0.4% succinate. 46
Figure 3.2 Growth curves for minimal media supplemented with 1X EZ Supplement and LB-L. 47
Figure 3.3 Relative expression and secretion titer in defined media. 48
Figure 3.4 Relative expression and secretion titer in NCE with various carbon sources. 49
Figure 3.5 Expression of SptP-Bla-GFP11-2xFLAG-6xHis in defined media with glucose, glycerol, or succinate. 51
Figure 3.6 Relative secretion titer of SptP-Bla-GFP11-2xFLAG-6xHis in defined media with glucose, glycerol, or succinate. 52
Figure 3.7 Analysis of cleavage product size after secretion for 24 hours. 54
Figure 3.8 Cleavage of SptP-tagged protein in the presence of the secreted fraction. 55
Figure 3.9 Cleavage of purified SptP-DH-2xFLAG-6xHis with and without protease inhibitors. 56
Figure 3.10 Zymogram of concentrated secreted fractions. 57
Figure 4.1 Coomassie quantification of SptP-DH-2xFLAG-6xHis secretion titer. 65
Figure 4.2 Relative secretion titer calculated from western blotting with different sample dilutions. 66
Figure 4.3 Representative western blots for Figure 4.2. 67

Figure 4.4 Secretion titers calculated from a quantitative Coomassie gel and a quantitative western blot. 68
Figure 4.5 Secretion titer relative to LB-L calculated from estimated titers. 69
Figure 4.6 Signal-to-noise ratios for various C-terminal TC tag arrangements. 71
Figure 4.7 Optimization of labeling time and amount of sample for labeling of various TC-tagged constructs with ReAsH. 72
Figure 4.8 Limit of detection and linear range for various TC-tagged constructs labeled with ReAsH or FlAsH. 73
Figure 4.9 Signal-to-noise ratio for secreted TC-tagged constructs with various excitation and emission filters. 74
Figure 4.10 FlAsH in-gel labeling and SYPRO Ruby post-stain of purified and secreted SptP-Bla. 75
Figure 4.11 Relative fluorescence for purified SptP-AP-GFP11-2xFLAG-6xHis. 76
Figure 4.12 Fluorescence of SptP-AP-GFP11-2xFLAG-6xHis secreted in various media. 77
Figure 4.13 Comparison of SptP-AP-GFP11 and SptP-DH-GFP11. 78
Figure 4.14 Relative secretion titer measured by fluorescence and densitometry for SptP-DH-GFP11-2xFLAG-6xHis. 80
Figure 4.15 Absorbance over time for varying amounts of purified SptP-AP and pNPP. 82
Figure 4.16 Initial rate of SptP-AP activity on pNPP. 83
Figure 4.17 Product formation over time for secreted SptP-AP-2xFLAG-6xHis. 84
Figure 5.1 Cartoon of the SPI-1 T3SS needle complex highlighting PrgI. 86
Figure 5.2 Relative secretion titer of SptP-DH-2xFLAG-6xHis from mutations at PrgI residue 41. 97
Figure 5.3 Absorbance over time and initial rate versus densitometry for secreted alkaline phosphatase expressed from a plasmid. 99
Figure 5.4 Product formation is slow with alkaline phosphatase secreted from a genomic copy 100
Figure 5.5 Product formation is slow with alkaline phosphatase secreted from a genomic copy 101
Figure 5.6 SptP-DH-2xFLAG-6xHis secretion in various media and culture formats. 102
Figure 5.7 Enzyme activity is not detectable in the secreted fraction of a 2 mL square 96-well deepwell plate culture with LB-IM. 103
Figure 5.8 Secretion with a genomic copy of SptP-AP is incompatible with 2 mL cultures in 24-well deepwell plates. 104
Figure 5.9 TB media and increased shaking speed produce measurable secreted enzyme activity from 2 mL square 96-well deepwell plate cultures. 105
Figure 5.10 One hour is a sufficient measurement range for secreted SptP-AP activity harvested from secretion in 0.5 mL TB, and 350 rpm in 2mL square 96-well deepwell plates 106
Figure 5.11 Overview of recombineering for genomic modification. 107

Figure 5.12 Distribution of position 41 mutants from a single recombineering event. . 108
Figure 5.13 Histogram of screened PrgI library members according to relative secretion titer. 109
Figure 5.14 Workflow for PrgI library assembly, screening, and analysis..... 110
Figure 5.15 Weighted-average secretion scores for all single amino acid variants of the modified residues of PrgI..... 112
Figure 5.16 Interior and exterior residues show different patterns for secretion fitness. 113
Figure 5.17 Substitution with large hydrophobic amino acids increased secretion titer at residues that line the hydrophobic groove of the needle interior. 114
Figure 5.18 Substitutions at residues not facing the interior of the needle likely affect secretion titer by changing inter- and intrachain interactions..... 115
Figure 5.19 Weighted-average secretion scores predict relative secretion titer. 116
Figure 5.20 Double mutants suggest epistatic effects. 117
Figure 5.21 Average secretion score by amino acid substitution. 118
Figure 5.22 Average secretion score of substituted residues differs in the N- and C-terminal helices. 118
Figure 6.1 Relative expression and secretion titer for individual SPI-1 effector knockouts. 125
Figure 6.2 Relative expression and secretion titer for *sip* knockouts..... 126
Figure 6.3 Relative expression and secretion titer for non-essential pathogenic gene knockouts. 127
Figure 6.4 Secretion titer and transcriptional activity of genomically modified strains compared to the state-of-the-art system..... 129
Figure 6.5 Secretion using the enhanced strains in an optimized medium..... 131
Figure 6.6 Histograms for flow cytometry of strains in Figures 6.4 and 6.5. 132

List of Tables

Table 1.1 Summary of Gram-negative bacterial secretion systems for heterologous protein production.... 4
Table 1.2 Features of Gram-negative bacterial secretion systems to consider when selecting a production platform. 5
Table 1.3 A selection of heterologous proteins secreted by Gram-negative bacteria.... 6
Table 1.4 Environmental conditions that control SPI-1.... 12
Table 1.5 Components of the SPI-1 T3SS apparatus. 14
Table 1.6 Native substrates and secretion tags for the SPI-1 T3SS. 15
Table 1.7 Selected virulence-enhancing genetic elements in *S. enterica* Typhimurium.16
Table 2.1 Strains used in Chapter 2.... 19
Table 2.2 Plasmids used in Chapter 2.... 19
Table 2.3 Primers used in Chapter 2.... 21
Table 2.4 Medium additives for Figure 2.3. 27
Table 2.5 Conductivity of LB supplemented with buffers, salts, and carbon sources. ... 29
Table 2.6 pH of the Secreted Fraction from Figure 2.4. 31
Table 3.1 Strains used in Chapter 3.... 43
Table 3.2 Plasmids used in Chapter 3.... 43
Table 3.3 Primers used in Chapter 3.... 43
Table 3.4 Growth rates, end OD, and pH in defined media and LB-L. 46
Table 3.5 Growth rates in defined media with 1X EZ Supplement compared to LB-L... 47
Table 3.6 Optical density at 8 hours in NCE with various carbon sources. 49
Table 3.7 Optical density at 8 and 24 hours in NCE, PCN, and PCN+P with glucose, glycerol, or succinate.... 53
Table 3.8 Strains used to generate the secreted fraction and analysis of cleavage of purified SptP-tagged protein.... 54
Table 3.9 Strains and fragmentation analysis. 56
Table 4.1 Strains used in Chapter 4.... 60
Table 4.2 Plasmids used in Chapter 4.... 60
Table 4.3 Primers used in Chapter 4.... 61
Table 4.4 Strains and purified proteins for FlAsH in-gel labeling.... 75
Table 5.1 Strains used in Chapter 5.... 88
Table 5.2 Plasmids used in Chapter 5.... 89
Table 5.3 Primers used in Chapter 5.... 91
Table 5.4 Pools for high throughput sequencing according to relative secretion titer.... 95
Table 5.5 PCR reactions for high throughput sequencing library preparation 95
Table 5.6 PCR cycling conditions for high throughput sequencing library preparation.. 96
Table 5.7 Recombination efficiency at the *prgI* locus. 107
Table 6.1 Strains used in Chapter 6.... 122
Table 6.2 Plasmids used in Chapter 6.... 123
Table 6.3 Primers used in Chapter 6.... 123

List of Abbreviations

AP – alkaline phosphatase
Bla – β-lactamase
BSA – bovine serum albumin
CHO – Chinese hamster ovary
CCM – comprehensive codon mutagenesis
DH – domain of human intersectin
EPEC – enteropathogenic *E. coli*
FlAsH – fluorescent arsenical hairpin binder
Fab – “fragment antigen binding” antibody fragment
fT3SS – flagellar T3SS
GFP – green fluorescent protein
GRAS – generally regarded as safe
HRP – horseradish peroxidase
IFN-α – interferon-α
IgG – immunoglobulin G
IL-2 – interleukin-2
IM – inner membrane
IPTG – isopropyl β-D-1-thiogalactopyrandoside
LEE – locus of enterocyte effacement
LMW – low molecular weight
MBP – maltose binding protein
MBS – MOPS-buffered saline
OM – outer membrane
OD – optical density
PBS – phosphate-buffered saline
PMF – proton motive force
POI – protein of interest
PI – protease inhibitor
pNPP – p-nitrophenyl phosphate
ReAsH – resorufin arsenical hairpin binder
RFU – relative fluorescence units
rhGH – recombinant human growth hormone
scFv – single chain variable fragment
sdAb – single domain antibody
Sec – general secretory pathway
SIEC – synthetic injector *E. coli*
SPI – Salmonella pathogenicity island
SRP – signal recognition particle
SF – secreted fraction
SFL – secretion fitness landscape
SyMAPS – systematic mutagenesis and particle selection
Tat – twin arginine translocation pathway
TC – tetracysteine
WCL – whole culture lysate

Acknowledgments

I owe thanks to many people for this document. First, to Danielle for being a great advisor to me over the course of five-plus years. Few people would have been a better match for my personality and working style. You let me explore and were always supportive of my choices, but you were critical when I needed you to be.

Much of the work in this thesis would not exist but for the previous students who laid the groundwork for it. Thank you to Eli and Kevin for getting the PrgI project started, and Kevin and Anum for being the founders of secretion subgroup. I must extend a special thank you to Kevin for being a great mentor when I started in the lab.

Much of the work in this thesis would also not exist without the undergraduate students I was lucky to work with over the course of my Ph.D. I am forever grateful to Stephanie, Leah, Emily, Jordan, Jinan, and Aliki for being the best students, people, and workers I could have asked for. You made me look great with all the data you generated, but more importantly, I loved working with you because you taught me as much as I taught you.

To my fellow DTE lab members, past and present—thank you for being great co-workers and friends. I'm grateful for all we experienced together: the critical feedback on actual work, the meandering conversations about a startlingly wide range of topics, and the entertainment we all provide each other. My time in graduate school would not have been the same without you. Nolan and Taylor, thanks for the free therapy.

I was fortunate enough to know the community of graduate students at two universities. Thanks to the Northwestern community for being so welcoming when and since we moved. To my friends and classmates at Berkeley—I truly would not have survived the first two years without you, and I've missed you since I left.

Thank you to everyone I worked with in the Shusta lab, at Genentech, and at Waisman Biomanufacturing—I wouldn't have developed into the scientist I am today without those experiences.

Kaitlin and Marie, I didn't expect to make lifelong friends with random roommates, but I am so thankful I met you. The year we lived together in Berkeley was hard on all of us, but you made our house a home and we had so many fun adventures together. I'm not sure I would have made it through my qualifying exam without you. I love you both!

I'm lucky to have two families. Doug, Kathy, Stu, Amelia, and Spencer—I love you all. You've been there for me since before this started, and I don't remember a time when I didn't feel your love and support. Thank you keeping me on my toes, sharing your dogs, and accepting me despite my deficiencies in dinosaur knowledge.

I wouldn't be here today without my parents and my brother. Mom and Dad, thank you for teaching me to be curious and encouraging me to learn about the world around me. Thank you for being my strongest cheerleaders and always reminding me that my potential is limitless. Eric, I couldn't have asked for a better person to grow up with. We are very different and yet so similar. I will never forget when you visited me in college and Amy and Sonja said, "oh my God, there's two of you!". I love you all so much.

Finally, Warren—this belongs to you too. It's been a long and difficult five years, but you put up with me and made sure I had a real life outside of lab. Thank you for being my partner and loving the mountains even more than me. I love you more than you know.

Chapter 1: Introduction

Portions of the following were adapted with permission from Burdette, L.A., Leach, S.A., Wong, H.T., Tullman-Ercek, D. Developing Gram-negative bacteria for the secretion of heterologous proteins. Microb Cell Fact ***17****, 196 (2018).*

1.1 Industrial Protein Production in Gram-negative Bacteria

The development of recombinant insulin and its production in *Escherichia coli* in the 1980s [1] launched an industry that comprises billion-dollar markets, with products ranging from protein biologics ($91B) [2] to industrial enzymes ($4.8B) [3,4]. In recent years, the chemical and structural diversity of protein products has expanded rapidly, with targets such as short anti-microbial peptides [5], antibody-like binding proteins for diagnostics [6], and protein biomaterials [7,8]. These vast differences in physicochemical properties call for a variety of strategies to ensure that production of heterologous proteins is robust and scalable.

Secreting protein to the extracellular space increases initial purity and thus can decrease the complexity of downstream processing. Eukaryotic hosts such as *Saccharomyces cerevisiae*, *Pichia pastoris* Chinese hamster ovary (CHO) cells, or HEK 293 cells secrete recombinant proteins through their native secretion machinery at high titers [9–12]. Mammalian cells, however, are slow growing, expensive to culture, have large batch to batch variations, and can be difficult to genetically engineer [10,13]. Yeasts typically are faster growing but can suffer from genetic instability and clonal variation [14].

Bacterial hosts offer fast growth, relative genetic simplicity, and genetic stability [15]. Gram-positive bacteria such as *Bacillus subtilis* and *Streptomyces lividans* allow protein secretion directly to the extracellular medium through the general secretory pathway, as they lack an outer membrane (OM). Secretion stress response systems and folding chaperones provide extracellular quality control. *B. subtilis* has a high secretion capacity—titers on the order of grams per liter have been achieved [16–18]. Compatibility with heterologous proteins is largely limited to enzymes, however, and extracellular proteolytic degradation is a significant product quality issue [19]. *S. lividans* has low extracellular protease activity while retaining a high secretion capacity. Nevertheless, the range of compatible proteins remains narrow, and growth of filamentous organisms such as *S. lividans* is difficult at scale. Progress in engineering these and other Gram-positive hosts for protein production is well-summarized in several review articles [18–22].

Gram-negative bacteria, including *E. coli*, are favored as bacterial hosts for protein production—they can express a broad range of heterologous proteins at high titers and are robust in industrial-scale culture [23]. A typical production process involves intracellular product expression, which requires cell lysis and multi-step purification to isolate the product from the cellular contents [1,2]. Though this process is used for many products at industrial scale—including insulin, interleukin-2 (IL-2), and human growth hormone (hGH)—intracellular expression introduces product and process development challenges [24]. For example, expression of IL-2 in *E. coli* required mutation of an internal cysteine to ensure proper disulfide bond formation after cell lysis [25]. Further, intracellular overexpression of recombinant proteins in bacteria often leads to the

formation of insoluble aggregates, or inclusion bodies. Initial product purity is higher in inclusion bodies compared to soluble cytosolic expression, but product recovery requires solubilization and refolding steps that must be optimized for each product [26]. Finally, some classes of proteins, particularly biomaterials, are toxic when expressed intracellularly in Gram-negative hosts, which yields low titers [27,28].

Gram-negative bacteria secrete few native proteins [29]. Thus, heterologous protein secretion from Gram-negative bacteria combines the advantages of bacterial production with broad substrate compatibility and the high baseline purity afforded by extracellular secretion. Gram-negative bacteria possess seven secretions systems—types I-VI and type VIII—that are known to secrete protein to the extracellular space [30]. Recombinant proteins are secreted successfully by five of these systems: type I, type II, type III, type V, and type VIII secretion systems (T1SS, T2SS, T3SS, T5SS, and T8SS, respectively) [24,31]. Moreover, some proteins are natively secreted by *E. coli* via other, as-yet unknown mechanisms and are used as fusion partners to facilitate secretion of recombinant proteins [32–34]. As another alternative, the outer membrane can be engineered to release periplasmic protein [35,36]. The following sections describe characteristics of Gram-negative bacterial secretion systems, outline general engineering strategies for improving protein secretion via those systems, and highlight the T3SS as a promising platform for heterologous protein production.

1.1.1 Characteristics that Differentiate Secretion Systems

Bacterial secretion systems can be divided into two classes, based on whether their cargo enters or bypasses the periplasm. One-step systems secrete proteins directly to the extracellular space from the cytosol, while two-step systems first export proteins to the periplasm through the general secretory (Sec) or twin arginine translocation (Tat) pathways. In this latter system, the target subsequently traverses the outer membrane to the extracellular medium. Table 1.1 provides a summary of each secretion system.

The type I and type III secretion systems are both one-step systems. In these pathways, protein cargo must be either completely or partially unfolded to secrete successfully. Still, several groups have shown that proteins exported by these systems can fold properly in the extracellular media, even when disulfide bridges or salt bridges are required [37–40]. In addition, one-step systems are not required for cell viability, which expands the engineering space available for these systems [41,42].

Two-step systems, including the type II and type V secretion systems, provide access to useful machinery found in the periplasm, including foldases, chaperones, and other protein-folding enhancers [43,44]. However, Sec is always essential for cellular viability, and Tat can be essential depending on the strain and growth conditions. Thus, using the Sec and Tat pathways can limit the engineering space of the system [31,45].

Before selecting a production platform, it is critical to understand the features of bacterial secretion systems that affect compatibility with a desired protein product (Table 1.2, Figure 1.1, Figure 1.2). While rigorous criteria have not been explicitly developed for any of the secretion systems, empirical evidence suggests that folding kinetics, folding complexity, and protein size can play a role in secretion efficiency [8,37,46]. Table 1.3 lists a selection of proteins that are successfully secreted by Gram-negative bacteria.

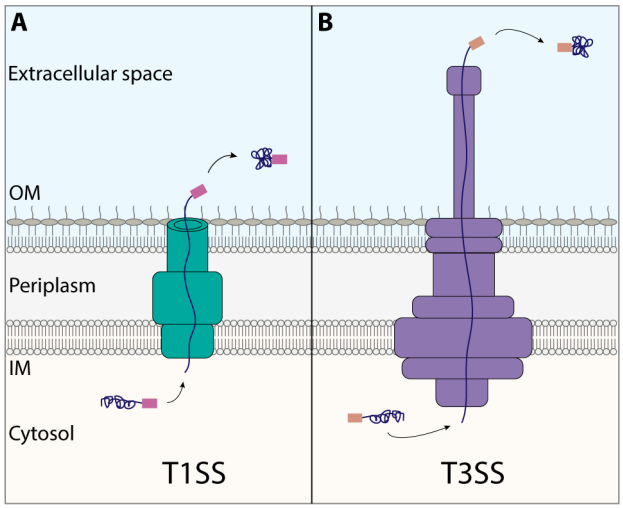


Figure 1.1 One-step secretion systems.
Proteins (dark blue) are translocated directly from the cytosol to the extracellular space in an unfolded state, bypassing the inner and outer membranes (IM and OM, respectively). A C-terminal (T1SS, **A**) or an N-terminal (T3SS, **B**) secretion tag is required for translocation and remains attached to the cargo upon exiting the secretion apparatus.

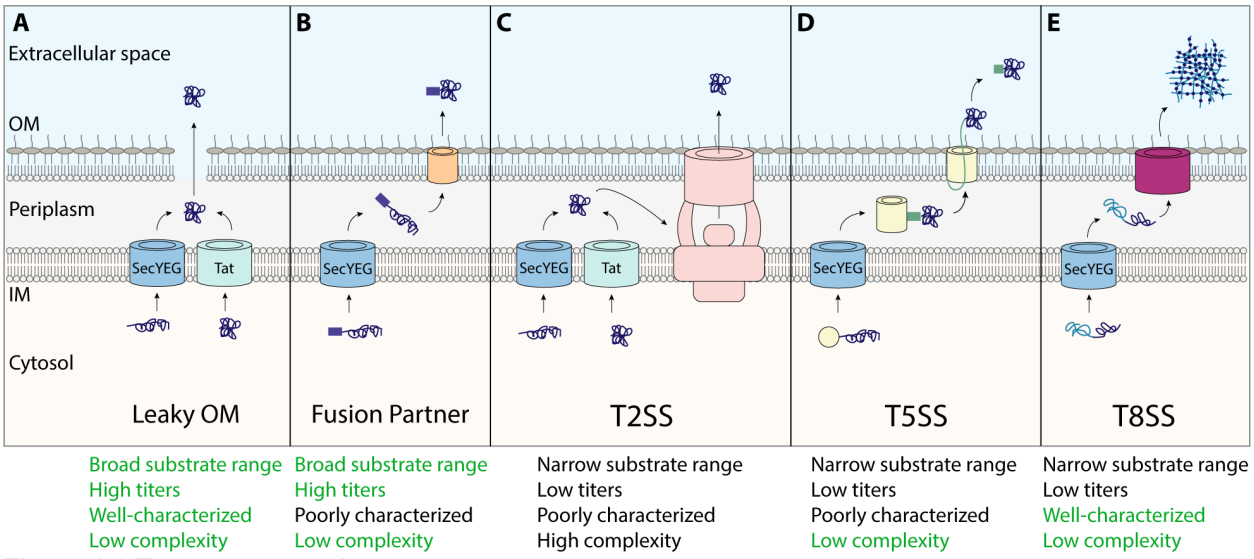


Figure 1.2 Two-step secretion systems.
A Proteins (dark blue) are exported through the inner membrane (IM) via either Sec or Tat before passively diffusing into the extracellular space. **B** An example of transport via a fusion partner. Export pathway specificity is unknown for many fusion partners, but YebF (purple box) is secreted only when it is exported through Sec. It is believed to translocate the outer membrane (OM) via a porin (orange). **C** Proteins are exported through either Sec or Tat before entering the pseudopilus apparatus (pink) that transports cargo across the OM. **D** The translocation domain-passenger domain fusion is exported through the Sec pathway. The translocation domain (yellow) inserts in the outer membrane and the passenger domain (green) is secreted through the pore. An autocleavage event releases the passenger domain in the class of T5SS discussed here. **E** Proteins fused to the curli subunit (teal) are exported through Sec and are thought to traverse the outer membrane via an entropy gradient in a chaperonin-like structure (magenta). In the

absence of the protein that anchors curli subunits to the outer membrane, fibers spontaneously polymerize and aggregate into networks in the extracellular space.

Table 1.1 Summary of Gram-negative bacterial secretion systems for heterologous protein production.

Secretion System	Model Organisms	Genetic Locus	Model Secretion Tag	Master Regulator
Type I	*Pseudomonas fluorescens*	*tliDEF*	TliA	
Type I	*E. coli*	*hlyA, hlyB, hlyD*	HlyA	
Type II	*E. coli*	*gsp*	Unknown	
Generalized Two-Step	*E. coli*	N/A	YebF	
Type III	*Salmonella enterica*	SPI-1	SptP	HilA
Type III	*Salmonella enterica*	*fli, flhDC*	FlgM (flagellar)	FlhD4C2 (flagellar)
Type III	Enteropathogenic *E. coli* (EPEC)	locus of enterocyte effacement (LEE)	EspA	Ler
Type III	*Shigella flexneri*	*Shigella* virulence plasmid	OspB	VirB, VirF
Type III	*Yersinia enterocolitica*	pYV virulence plasmid	YopE	VirF
Type V	*E. coli*	*tps*	Pet	
Type VIII	*E. coli*	*csg*	CsgA	CsgD (not necessary for synthetic induction)

1.1.2 Overview of Engineering Strategies

In addition to dividing bacterial secretion systems into two classes, the engineering strategies for increasing protein secretion can be divided into four categories: 1) modifying the secretion tag that targets a protein for secretion; 2) engineering the secretion system machinery; 3) transferring a secretion system to a more desirable production strain; and 4) manipulating the genetic regulation of the secretion system.

All secretion systems identified here require a signal sequence, a peptide secretion tag that is appended to the N- or C-terminus of the protein cargo, to facilitate secretion by the desired system [47]. Much engineering effort on the Sec and Tat pathways has uncovered clear design rules and minimal signal sequences that facilitate secretion through these systems. These rules form the basis of prediction algorithms that determine whether a given sequence is likely to be exported via one of these systems [47–49]. Similar approaches have been applied to other bacterial secretion systems, and the results suggest that secretion efficiency is highly dependent on the heterologous protein, though concrete design rules remain elusive [50]. Continuing this work to define design rules for secretion signal sequences will expedite process optimization for new protein products.

Altering the secretion apparatus can lead to higher secretion titer by improving the properties or increasing the capacity of the secretion system, though this approach is more feasible for secretion systems that are not essential to growth or survival [51,52]. Traditional protein engineering tools, such as random mutagenesis, have been used in the directed evolution of the secretion apparatus with some success [50,53]. However, limited sequence space can be surveyed using this approach. For future efforts, the development of high throughput screens and selections is critical to increase the utility of this strategy [54–56].

The ability to synthetically overexpress the secretion machinery is an important tool to maximize secretion titer [57]. This is especially the case for secretion systems that are not essential for growth and thus not constitutively expressed or highly expressed, such as the Tat pathway and the type III secretion system [41,58–60]. Successes in this area include overexpressing master regulators that control the expression of the system [41,58,59] and synthetic expression of individual system components [60,61]. This strategy increases the pool of secretion system machinery available for heterologous protein secretion.

Finally, many bacterial secretion systems are natively part of pathogenesis, and yet industrial-scale production requires safe and, ideally, non-pathogenic strains. Two approaches are being studied for using these systems in a non-pathogenic context. One strategy is to move the secretion machinery into a strain of *E. coli* that is already optimized for industrial protein production [58,61–64]. Another viable strategy is to derive a non-pathogenic production strain from the pathogenic parent strain through genomic engineering [65,66].

Table 1.2 Features of Gram-negative bacterial secretion systems to consider when selecting a production platform.

Secretion System	Number of Steps	Secretion Tag Cleavage	Pathogenic Origin	Access to Folding Chaperones	Secretion System Induction Scheme
Type I	One	No	Yes	No	Synthetic
Type II	Two	Unknown	Yes	Yes	Constitutive
Leaky Membrane	Two	No tag needed	No	Yes	Constitutive
Unknown Mechanism	Two	No	No	Yes	Constitutive
Type III	One	No	Yes	No	Synthetic
Type V	Two	Partial	Yes	Yes, but unfolded for transport across OM	Synthetic
Type VIII	Two	No	Yes	Yes	Synthetic

Table 1.3 A selection of heterologous proteins secreted by Gram-negative bacteria.

Secretion System	Mechanism	Secretion Product	Product Type	Extracellular Titer	Reference
Type I	HlyA	scFv	antibody	2 mg/L	[67]
Type I	HlyA	CGTase	industrial	0.58 mg/L	[68]
Type I	HlyA	Cutinase	industrial	334 U/mL	[69]
Type I	HlyA	Interferon alpha 2	therapeutic	6 mg/L	[37]
Type I	TliDEF, HlyA	Lipase	industrial	8450 U/mL, 3 mg/L	[37],[61]
Type I	TliDEF	Metalloprotease	industrial	789 mg/L	[62]
Type I	TliDEF	Endo-β-1,4-mannanase	industrial	4.65 mg/L	[70]
Type II and generalized two-step	TatExpress	hGH	therapeutic	30 mg/L	[60]
Type II and generalized two-step	TatExpress	scFv	antibody	N/A*	[60]
Type II and generalized two-step	Periplasmic expression via SRP	IgG	antibody	236.5 mg/L	[71]
Type II and generalized two-step	Leaky OM	Human parathyroid hormone	therapeutic	680 mg/L	[36]
Type II and generalized two-step	Leaky OM	Fab	antibody	6 g/L	[72]
Type II and generalized two-step	Leaky OM	IFN-α	therapeutic	N/A*	[73]
Type II and generalized two-step	Unknown, Cel-CD fusion	CGTase	industrial	637.4 U/ml, 348 mg/L	[34],[74]
Type II and generalized two-step	Cel-CD fusion	Carbohydrate binding domain	industrial	348 mg/L	[34]
Type II and generalized two-step	Cel-CD fusion	Neuritin	therapeutic	211 mg/L	[34]
Type II and generalized two-step	YebF	α-amylase	industrial	150 μmol glucose/min/mg protein	[54]
Type II and generalized two-step	YebF	Interleukin-2	therapeutic	43,800 U/mL	[75]
Type II and generalized two-step	YebF	NanH2 Sialidase	therapeutic	N/A*	[76]
Type II and generalized two-step	OsmY	Human leptin	therapeutic	250 mg/L	[77]
Type II and generalized two-step	OsmY	Osteopontin	therapeutic	3.6 mg/L	[78]
Type III	flagellar	Interferon alpha 2	therapeutic	0.6 mg/L	[50]
Type III	flagellar	Lipase	industrial	~420 U/L	[79]
Type III	flagellar	Enfuvirtide	therapeutic	13.4 mg/L	[80]
Type III	flagellar	apical membrane antigen 1	therapeutic	N/A*	[81]
Type III	flagellar	δ-SVIE	therapeutic	N/A*	[81]
Type III	flagellar	MrVIA	therapeutic	N/A*	[81]
Type III	flagellar	NCR peptide	therapeutic	N/A*	[81]
Type III	injectisome	Spider silk monomer	material	14 mg/L	[27]

Secretion System	Mechanism	Secretion Product	Product Type	Extracellular Titer	Reference
Type III	injectisome	Elastin	material	20 mg/L	[27]
Type III	injectisome	Resilin	material	20 mg/L	[27]
Type V	Type V	Pertactin	therapeutic	1 mg/L	[39]
Type V	Type V	Ag85B	therapeutic	N/A*	[39]
Type VIII	Type VIII	sdAb	antibody	N/A*	[82]
Type VIII	Type VIII	Cecropin A	therapeutic	294 mg/L (after purification)	[83]
Type VIII	Type VIII	Mussel foot protein	material	N/A*	[84]

*The study did not report secretion titer

1.2 Engineering the T3SS for Heterologous Protein Production

One of the best characterized systems for heterologous protein secretion in bacteria is the type III secretion system. A wide variety of proteins are successfully secreted by the T3SS (Table 1.3), and the T3SS is not required for cell viability, which makes it an attractive engineering target. The following sections provide a broad overview of the T3SS and summarize efforts to engineer it for heterologous protein secretion.

1.2.1 System Structure and Overview

There are two classes of type III secretion systems: the injectisome and the flagellar T3SS (fT3SS). Pathogens such as *Salmonella enterica*, *Yersinia enterocolitica*, and *Shigella flexneri* use the needle-like injectisome to secrete proteins that facilitate virulence into host cells. The fT3SS natively secretes and assembles the bacterial flagellum. Both T3SSs are multimeric protein structures that span the inner and outer membranes (Figure 1B) [85]. Proper T3SS function requires additional regulatory proteins and sRNAs, so it is difficult to transfer the T3SS to a plasmid, much less a non-native host. An N-terminal secretion signal targets proteins to the secretion apparatus, where they are secreted in an unfolded state directly from the cytosol to the extracellular space. Some secretion signals require a cognate chaperone to carry out this function. The secretion signal remains attached to the secreted protein [86].

Researchers have successfully secreted high titers of a wide variety of proteins using these systems. Proteins such as scFvs, alkaline phosphatase, and spider silk monomers are secreted at titers up to 130 mg/L via the injectisome [8,24,38]. The flagellar T3SS secretes a variety of substrates as well, including GFP, neuroactive peptides, and α-enolase at titers up to 15 mg/L [87,88]. The T3SS is particularly attractive for its ability to express and secrete otherwise difficult-to-express proteins [27].

Though researchers achieved high titers using the T3SS, design rules for compatible substrates and secretion tags remain elusive. Regulation of the T3SS is tightly controlled and sensitive to environmental input and growth phase [86]. Moreover, the T3SS injectisome in particular is found in pathogenic bacteria, which are undesirable hosts from an industrial perspective. Recent research has focused on characterizing the secretion signal, transferring the T3SS to non-pathogenic organisms, and overriding native control of the system to increase titer and decrease sensitivity to environmental input [38,50,58].

1.2.2 Flagellar T3SS

1.2.2.1 Signal Sequence Engineering

The library of available flagellar secretion tags are empirically-determined truncations of native substrates [41,59]. However, in two recent studies, researchers sought to characterize the secretion tags through a more systematic approach [50,89]. In general, these studies revealed that secretion recognition and regulation via the fT3SS is complex, and that design rules might be challenging to define. For example, truncations of FlgB and FlgE fused to β-lactamase (Bla) required a native *flgB* 5'UTR for efficient secretion, while secretion of a FlgK-Bla fusion remained unchanged in the presence of the native *flgB* 5'UTR [89]. A long truncation of FliC fused to interferon-α (IFN-α) secreted better than a short truncation, while the inverse was true for MBP, and there was no difference in secretion titer for GFP [50]. Secretion efficiency depended strongly on the combination of the secretion tag and the heterologous protein, and no correlations have yet been identified.

1.2.2.2 Manipulating Native Regulation

The flagellar T3SS integrates numerous inputs in its complex regulatory circuit. Identifying critical nodes of control can increase secretion efficiency, as described above for the T3SS injectisome. A systematic evaluation of factors that could affect secretion titer via the *S. enterica* fT3SS determined that master regulator overexpression, increased chaperone stability, and high ionic strength (e.g. 200 mM NaCl) all increased secretion titer. Both knocking out native substrates and deleting SPI-1, which is known to participate in regulatory crosstalk with the fT3SS, had no effect or caused a slight decrease in secretion titer. However, combining the factors that improved secretion into an optimized strain facilitated secretion of difficult-to-express proteins such as conotoxins δ-SVIE and MrVIA, nodule-specific, cysteine-rich antimicrobial peptides (NCR), and a malaria surface antigen domain of apical membrane antigen AMA-1 [81].

1.2.2.3 Transferring the Secretion Machinery

The flagellar system is largely conserved throughout the domain of Eubacteria and thus improvements in one flagellar system can be applied to another. Many strains of *E. coli* currently used for biotechnological purposes lost their flagellar machinery during the domestication of those strains [90]. Auer and colleagues sought to restore flagellar machinery by inserting the fT3SS operon into the genome of the biotechnologically-relevant HMS174(DE3) strain and placing it under the control of the P_{lacUV5} promoter [59]. This construct was insufficient to generate observable FlgM secretion, but adding a second copy of the operon to increase gene products of the flagellar machinery and swapping the P_{lacUV5} promoter with the stronger T7 promoter produced measurable secretion titer. Absolute titer was not reported, making direct comparisons to the native fT3SS in *S. enterica* challenging.

1.2.3 Injectisome

1.2.3.1 Strain Engineering

Native T3SS injectisome machinery integrates a variety of input signals to ensure that T3SS machinery is constructed only when required for its pathogenic function. Although this strategy confers obvious evolutionary benefit, it often works counter to the goal of maximizing secretion titer. Thus, researchers have sought to manipulate and

engineer regulatory pathways to exploit positive regulation and remove or circumvent negative regulation.

T3SS activation and assembly is controlled by at least one master regulator. Overexpressing a master regulator allows synthetic induction and can increase secretion activity in the population. In *S. enterica*, synthetic overexpression of the SPI-1 T3SS master regulator HilA increased the proportion of cells expressing the T3SS from around 30% to near 100% and increased secretion titer of a model protein tenfold [41]. This modification also promoted secretion of other heterologous proteins that were previously considered incompatible with the T3SS. Crucially, this increase in secretion titer and activity occurred in conditions that are known to repress the T3SS, such as high oxygenation, which is desirable to increase cell density and therefore bulk secretion titer.

Subsequent studies combined this work with strain engineering to secrete a variety of biopolymers at titers up to 130 mg/L [27,52]. This titer was achieved by knocking out a protein present in the needle tip complex, SipD, and adding purified SipD exogenously. The *ΔsipD* strain increased titer 50-fold relative to wild type, and exogenous addition of purified SipD increased titer an additional twofold. The mechanism of the increase in secretion as a result of exogenous SipD addition has not yet been determined conclusively.

To remove all native regulatory inputs to the *S. enterica* T3SS, Song *et al.* recently reported a remarkable bottom-up approach to refactor the entire 35 kb SPI-1 locus, essentially rebuilding it from scratch [65]. The researchers removed non-essential genes, codon optimized certain genes, scrambled the gene order, and replaced non-coding genes with synthetic, controllable, well-studied genetic parts. The result was a minimal 16 kb SPI-1 that was about half the nucleotide length of wild type. The refactored system was completely controlled by synthetic regulation, and it was active in T3SS-repressing conditions. In addition to fully decoupling the T3SS from native regulation, this bottom-up approach revealed new information about necessary regulatory elements in the native system. For example, the researchers were able to identify at least two new essential factors—the small RNA, InvR, and an internal start site that generated a shorter version of the structural protein SpaO.

1.2.3.2 Engineering Non-Pathogenic Secretion Strains

The most popular approach for removing the T3SS from a pathogenic context is transferring the T3SS genes to a non-pathogenic host. This is a non-trivial endeavor because of its size and complexity, so a number of avenues have been explored. Much of this work is motivated by vaccine delivery efforts, as well-formed T3SSs can inject proteins directly into the cytosol of mammalian cells. The T3SS injectisomes of *Vibrio parahaemolyticus*, *Yersinia pestis*, and *S. enterica* can be expressed from plasmids in non-pathogenic *E. coli*; however, genomic integration of a synthetically controlled T3SS would create a more robust, programmable production strain [63,91,92]. One such strategy involved constructing a series of five rationally designed transcriptional units (TUs) that encoded the structural genes from the enteropathogenic *E. coli* (EPEC) locus of enterocyte effacement (LEE). Each TU was placed under the control of a synthetic Ptac promoter and integrated into a specific location in the *E. coli* K-12 genome to create a synthetic injector *E. coli* (SIEC) strain. The SIEC strain was capable of secreting native substrates expressed from a plasmid into HeLa cells [64].

Reeves *et al.* pursued yet another approach to transfer the *S. flexneri* T3SS into *E. coli* DH10B [58]. The researchers cloned the 31 kb of DNA necessary to express a functional T3SS from the *Shigella* virulence plasmid to a 44 kb synthetic plasmid that also contained the genetic elements necessary to transfer the T3SS to a defined "landing pad" on the DH10B chromosome. The resulting strain, mT3 *E. coli*, was secretion-active upon synthetic induction of a master regulator present on a separate plasmid. The strain was capable of secreting equal levels of endogenous proteins as native *S. flexneri* and successfully translocated proteins into HeLa cells. The researchers used this programmable, non-pathogenic strain to identify signal sequences that promoted secretion of therapeutically relevant proteins such as mammalian reprogramming factors and TALENs.

Creating a replication-deficient host can also abstract the T3SS from a pathogenic context, as demonstrated by Carleton *et al.* in *ΔminD S. enterica* [93]. The *ΔminD* strain divided aberrantly to produce minicells that lacked chromosomes and therefore could not replicate further. The minicells were metabolically active and could synthesize proteins from a plasmid introduced in the parent strain. The researchers created T3SS-active minicells by overexpressing the master regulator HilA and expressing the secretion target from a plasmid.

Ittig *et al*. engineered the host strain directly to reduce pathogenicity [66]. In this study, all native substrates of the *Yersinia enterocolitica* T3SS were knocked out, and a synthetic growth dependency on meso-2,6-diaminopimelic acid was engineered, making the strain consistent with biosafety level 1 standards. The engineered strain was capable of secreting a variety of heterologous proteins expressed from plasmids into HeLa cells.

1.2.3.3 High Throughput Screening Methods

Some T3SS injectisomes, such as those of *S. flexneri* and *Y. enterocolitica*, are activated by the addition or depletion of small molecules [94,95]. Lesser and colleagues recently showed that this property could be exploited to establish high-throughput screening methods for rapid engineering of the T3SS [56]. The *S. flexneri* T3SS is activated in the presence of the dye Congo red, so the researchers developed a solid plate-based assay in which liquid cultures were spotted onto solid media that contained Congo red and then transferred onto another solid plate that contained IPTG to induce the expression of the desired secretion product. A nitrocellulose membrane was attached to the plate that contained IPTG, incubated overnight, and probed with antibodies to detect secretion. This method allowed them to rapidly screen secretion titers of *S. flexneri* T3SS substrates in a multitude of conditions. The mT3 *E. coli* strain described previously was also compatible with this system. Though only a subset of T3SSs can be activated with a small molecule, the design of this screen could be adapted to other systems to enable rapid engineering.

1.3 SPI-1 T3SS Biology: Considerations for Developing Engineering Strategies

As the preceding sections outline, the SPI-1 T3SS of *S. enterica* Typhimurium has been a frequent and successful target of engineering efforts to adapt it for heterologous protein secretion [8,27,38,41,52]. Though the refactoring effort of Song *et al.* stripped it

of its native biological regulation [65], most engineering efforts focused on co-opting its native regulation to increase heterologous protein secretion. The following sections provide a more detailed overview of the native regulation and structure to highlight considerations for engineering the SPI-1 T3SS as a heterologous protein secretion platform.

1.3.1 SPI-1 T3SS Regulation

The 35 kb SPI-1 genetic locus contains most of the genes that code for regulatory, structural, and secreted proteins of the SPI-1 T3SS [96]. SPI-1 genes are organized in the *inv*, *prg*, and *sic* operons (Figure 1.3). The master regulator HilD is proposed to participate in a feed-forward loop with HilC and RtsA to control expression of the master regulator HilA and induce a signal cascade that assembles and activates the T3SS [97]. HilD appears to be the dominant regulator—its absence decreased *hilA* expression tenfold, while lack of HilC or RtsA decreased *hilA* expression only twofold.

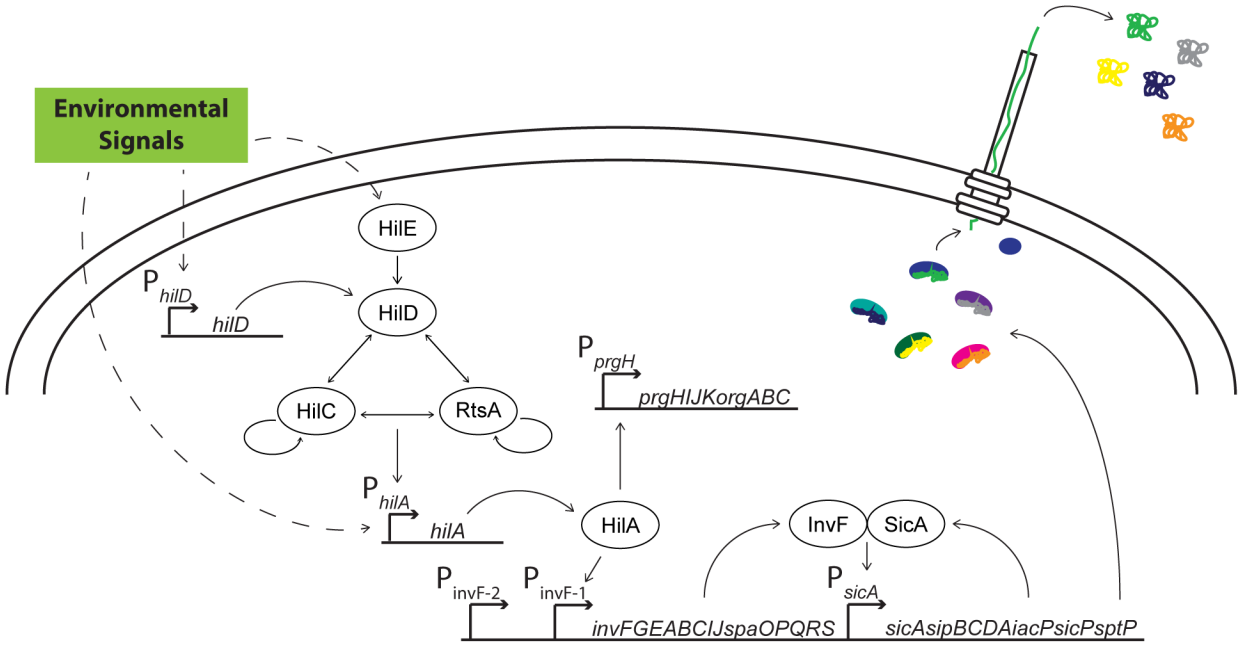


Figure 1.3 *Salmonella* pathogenicity island 1 regulation.
The *Salmonella* pathogenicity island 1 contains regulatory and structural genes for the T3SS. The master regulator HilA integrates environmental cues to trigger a signal activation cascade that builds and activates the T3SS. Environmental signals travel through two component systems or more complex signal cascades to control HilD expression at the transcriptional, translational, and post-translational levels. Some signaling pathways control *hilA* transcription directly [98]. The figure was adapted from [96] and [98].

Expression of the genes within SPI-1 is controlled by a confluence of environmental signals to ensure that cells express the ~4 MDa apparatus only when they exist in the correct environment to use it [96,98–100]. Natively, *S. enterica* uses the SPI-1 T3SS to prepare mammalian gut epithelial cells for invasion, an early step in its pathogenic program. Thus, environmental cues that signal proximity to the gut epithelium induce SPI-1 T3SS expression, while environmental cues that signal that the cells have advanced to the next stage of the infection program repress SPI-1 genes. A sample of environmental conditions that control SPI-1 gene expression are listed in Table 1.4.

Table 1.4 Environmental conditions that control SPI-1.

Environmental Factor	System that Transmits Signal to SPI-1	Point of Control in SPI-1	Reference
Divalent cations	PhoPQ	*hilA* transcription	[99,101]
Phosphate	PhoBR BarA/SirA	*hilA* transcription *hilD* translation	[99,101] [102]
Iron	Fur	*hilD* autoinduction	[98]
Glucose	Csr via BarA/SirA	*hilD* translation	[102,103]
Short chain fatty acids	BarA/SirA	*hilD* translation	[104]
Oxygen	sRNAs FnrS, ArcZ	*hilD* translation	[105]
Osmolarity	BarA/SirA	hilA, hilD transcription	[106]
Temperature	H-NS, Hha	*hilA, hilD, hilC, rtsA* transcription	[107,108]
pH	Unknown	*hilA* transcription	[100]

1.3.2 SPI-1 T3SS Structure

The ~4 MDa T3SS injectisome is composed of 1-100 copies each of >20 different proteins that are precisely assembled to form a basal body, needle, and tip complex that spans the inner and outer membranes of the bacterial cell [109] (Figure 1.4). The components of the basal body are exported through the Sec pathway. After the basal body assembles correctly, the protein chains that compose the needle are secreted through the nascent T3SS structure and spontaneously polymerize to form the lumen through which substrates are secreted [110]. A tip complex is the last structural component to be added. The order of substrate secretion and the needle length are tightly controlled, though the mechanisms of control are disputed [111–116]. The proteins that compose the SPI-1 T3SS are listed in Table 1.5.

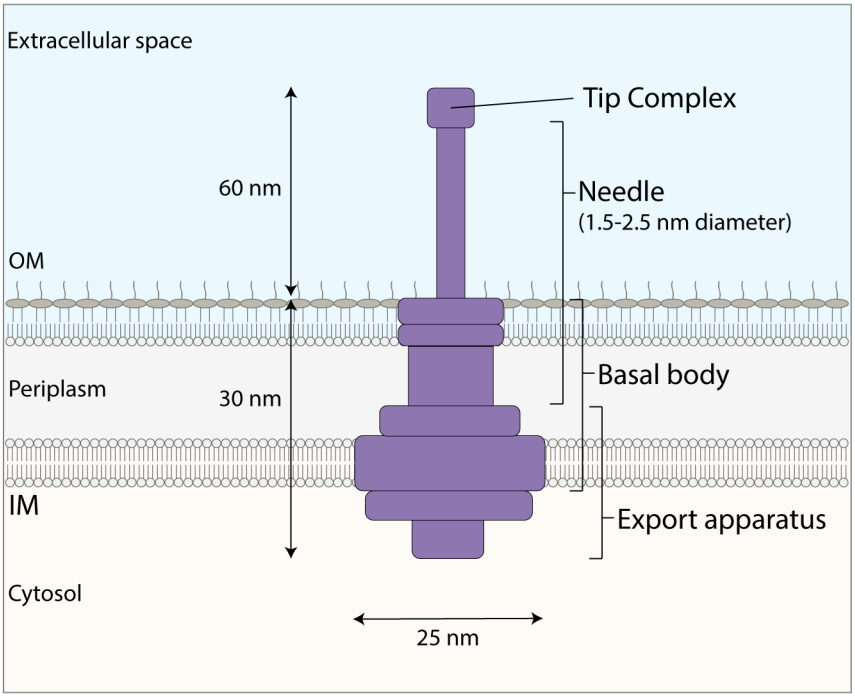


Figure 1.4 Structure of the T3SS.
The T3SS is composed of an export apparatus and basal body that span the inner and outer membranes, a needle that is anchored in the basal body and extends into the extracellular space, and a tip complex that caps the needle.

Table 1.5 Components of the SPI-1 T3SS apparatus.

Part of Apparatus	Component	Function	Reference
Export apparatus	InvA	Gate	[109],[117]
Export apparatus	InvC	ATPase	[109],[117]
Export apparatus	InvE	Gatekeeper	[109],[117]
Export apparatus	InvI	Stalk	[109],[117]
Export apparatus	OrgB	Stalk	[109],[117]
Export apparatus	SpaO	Sorting platform	[109],[117]
Export apparatus	SpaP	Sorting platform	[109],[117]
Export apparatus	SpaQ	Sorting platform	[109],[117]
Export apparatus	SpaR	Sorting platform	[109],[117]
Export apparatus	SpaS	Substrate specificity switch	[109],[117]
Basal body	InvG	OM ring (secretin)	[109],[117]
Basal body	InvH	Pilotin	[109],[117]
Basal body	PrgH	IM ring	[109],[117]
Basal body	PrgJ	Inner rod	[109],[117]
Basal body	PrgK	IM ring	[109],[117]
Needle	PrgI	Filament	[109],[117]
Tip	SipB	Translocon	[109],[117]
Tip	SipC	Translocon	[109],[117]
Tip	SipD	Tip	[109],[117]

1.3.3 SPI-1 T3SS Protein Secretion

Substrates of the type III secretion system are targeted for secretion by an N-terminal peptide sequence [8,86]. Each N-terminal tag contains a 15-20 amino acid signal sequence and a 20-154 amino acid chaperone binding domain [118]. The chaperone binding domain is required for targeting to the T3SS, and the secretion signal is required for export [86,118]. Secretion titer increases for polypeptides targeted to the T3SS when the cognate chaperone is co-expressed [8]. The chaperone is thought to perform a dual role by targeting cargo to the correct T3SS and maintaining an unfolded state, “priming” the cargo for secretion [118,119]. A list of SPI-1 T3SS substrates and their secretion signals is provided in Table 1.6. The various secretion signal-chaperone combinations show differential secretion of heterologous cargo, but this effect has not been systematically evaluated [8].

Table 1.6 Native substrates and secretion tags for the SPI-1 T3SS.

Substrate	Full-length Molecular Weight (kDa)	Total Amino Acids	Amino Acids in Secretion Signal	Chaperone	Used for Heterologous Protein Secretion?	Reference
AvrA	34	301	-	Unknown	No	[120]
InvJ	36	336	15	Unknown	Yes	[121]
SipA	74	685	169	InvB	Yes	[122],[8]
SipB	62	593	160	Possibly SicA	Yes	[123],[124]
SipC	43	409	121	SicA	Yes	[123],[125],[8]
SipD	37	343	-	Unknown	No	[126]
SopA	87	782	96	InvB	Yes	[127],[8]
SopB	62	561	168	SigE	Yes	[128],[129],[8]
SopD	36	317	40	None (self)	Yes	[130],[8]
SopE2	26	240	105	InvB	Yes	[131],[8]
SptP	60	543	159	SicP	Yes	[132],[8]

The overall mechanism of protein secretion via the T3SS is not well described, though several components of the process have been investigated [111,115]. Natively, rapid secretion of non-structural substrates is thought to be activated by direct contact with the mammalian epithelium [133]. Schlumberger *et al.* showed that docking to the mammalian epithelium triggered SipA and SopE2 secretion within one minute of contact [134]. That study also provided quantitative evidence that the number of secreted copies of each substrate is tightly controlled.

The energy for protein secretion via the SPI-1 T3SS is hypothesized to be provided by a combination of the ATPase InvC and the proton motive force (PMF) [135,136]. The substrate-chaperone complex binds the ATPase directly, inducing dissociation of the chaperone from the substrate and substrate unfolding to prepare for secretion [137]. However, ATPase function is dispensable for secretion via the T3SS, while the PMF is essential [138]. Though neither the contributions of the $\Delta\psi$ and ΔpH components of the PMF nor the exact mechanism of PMF influence on SPI-1 T3SS function have been investigated rigorously, the requirement for a functional PMF provides yet another environmental point of control.

1.4 Other Virulence Factors of *Salmonella enterica* Typhimurium

The SPI-1 T3SS is but one component of the pathogenic program of *S. enterica* Typhimurium. Genes that encode a second injectisome, promote adhesion to host cells, enhance colonization, and promote long-term survival within host cells are organized into four other pathogenicity islands and a virulence plasmid, pSLT [139,140]. Additionally, virulence-enhancing genes that encode fimbrial surface structures, the flagellum, and other biofilm-related structures are distributed throughout the genome [139]. Table 1.7

lists operons and pathogenicity islands that contribute to the virulence of *S. enterica* Typhimurium. These virulence factors and their interaction with SPI-1 are important to consider in the development of the SPI-1 T3SS as a heterologous protein production platform.

Table 1.7 Selected virulence-enhancing genetic elements in *S. enterica* Typhimurium.

Name	Length (kb)	Function	Notable Genes (Function)	Reference
SPI-2	40	SPI-2 T3SS; tetrathionate reduction pathway	*ssrAB* (master regulators)	[141]
SPI-3	17	Various	*mgtCB* (survival in macrophages)	[142]
SPI-4	25	Adhesion	*siiE* (adhesion to epithelium)	[143]
SPI-5	7	Various	*sopB* (SPI-1 T3SS substrate)	[144]
Flagellar operons	Various	fT3SS, chemotaxis	*flhDC* (master regulators); *fliZ* (SPI-1 T3SS regulator)	[98,145]
fim operon	9	Type I fimbriae; SPI-1 regulation via *fimYZ*	*fimYZ* (master regulators)	[146]
csg operon	5	T8SS (curli), biofilm formation	*csgD* (master regulator)	V
pSLT plasmid	95	Various	*spvR* (master regulator)	[147,148]

1.5 Conclusions and Engineering Goals

Despite remarkable advancements in biotechnology over the past forty years, many new protein products and those in development are challenging to produce in a cost-effective manner using the current suite of production organisms. Furthermore, the emerging concept of point-of-care, on-demand manufacturing for personalized medicine will require new production organisms and technologies because of the unique challenges of such a system [129,130]. Finally, much of the cost of a new protein product is derived from its product development. Secretion systems in Gram-negative bacteria have the potential to address these needs.

The SPI-1 T3SS shows particular promise as a heterologous protein secretion platform. It can secrete a wide range of heterologous proteins at titers up to hundreds of milligrams per liter, and secreted proteins can be recovered in pure, active forms after only one chromatography step [27,38]. Several challenges remain in adapting it for large-scale use, however. The “platform” is not established—design rules for its secretion tags, its compatibility with desired heterologous proteins, and ideal process conditions are not optimized or broadly characterized. Further, high-throughput screening methods and selections have not been developed for the system. Finally, using the SPI-1 T3SS as a protein production platform will require a production strain that is generally regarded as safe (GRAS). The SPI-1 T3SS has not been successfully expressed in a GRAS strain, and even the refactored T3SS produced by Song *et al.* exists in a strain with the remaining pathogenicity islands and other virulence-enhancing operons intact [65].

This thesis outlines efforts to advance the SPI-1 T3SS beyond initial engineering successes to create a robust protein production platform. First, I discuss design rules for rich growth media that increase secretion titer (Chapter 2). I extend those design rules to chemically defined media and discuss new process optimization challenges that arose from extending secretion beyond the standard eight-hour timeline (Chapter 3). In Chapter 4, I describe the challenges, successes, and failures of attempting to establish a universal high-throughput assay to measure secretion titer. I apply the process optimization knowledge acquired in Chapters 2-4 to secretion system optimization by creating a library construction and screening method to find high-secreting variants of the T3SS needle protein, PrgI (Chapter 5). Finally, I identify strain modifications that advance the development of a GRAS strain derived from *S. enterica* Typhimurium and simplify secretion system activation while retaining high secretion titers (Chapter 6).

Chapter 2: Medium Composition Affects Secretion Titer

The following is in preparation (Burdette, Wong, Tullman-Ercek)

2.1 Introduction

The ever-increasing diversity of commercial protein products coupled with the high cost of developing these products is driving engineering efforts to develop low-cost, easy-to-manipulate production systems. Bacterial hosts are robust, genetically tractable, and inexpensive to cultivate, but traditional intracellular expression strategies present several challenges. Products must be recovered from a complex lysate mixture, which requires several downstream purification steps [1,2]. Further, intracellular overexpression of heterologous proteins often causes aggregation in insoluble inclusion bodies. Initial product purity is often higher in inclusion bodies, but resolubilization and refolding processes must be developed for each product [26]. Finally, many heterologous proteins, including biomaterials such as spider silk, are difficult to express at high levels because they are toxic to the host [27,28]. Engineering bacteria to secrete heterologous proteins into the extracellular space eliminates these constraints.

The type III secretion system (T3SS) is a multimeric protein needle complex that spans the inner and outer membranes of the bacterial cell and secretes proteins from the cytoplasm to the extracellular space in a single step [86]. The T3SS is not required for cell viability, which facilitates engineering efforts and allows it to be used solely for heterologous cargo. The Salmonella pathogenicity island 1 (SPI-1) T3SS of *Salmonella enterica* Typhimurium has been successfully engineered to secrete high titers of several heterologous proteins, including spider silk monomers, antimicrobial peptides, and single chain variable fragments (scFvs) [8,27,38]. Despite these successes, secretion titers remain below 1 g/L, and T3SS function is sensitive to numerous extracellular inputs, including osmolarity, pH, and nutrient concentrations [98–100] (Table 1.4).

The SPI-1 regulatory network is complex—it receives input from regulatory systems that are essential for normal cellular function such as the DNA-binding protein HN-S, the Csr carbon storage regulatory system, and the osmolarity sensor OmpR/EnvZ [103,107,149]. Given this complexity, we hypothesized that instead of attempting to relieve T3SS repression at the genetic level, we could take an “outside-in” approach and engineer growth medium composition to identify and optimize environmental factors that promote consistently high secretion titers. Here, we screened common rich bacterial growth media and found that carbon source choice, buffering agents, and ionic content were critical factors for T3SS activity and secretion titer. We found that individually, non-ionic carbon sources repressed secretion via the T3SS while high ionic content increased secretion titers. The combination of a carbon source, a buffering agent, and high ionic content, however, had a synergistic effect to boost secretion titer further. SPI-1 T3SS transcriptional activity showed that this optimal combination increased secretion titer in part by elevating transcriptional activity and prolonging secretion system activation. An optimized combination of glycerol, potassium phosphate, and sodium chloride in LB medium increased secretion titer at least fourfold for several model proteins.

2.2 Methods

2.2.1 Strains and Growth Conditions

Strains and plasmids used in this work are listed in Table 2.1 and Table 2.2. Secretion experiments were started by growing a single colony from a fresh streak of a frozen glycerol stock in the lysogeny broth Lennox formulation (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) with appropriate antibiotics (34 µg/mL chloramphenicol, 50 µg/mL kanamycin) for 12-16 hours overnight in an orbital shaker at 37°C and 225 rpm. Overnight cultures were diluted 1:100 into the appropriate medium supplemented with 100 µg/mL isopropyl β-D-1-thiogalactopyranoside (IPTG) with appropriate antibiotics. All culturing steps were performed in 24-well deepwell plates (Axygen). Secretion was performed for 8 hours at 37°C and 225 rpm in an orbital shaker. Whole culture lysate samples for SDS-PAGE were prepared by adding cell suspension to Laemmli buffer [150] at the end of secretion. The secretion fraction was harvested by centrifuging cultures at 4000 *x g* for 10 minutes. SDS-PAGE samples for the secretion fraction were prepared by adding supernatant to Laemmli buffer. All SDS-PAGE samples were boiled at 95°C for 5 minutes immediately after preparation.

Table 2.1 Strains used in Chapter 2.

Strain Name	Comment	Reference
ASTE13	LT2-derived lab strain similar to DW01	This study; DW01 [65]
ASTE13 *invE::GFPmut2*	*GFPmut2* inserted immediately downstream of *invE* coding sequence	This study
ASTE13 *prgH::GFPmut2*	*GFPmut2* inserted immediately downstream of *prgH* coding sequence	This study
ASTE13 s*ipC::GFPmut2*	*GFPmut2* inserted immediately downstream of *sipC* coding sequence	This study
ASTE13 Δ*phoB*	*phoB* knockout	This study
ASTE13 Δ*sipD*	*sipD* knockout	Based on [52], newly constructed for this study

Table 2.2 Plasmids used in Chapter 2.

Plasmid Name	ORFs under inducible control		ORI	ab^R	Reference
P_{sic} *DH*	*sicP*	*sptP-DH-2xFLAG-6xHis*	colE1	cam	[41]
P_{sic} *MAG1*	*sicP*	*sptP-MAG1-2xFLAG-6xHis*	colE1	cam	This study
P_{sic} *14B7**	*sicP*	*sptP-14B7*-2xFLAG-6xHis*	colE1	cam	[38]
P_{sic} *rhGH*	*sicP*	*sptP-rhGH-2xFLAG-6xHis*	colE1	cam	This study
P_{lacUV5} *hilA*	*hilA*		p15a	kan	[41]

2.2.2 Medium Formulations

"LB" refers to a base medium formulation of 10 g/L tryptone and 5 g/L yeast extract, and "TB" is the standard Terrific Broth formulation: 12 g/L tryptone, 24 g/L yeast extract, 9.4 g/L K_2HPO_4, 2.2 g/L KH_2PO_4, and 0.4% w/v glycerol. LB-L is 10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl. Carbon sources were prepared as 20% w/v solutions, sterile filtered, and diluted into the medium to final concentrations of 0.4% w/v at the time of subculture. Media with buffers and salts were formulated by autoclaving LB medium at a 1.2X concentration and adding the appropriate volumes of 1M K_2HPO_4, 1M KH_2PO_4, 1M MOPS, or 5 M NaCl. Ultrafiltered water was added as necessary to achieve a final 1X LB concentration. The pH of all media was adjusted to 7.4 using HCl or NaOH as appropriate. Tryptone and yeast extract were sourced from BD Bacto. Conductivity and pH were measured using a Fisher Scientific accumet AB150 meter.

2.2.3 Plasmid Construction

PCR was performed with Phusion DNA polymerase using the primers listed in Table 2.3. Golden gate cloning was used to construct plasmids for this study [151]. Genes for proteins to be secreted were inserted into a modified pPROTet.133 backbone vector (BD Clontech) under the control of the *sic* promoter [8]. All secretion plasmids expressed the SptP chaperone *sicP* and the *sptP* secretion signal (nucleotides 1 to 477). The SptP secretion tag was fused N-terminal to the protein of interest, and 2xFLAG and 6xHis tags were fused C-terminal to the protein of interest. The gene for rhGH was ordered from Twist Biosciences with overhangs compatible with golden gate cloning. All cloning was done in *E. coli* DH10B cells, and all DNA sequences were confirmed by Sanger sequencing (Quintara).

2.2.4 Recombineering

Recombineering was performed in *S. enterica* Typhimurium ASTE13 as described by Thomason *et al.* [152]. Briefly, a *cat-sacB* cassette conferring chloramphenicol resistance and sucrose sensitivity was amplified using primers with 40 bp of homology 5' and 3' to the locus of interest. The *GFPmut2* gene was amplified using primers containing the same 40 bp of homology 5' and 3' to the locus of interest as used for the *cat-sacB* cassette. PCR was performed with Phusion DNA polymerase with the primers listed in Table 2.3. *S. enterica* Typhimurium ASTE13 was first transformed with pSIM6. A first round of recombineering was performed to insert the *cat-sacB* cassette at the locus of interest, and a second round of recombineering replaced the *cat-sacB* cassette with an appropriate DNA product. The replacement DNA was *GFPmut2* for transcriptional fusions, a 60 bp oligo containing the first and last 30 bp of *phoB* for the *phoB* gene knockout, and a 200 bp double-stranded PCR product containing the first and last 30 bp of *sipD* flanked 5' and 3' by 70 bp of homology to the *sipD* genetic locus. The genomic modifications were confirmed by Sanger sequencing (Quintara), and the strains were cured of pSIM6. Primers used for recombineering are listed in Table 2.3.

Table 2.3 Primers used in Chapter 2.

Sequence	Amplicon	Used to Construct
AATGGCAGAACAGCGTCGTACTATTGAAAAGC TGTCTTAA tgtgacggaagatcacttcg	*cat-sacB*	ASTE13 *invE::gfpmut2*
GAGAAAGCAGCACTATAGGTATCCTGTTAATA TTAAA atcaaagggaaaactgtccatat	*cat-sacB*	ASTE13 *invE::gfpmut2*
AATGGCAGAACAGCGTCGTACTATTGAAAAGC TGTCTTAA attaaagaggagaaaggtcatgag	*gfpmut2*	ASTE13 *invE::gfpmut2*
GTAGAGAAAGCAGCACTATAGGTATCCTGTTA ATATTAAA ttatttgtatagttcatccatgccatg	*gfpmut2*	ASTE13 *invE::gfpmut2*
AATGAGCCCAGGCCATTGGTATTTCCCAAGCC CACTTTAA tgtgacggaagatcacttcg	*cat-sacB*	ASTE13 *prgH::gfpmut2*
AAGGTGTTGCCATAATGACTTCCTTATTTACGT TAAA atcaaagggaaaactgtccatat	*cat-sacB*	ASTE13 *prgH::gfpmut2*
AATGAGCCCAGGCCATTGGTATTTCCCAAGCC CACTTTAA attaaagaggagaaaggtcatgag	*gfpmut2*	ASTE13 *prgH::gfpmut2*
ACCAAGGTGTTGCCATAATGACTTCCTTATTTA CGTTAAA ttatttgtatagttcatccatgccatg	*gfpmut2*	ASTE13 *prgH::gfpmut2*
ATCCGCACTCGCTGCTATCGCAGGCAATATTC GCGCTTAA tgtgacggaagatcacttcg	*cat-sacB*	ASTE13 *sipC::gfpmut2*
AATCACACCCATGATGGCGTATAGATGACCTT TCAGA atcaaagggaaaactgtccatat	*cat-sacB*	ASTE13 *sipC::gfpmut2*
ATCCGCACTCGCTGCTATCGCAGGCAATATTC GCGCTTAA attaaagaggagaaaggtcatgag	*gfpmut2*	ASTE13 *sipC::gfpmut2*
TTAAATCACACCCATGATGGCGTATAGATGAC CTTTCAGA ttatttgtatagttcatccatgccatg	*gfpmut2*	ASTE13 *sipC::gfpmut2*
AAATTATGGCGAGACGTATTCTGGTCGTAGAA GATGAGGC tgtgacggaagatcacttcg	*cat-sacB*	ASTE13 Δ*phoB*
GGCATTAAAAGCGGGTCGAAAAACGATACCCT GTCCCGC atcaaagggaaaactgtccat	*cat-sacB*	ASTE13 Δ*phoB*
ATGGCGAGACGTATTCTGGTCGTAGAAGATGA GGCGCGGGACAGGGTATCGTTTTTCGACCCG CTTTTAA	N/A	ASTE13 Δ*phoB*
TTTAATCGCGCTCCTGATGGCGAACTGGGGAT ATTATGCTTAATATTCAA tgtgacggaagatcacttcg	*cat-sacB*	ASTE13 Δ*sipD*
CTTACACTTGTAACCATTATTAATATCCTCTTC TGTTATCCTTGCAGGAA atcaaagggaaaactgtccatat	*cat-sacB*	ASTE13 Δ*sipD*
TCTGAAAGGTCATCTATACGCCATCATGGGTG TGATTTAATCGCGCTCCTGATGGCGAACTGGG GATATT atgcttaatattcaattcctgcaaggataa	self	ASTE13 Δ*sipD*
TCTGCATACCTGGCATTATGACGGGGGGCTG AGTCCTTACACTTGTAACCATTATTAATATCCT CTTCTG ttatccttgcaggaattgaatattaagcat	self	ASTE13 Δ*sipD*

2.2.5 Protein Separation and Western Blotting

Samples were separated by SDS-PAGE and transferred to a polyvinylidene fluoride membrane (PVDF, Millipore) for western blotting using the Bio-Rad Criterion blotter. Samples were diluted such that all band signals were within twofold of the average signal across the blot. Membranes were probed with mouse anti-FLAG per manufacturer's instructions (Sigma Aldrich). To facilitate chemiluminescent detection, a secondary labeling step was performed with goat anti-mouse IgG (H+L) HRP conjugate according to manufacturer's instructions (Thermo Fisher). Bands were detected with the SuperSignal West Pico Plus substrate (Thermo Fisher) and a ChemiDoc XRS + imaging system (Bio-Rad).

2.2.6 Protein Quantification

All relative protein quantities were calculated by performing densitometry using Image Lab software (Bio-Rad) and normalizing to the average of the replicates of the specified normalization condition. Relative protein amounts were corrected for dilution if appropriate. Absolute secretion titers were measured by performing SDS-PAGE, staining with Coomassie according to Studier [153], and calculating densitometry relative to a bovine serum albumin standard curve (Thermo). Background was calculated by averaging the signal at the same molecular weight as the protein of interest across all other lanes containing secreted fractions and then subtracted from the signal of the protein of interest. Error bars are standard deviation on three biological replicates unless otherwise specified.

2.2.7 Flow Cytometry

ASTE13 strains carrying *invE::gfpmut2*, *prgH::gfpmut2*, or *sipC::gfpmut2* transcriptional fusions and P_{lacUV5} *hilA* were grown and induced as specified in "Strains and Growth Conditions". Samples were prepared by diluting cultures to an optical density at 600 nm of 0.005 to 0.05 in PBS with 1 mg/mL kanamycin in round-bottom 96-well plates (Greiner Bio-One #650101). Plates were sealed and stored at 4°C for analysis. For each sample, an Attune NxT flow cytometer (Life Technologies) was used to collect 10,000 events within a gated population determined to be cells. Data was analyzed using FlowJo 10.5.3 (TreeStar, Inc.). The experiment was performed in biological triplicate, and error bars represent standard error.

2.3 Results

2.3.1 Growth medium affects secretion titer

Traditionally, secretion via the SPI-1 T3SS is activated by shifting cells to high osmolarity, low aeration growth conditions [8,154]. These growth conditions yield low secretion titers, however, because maximum cell density is low and only ~33% of cells have T3SS activity. In prior work, we showed that overexpressing the SPI-1 T3SS master regulator HilA solved these challenges and increased secretion titer by an order of magnitude [41]. HilA overexpression activated the T3SS in more than 90% of the population, and it provided higher cell densities by facilitating secretion in low-salt (5 g/L NaCl) LB-Lennox (LB-L) medium with moderate aeration. In subsequent work, combining HilA overexpression, moderate aeration, and the high osmolarity medium traditionally used to induce the SPI-1 T3SS, LB-IM (17 g/L NaCl), increased secretion titer further [38].

An optimal medium for protein production is composed of nutrients that maximize cell density and protein production per cell, or specific productivity. The increase in secretion titer observed in LB-IM with HilA overexpression highlights a unique challenge for protein production via the T3SS, however—secretion titer is increased in high-salt conditions that limit cell density. As a result, an ideal growth medium for protein production via the T3SS will need to strike the optimal balance between maximizing cell density and specific productivity.

Though numerous environmental inputs to the T3SS are documented (Table 1.4), they were not characterized in the context of our engineered system. Thus, to expand our knowledge of medium components that affect secretion titer, we began simply by measuring expression and secretion titer of a model protein in the common bacterial growth media 2X YT and terrific broth (TB). The model protein, or protein of interest (POI), was the soluble catalytic DH domain from the human protein intersectin-1L fused C-terminal to the T3SS secretion tag SptP to facilitate secretion [8,155]. SptP-DH-2xFLAG-6xHis was co-transformed with a HilA overexpression plasmid to activate the T3SS and maximize secretion titer in ASTE13, a derivative of *S. enterica* Typhimurium DW01 [65]. The native SPI-1 promoter P_{sic} controlled SptP-DH-2xFLAG-6xHis expression so that T3SS and POI expression could be activated with a single induction event [8,41]. Secretion titers were measured relative to LB-L using semi-quantitative western blotting.

Bulk secretion titer increased threefold in 2X YT and fivefold in TB (Figure 2.1A*i*). 2X YT and TB increased expression of SptP-DH threefold and fourfold (Figure 2.1A*ii*), suggesting that an increase in secretion titer is correlated with an increase in expression of the target protein. The relative increase in secretion titer was higher than the relative increase in SptP-DH expression in TB, indicating that some components in TB caused a higher percentage of expressed protein to be secreted. We will henceforth refer to the fraction of expressed protein that is secreted as “secretion efficiency”.

TB contains two defined components that are not present in LB-L: 0.4% w/v glycerol and 89 mM potassium phosphate (Figure 2.1B). We hypothesized that adding those components to LB-L could recapitulate the increase in secretion titer and efficiency observed in TB. Surprisingly, separate addition of those components to LB-L revealed a competing dynamic—glycerol repressed secretion, while phosphate buffer promoted secretion. Combining the two components in LB-L, however, had a synergistic effect, matching the fivefold increase in secretion titer observed in TB.

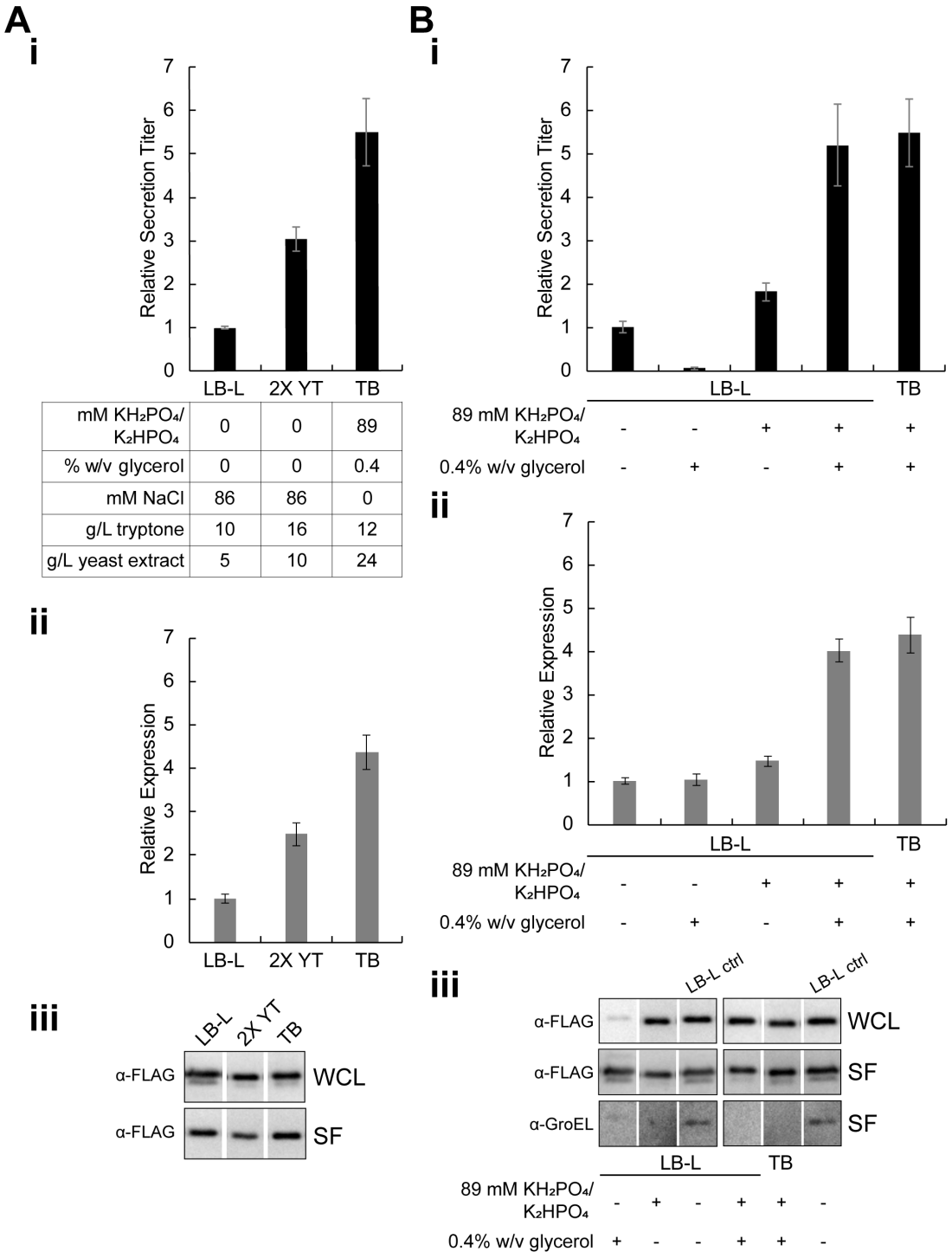


Figure 2.1 Medium composition affects secretion titer.
A Relative bulk secretion titer (i), expression (*ii*), and representative western blots (*iii*) of SptP-DH-2xFLAG-6xHis in LB-L, 2X YT, and TB. **B** Bulk secretion titer (i), expression (*ii*), and representative western blots (*iii*) of SptP-DH-2xFLAG-6xHis secreted LB-L supplemented with the defined components in TB. All western blots are representative of four biological replicates. Samples were diluted to fall within the linear range of the LB-L signal. Boxed bands are from the same blot but were rearranged for clarity. "WCL" is whole culture lysate and "SF" is secreted fraction. Error bars represent one standard deviation.

2.3.2 Non-ionic carbon sources decrease secretion titer

SPI-1 T3SS transcriptional activity is repressed in the presence of glucose [102], so the discovery that glycerol negatively impacted secretion titer led us to screen a panel of carbon sources in LB-L to determine the nature of the effect of carbon sources on T3SS activity. We included sodium succinate because we use it in minimal media formulations for *S. enterica* growth [156]. We evaluated SptP-DH-2xFLAG-6xHis secretion titer, expression, and T3SS transcriptional activity for each carbon source in LB-L. Transcriptional activity over time was measured by performing flow cytometry on strains with GFP integrated into one of the SPI-1 *inv*, *prg*, or *sic* loci. The *inv*, *prg*, and *sic* operons code for regulatory, structural, and natively secreted proteins, respectively [96].

SptP-DH-2xFLAG-6xHis secretion titer decreased relative to no added carbon source in the presence of all carbon sources except succinate, and SptP-DH-2xFLAG-6xHis expression decreased in the presence of all carbon sources except glycerol and succinate (Figure 2.2A). The trends in SptP-DH-2xFLAG-6xHis expression generally correlated with the maximum level of transcriptional activation at SPI-1 T3SS loci (Figure 2.2B), though increased transcriptional activity relative to no added carbon source did not translate to increased expression. Transcriptional activity was less explanatory for secretion titer. Low transcriptional activity did trend with low secretion titer, but glycerol and sorbitol had transcriptional activity similar to no added carbon source and reduced secretion titer. Transcriptional activity in cultures containing glucose and galactose ceased at least an hour earlier than the other conditions tested, which might contribute to the negligible secretion titer observed in those cultures.

The decrease in secretion titer observed in the presence of all carbon sources except succinate might also be explained by acidification of the medium (Figure 2.2A*i*), an environmental condition known to repress SPI-1 T3SS activity [99,100]. The increase in pH in cultures containing succinate suggests that amino acid catabolism occurred, which could result from succinate being exhausted or succinate not being catabolized at all. If the latter was true, its distinct could be solely due to its ionic character.

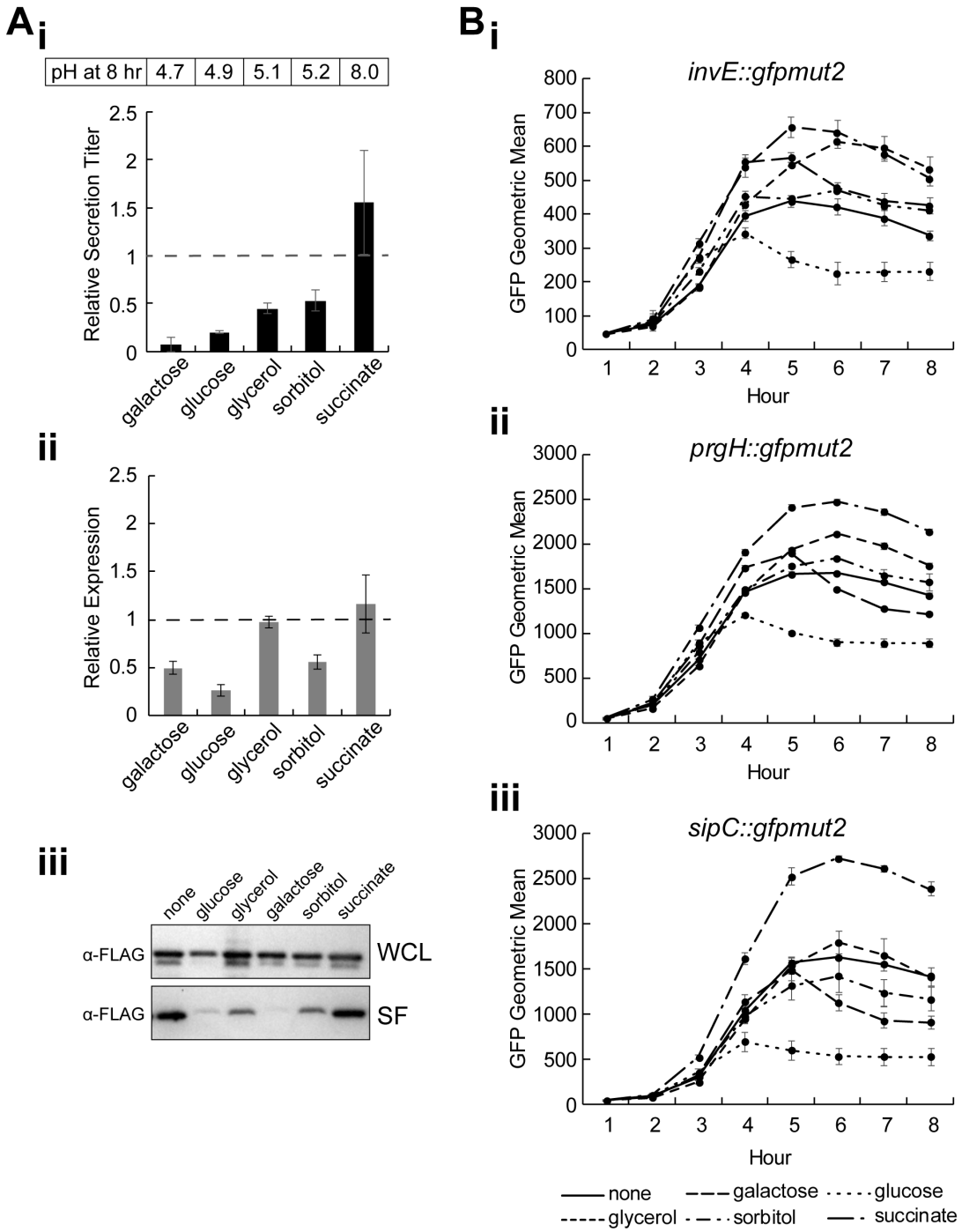


Figure 2.2 Non-ionic carbon sources decrease secretion titer in LB-L.
A Relative bulk secretion titer (*i*), expression (*ii*), and representative western blots (*iii*) of SptP-DH-2xFLAG-6xHis. Error bars represent one standard deviation. Expression and secretion titer were normalized to LB-L with no carbon source using densitometry measurements on western blots. Western blots are representative of three biological replicates. The pH of the secreted fraction was recorded at the end of the experiment. **B** Transcriptional activity at the *inv* (*i*), *prg* (*ii*), and *sic* (*iii*) loci was measured by performing flow cytometry on strains with transcriptional fusions of GFPmut2 at the appropriate locus. The *invE::gfpmut2* data was rescaled to highlight differences. Error bars represent standard error of the GFP geometric mean for three biological replicates.

2.3.3 Secretion titer increases with ionic content

Phosphate is a unique buffer species because in addition to providing buffer capacity, it contributes significantly to the ionic content of the media. To decouple the effects of buffering and increased ionic content, we compared secretion titer in media containing a base of 10 g/L tryptone and 5 g/L yeast extract supplemented with a range of concentrations of potassium phosphate, MOPS, sodium chloride, or MOPS supplemented with sodium chloride (Table 2.4). We selected MOPS as an alternative buffer species because it is the main buffer component in a defined medium explicitly designed for *S. enterica* cultivation [157], it has a minimal contribution to ionic strength, and it is one of Good's buffers [158]. To mimic the simultaneous contributions of buffering and ionic content inherent to phosphate, we supplemented MOPS with sodium chloride. Finally, sodium chloride changes ionic content without a buffering effect. We monitored ionic content by measuring conductivity. The concentrations of sodium chloride with and without MOPS were chosen to match the conductivities of the specified concentrations of potassium phosphate. Expression and secretion titer were compared to LB-L with no additives using semi-quantitative western blotting.

Table 2.4 Medium additives for Figure 2.3.

Medium	[KH_2PO_4/ K_2HPO_4] (mM)	[MOPS] (mM)	[NaCl] (mM)	Conductivity (mS/cm)
1	10	0	0	4.0
2	20	0	0	5.6
3	40	0	0	8.4
4	80	0	0	14
5	160	0	0	25
6	0	10	0	2.5
7	0	20	0	2.8
8	0	40	0	3.5
9	0	80	0	4.7
10	0	160	0	7.1
11	0	10	12	3.8
12	0	20	24	5.4
13	0	40	48	8.1
14	0	80	97	14
15	0	160	193	23
16	0	0	14	3.5
17	0	0	28	4.9
18	0	0	56	7.5
19	0	0	112	13
20	0	0	224	23
LB-L	0	0	86	11
LB-IM	0	0	291	31

Expression and secretion titer increased linearly with conductivity. The correlation was stronger for secretion titer (R^2 = 0.9) than expression (R^2 = 0.8). LB-IM and the highest concentrations of potassium phosphate and MOPS+NaCl provided the highest secretion titers. Expression was maximal in the highest concentrations of sodium chloride, MOPS+NaCl, and potassium phosphate.

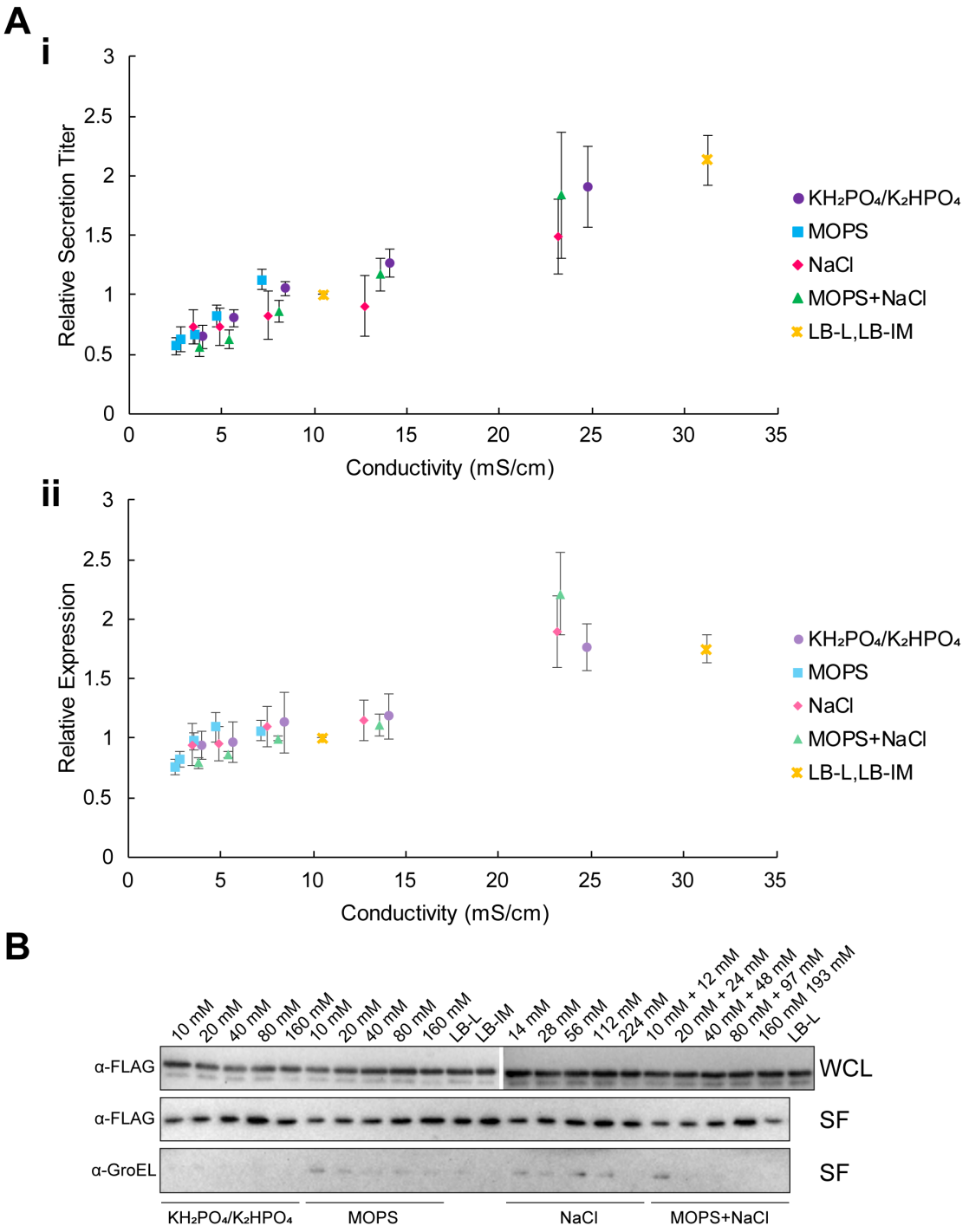


Figure 2.3 Expression and secretion increase with ionic content.
A Relative bulk secretion titer (*i*) and expression (*ii*) of SptP-DH-2xFLAG in media containing 10 g/L tryptone, 5 g/L yeast extract, and the additives specified in Table 2.4 plotted versus medium conductivity. Bulk expression and secretion titer were normalized to LB-L with no additives (dotted lines). Error bars represent one standard deviation. **B** Western blots are representative of three biological replicates. Samples at the highest additive concentrations were diluted 0.5X to fall within the linear range of the LB-L signal. Boxed bands are from the same blot but were rearranged for clarity. "WCL" is whole culture lysate and "SF" is the secreted fraction.

2.3.4 Carbon sources and buffers have a synergistic effect on expression and secretion titer

Independent addition of buffers, salts and carbon sources showed that secretion titer increases with ionic content and decreases with carbon sources that cause acidification of the culture. If acidification of the culture was the sole cause of decreased secretion titer, addition of a buffer would return secretion titer to levels at least equivalent to no added carbon source. Potassium phosphate and glycerol in LB-L had a synergistic effect, however, suggesting that high ionic strength, buffering, and added carbon sources are a critical combination for increased secretion titer.

To determine if the synergistic effect could be a result of any combination of carbon source and buffer at high ionic strength or whether the effect was specific to the combination of phosphate and glycerol, we measured secretion of SptP-DH-2xFLAG-6xHis in LB media containing potassium phosphate, MOPS, sodium chloride, and MOPS plus sodium chloride with added glycerol, glucose, or succinate. The potassium phosphate and MOPS concentrations were 90 mM to match the concentration of potassium phosphate in TB, and the concentrations of sodium chloride with and without MOPS buffer were chosen to approximate the conductivity of 90 mM potassium phosphate (Table 2.5). Expression and secretion titer of SptP-DH-2xFLAG-6xHis were measured in comparison to LB-L with no added carbon source using semi-quantitative western blotting.

Table 2.5 Conductivity of LB supplemented with buffers, salts, and carbon sources.

Medium	[KH_2PO_4/K_2HPO_4] (mM)	[MOPS] (mM)	[NaCl] (mM)	Conductivity (mS/cm)			
				None	Glucose	Glycerol	Succinate
KH_2PO_4/K_2HPO_4	90	0	0	14	14	14	18
MOPS	0	90	0	5.0	5.4	5.4	9.9
MOPS+NaCl	0	90	199	20	20	20	24
NaCl	0	0	234	22	22	21	26

Buffering was essential to maintain secretion in the presence of glucose and glycerol, and the combination of buffering and increased ionic content was necessary to increase expression and secretion titer (Figure 2.4A). LB-MOPS with glucose or glycerol produced secretion titers similar to LB-L with no added carbon source, while secretion titer decreased in LB-NaCl with glucose or glycerol. The decrease in secretion was likely due to acidification of the extracellular environment, as observed in Figure 2.2 (Table 2.6). Both LB-(KH_2PO_4/K_2HPO_4) and LB-(MOPS+NaCl) with added glucose or glycerol provided at least a threefold increase in secretion titer. The combination of phosphate and glycerol appeared to have a specific beneficial effect on secretion titer—it provided the largest increase in secretion titer even though expression in LB-(MOPS+NaCl) with glycerol increased by a larger margin (Figure 2.4B).

We want to maximize secretion efficiency in addition to bulk secretion titer, so we evaluated expression and secretion on a per cell basis for the combinations of media additives. Total protein expression per cell increased with ionic content if solution pH remained above 5.0 (Table 2.6) and was highest in LB-(MOPS+NaCl) with glucose and glycerol. Secretion per cell showed a more complex pattern. LB-(KH_2PO_4/K_2HPO_4) and

LB-(MOPS+NaCl) provided similar increases in secretion per cell in combination with glucose, but glycerol and LB-(KH_2PO_4/K_2HPO_4) had a greater synergistic effect than glycerol and LB-(MOPS+NaCl). Secretion efficiency in LB-(KH_2PO_4/K_2HPO_4) was also superior to that in LB-(MOPS+NaCl) with glycerol. Relative secretion per cell was higher than relative bulk secretion for media containing succinate because the high ionic strength of those media limited cell density.

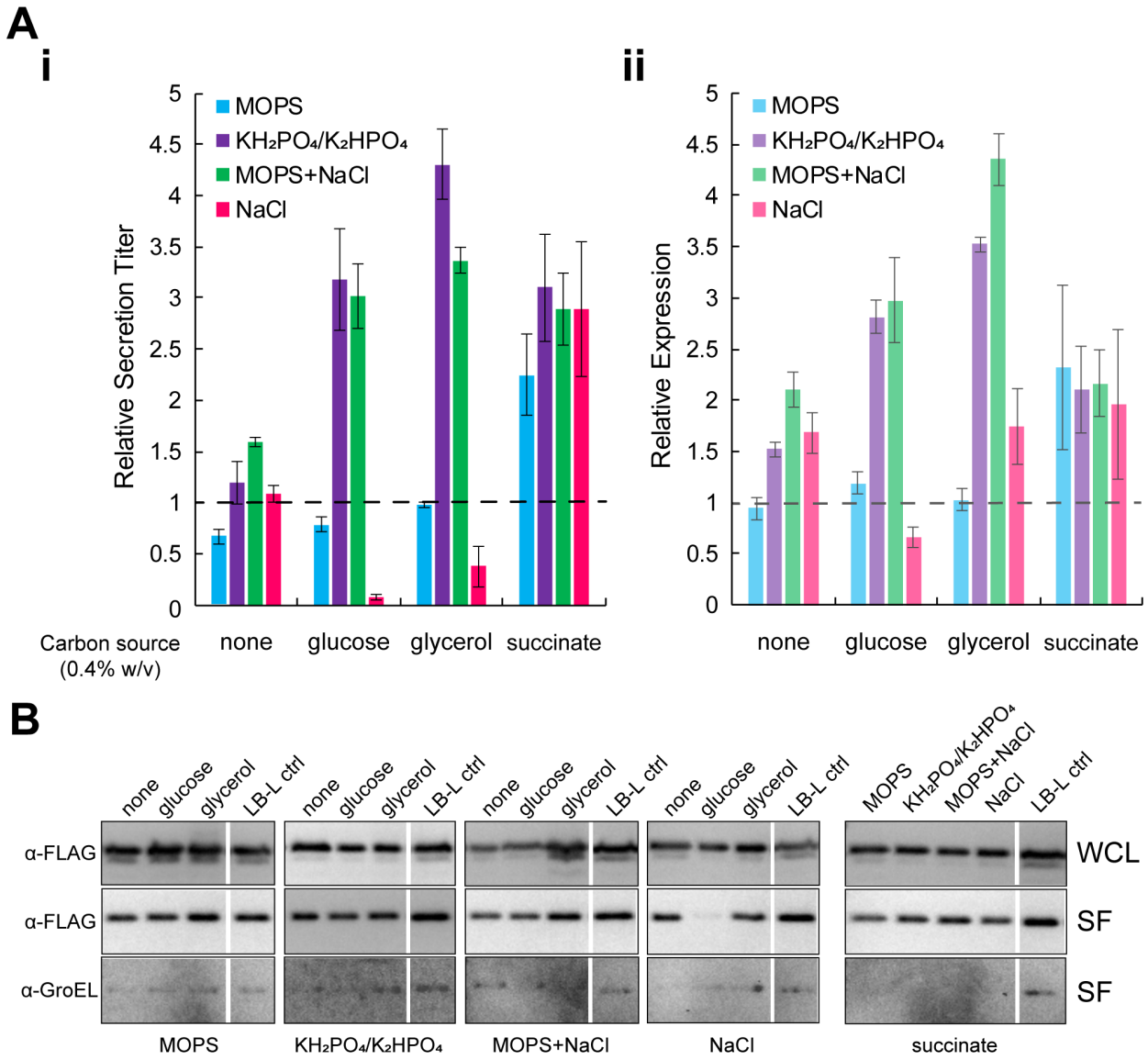


Figure 2.4 Sugars and buffers have a synergistic effect on expression and secretion titer.
A Relative bulk secretion titer (*i*) and expression (*ii*) of SptP-DH-2xFLAG-6xHis secreted in media containing 10 g/L tryptone, 5 g/L yeast extract, and the additives listed in Table 2.5. Bulk expression and secretion titer were normalized to LB-L with no additives (dotted lines). **B** Western blots are representative of three biological replicates. Samples were diluted to fall within the linear range of the LB-L signal. Boxed bands are from the same blot but were rearranged for clarity. "WCL" is whole culture lysate and "SF" is secreted fraction.

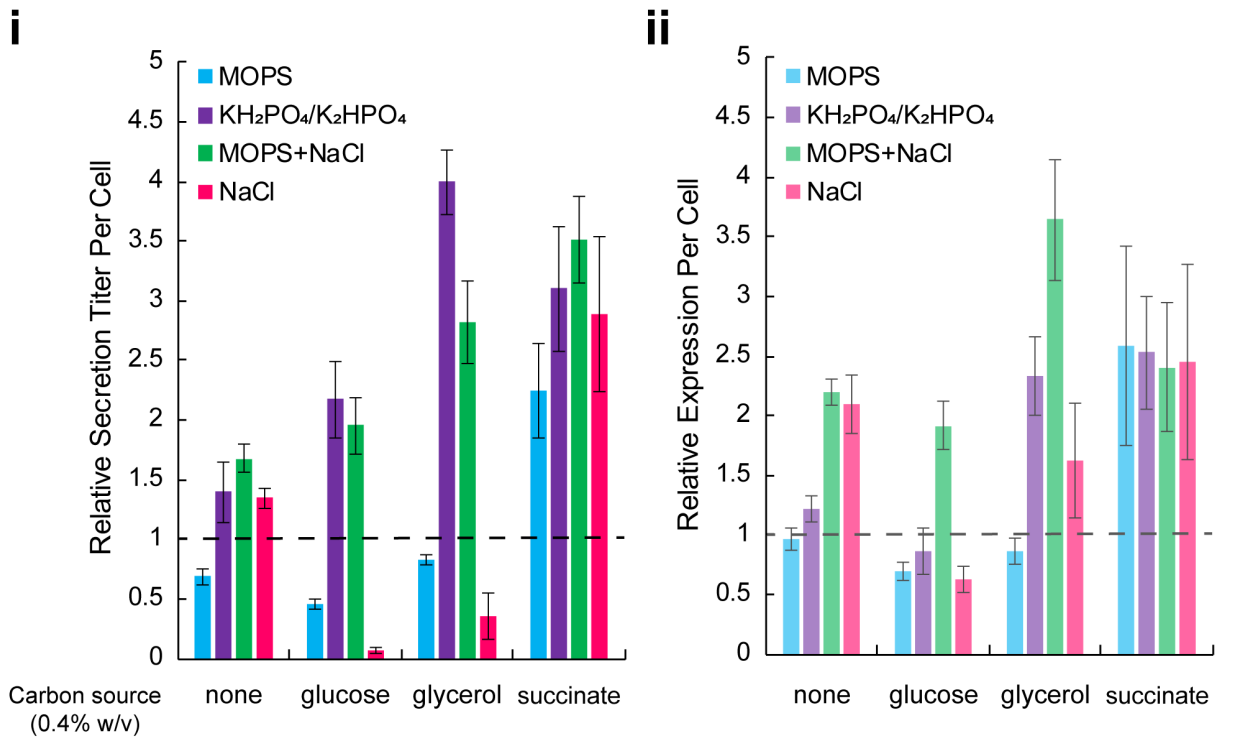


Figure 2.5 Expression and secretion per cell vary among combinations of buffers, salts and carbon sources.
(*i*) Relative secretion titer per cell and (*ii*) relative expression per cell of SptP-DH-2xFLAG-6xHis in the media listed in Table 2.5. This is the same data set as Figure 2.4. Expression and secretion per cell were calculated by dividing densitometry by OD_{600nm} and normalizing to LB-L with no additives. Error bars represent one standard deviation for three biological replicates.

Table 2.6 pH of the Secreted Fraction from Figure 2.4.

Medium	[KH_2PO_4/K_2HPO_4] (mM)	[MOPS] (mM)	[NaCl] (mM)	pH of Secreted Fraction			
				None	Glucose	Glycerol	Succinate
MOPS	0	90	0	7.6±0.0	6.7±0.1	6.8±0.0	7.6±0.2
KH_2PO_4/K_2HPO_4	90	0	0	7.5±0.0	6.5±0.0	6.6±0.0	7.6±0.1
MOPS+NaCl	0	90	199	7.3±0.0	6.3±0.0	6.8±0.0	7.4±0.1
NaCl	0	0	234	7.9±0.1	4.6±0.0	4.8±0.0	8.0±0.1

2.3.5 Carbon sources and buffers affect T3SS transcriptional activity

Flow cytometry performed on ASTE13 strains containing transcriptional fusions of GFP at the *sic* (Figure 2.6), *prg*, and *inv* (Figure 2.7) loci in the media listed in Table 2.5 showed that SPI-1 transcriptional activity varied with medium composition. Transcriptional activity increased with conductivity for no added carbon source or succinate. In media containing glucose or glycerol, the trends were more complex. Transcriptional activity was repressed in LB-NaCl and highest in LB-(MOPS+NaCl). The positive effect of LB-(KH_2PO_4/K_2HPO_4) with glucose or glycerol was stronger at the *prg*

operon than either the *inv* or *sic* operons. Relative to media with no added carbon source or succinate, glucose and glycerol also prolonged transcriptional activation in buffered media while causing an early decrease in unbuffered media. Histograms for the data in Figure 2.6 and Figure 2.7 are available in Appendix A.

Flow cytometry measures transcriptional activity on a per cell basis, so it should be compared to expression and secretion per cell to identify correlations. Increased transcriptional activity relative to LB-L with no added carbon source corresponded with increased expression per cell of SptP-DH-2xFLAG-6xHis (Figure 2.5, Figure 2.6, Figure 2.7). As in Figure 2.2, transcriptional activity was less explanatory for secretion titer, but media that prolonged and increased transcriptional activation also increased secretion titer. Transcriptional activity did not explain why potassium phosphate and glycerol produced a higher increase in secretion titer than LB-(MOPS+NaCl) and glycerol—the transcriptional activity better corresponds with the expression data, and both media prolonged transcriptional activity compared to LB-L, LB-MOPS, and LB-NaCl.

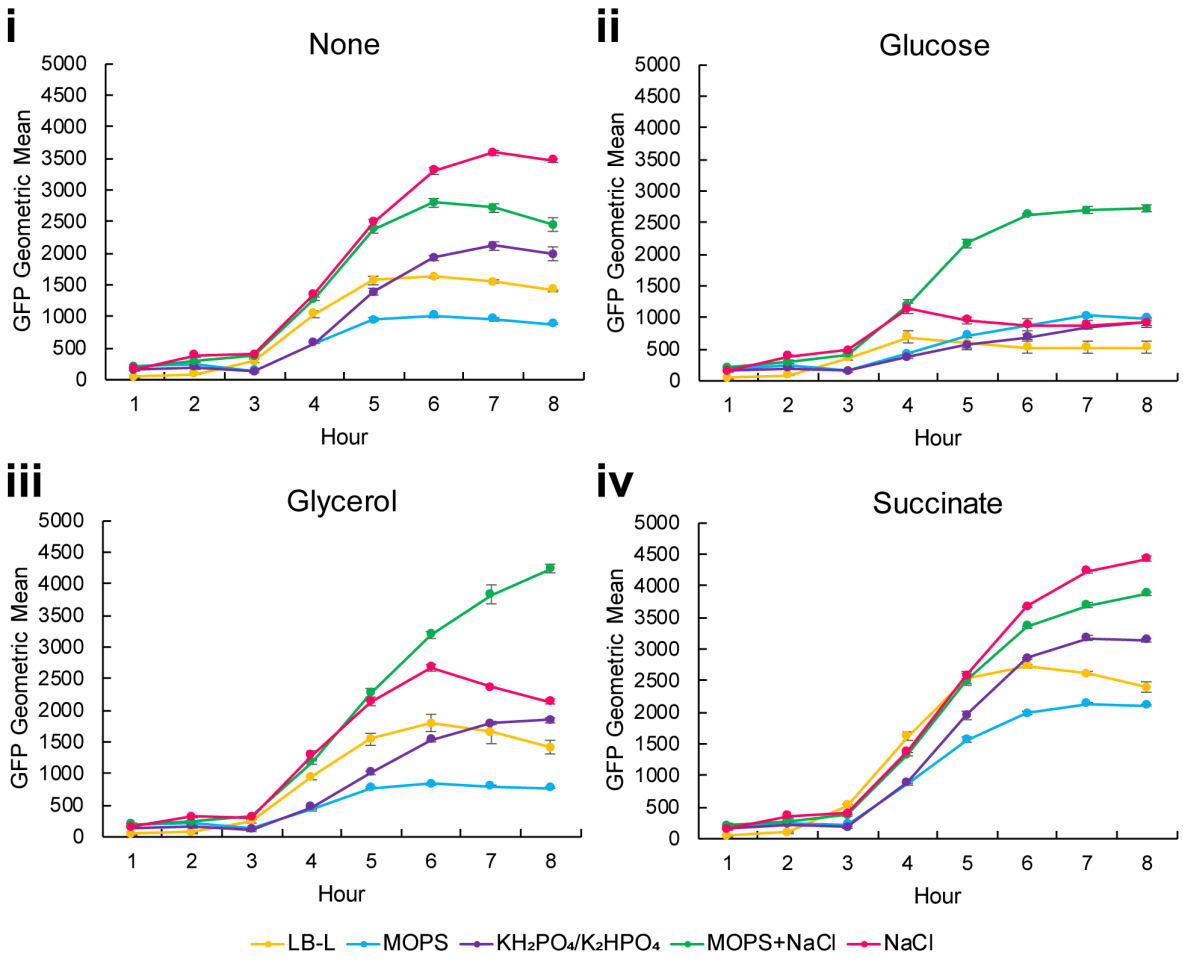


Figure 2.6 Carbon sources and buffers affect T3SS transcriptional activity.
Transcriptional activity at the *sic* locus for (i) no added carbon source, (ii) glucose, (iii) glycerol, and (iv) succinate in the media listed in Table 2.5 was measured over time by performing flow cytometry on ASTE13 strains containing a *sipC*::*gfpmut2* transcriptional fusion. Error bars represent standard error of the GFP geometric mean for three biological replicates.

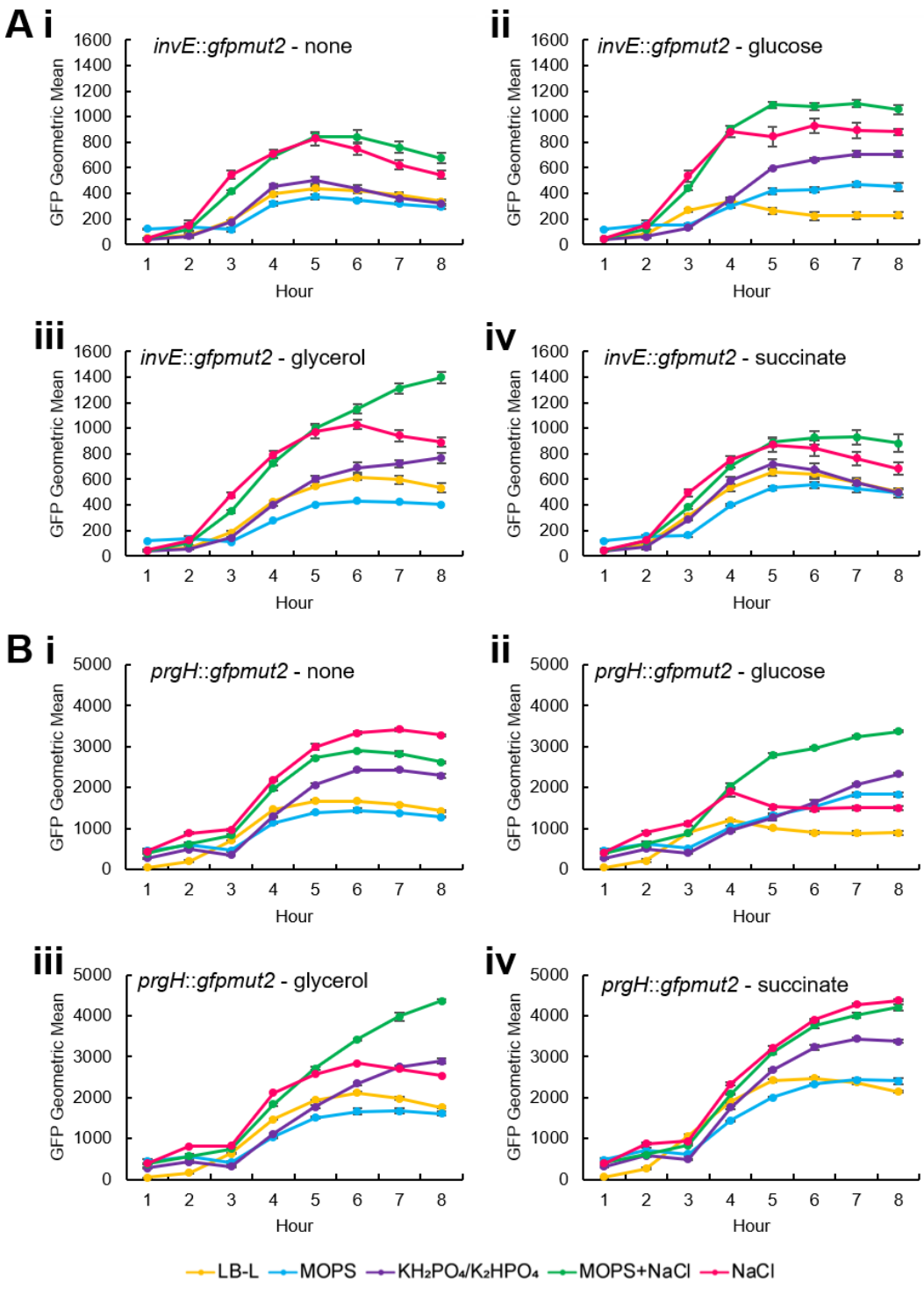


Figure 2.7 Carbon sources and buffers affect T3SS transcriptional activity.
A Transcriptional activity at the *inv* locus in (*i*) no added carbon source, (*ii*) glucose, (*iii*) glycerol, and (*iv*) succinate. **B** Transcriptional activity at the *prg* locus in (*i*) no added carbon source, (*ii*) glucose, (*iii*) glycerol, and (*iv*) succinate. Flow cytometry was performed on ASTE13 strains containing *invE*::*gfpmut2* (A) or *prgH*::*gfpmut2* (B) transcriptional fusions. Note that the *invE::gfpmut2* data has a different scale. Error bars represent standard error of the GFP geometric mean for three biological replicates.

2.3.6 The phosphate-sensing system PhoBR has no effect on expression or secretion titer with the engineered T3SS

Phosphate could affect secretion titer in three ways—by buffering the media, by increasing ionic content, or by inducing a specific regulatory cascade to increase SPI-1 activation. The two-component system PhoBR has been implicated in repressing SPI-1 activity in response to phosphate starvation [99,101]. We are studying phosphate in great excess of the starvation condition, so we hypothesized that our conditions were more likely to impact SPI-1 activity indirectly via metabolic pathways or osmolarity sensing [106]. Nevertheless, we compared secretion titer of SptP-DH-2xFLAG-6xHis in LB-(KH_2PO_4/K_2HPO_4) with 0.4% w/v glycerol in WT and Δ*phoB* strains (Figure 2.8). We were surprised to discover that deleting PhoB had a small negative impact on secretion titer in LB-(KH_2PO_4/K_2HPO_4) with glycerol. It had no other significant effects.

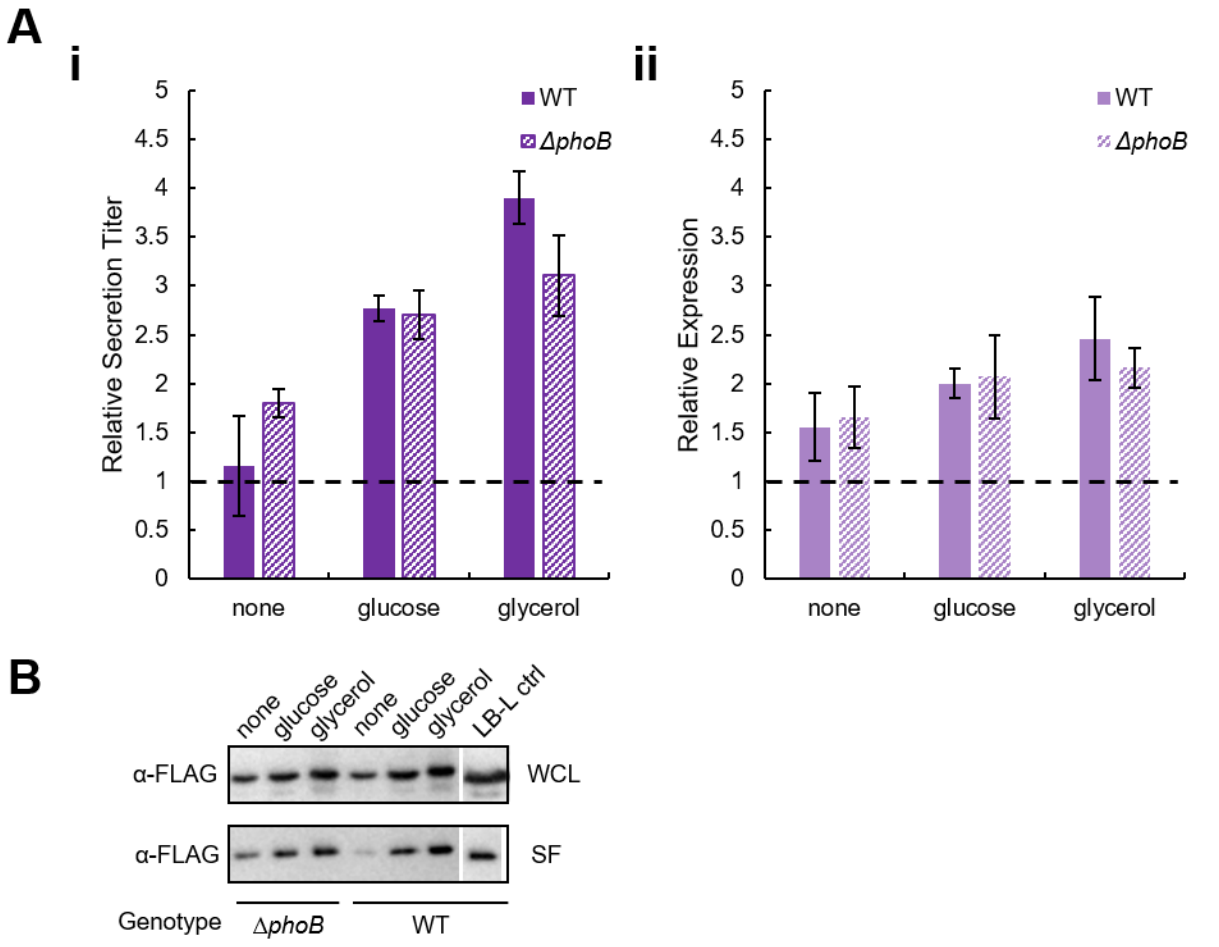
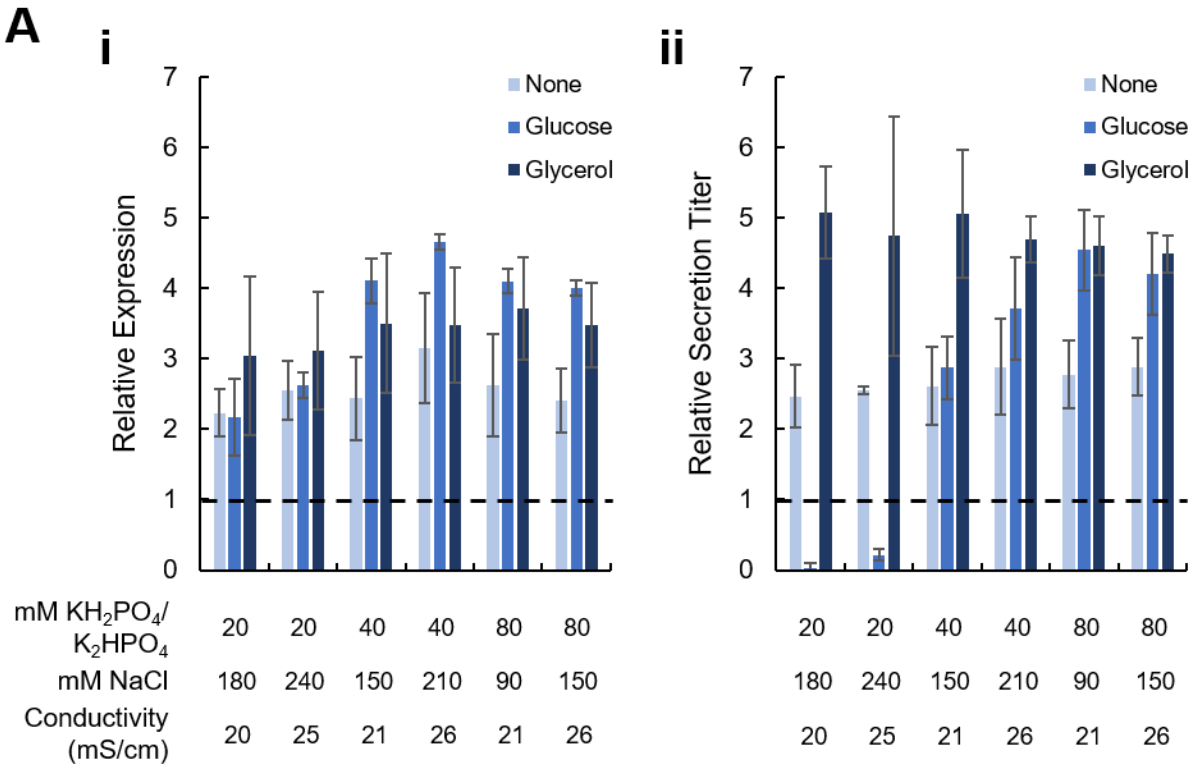


Figure 2.8 Expression and secretion titer are not affected by PhoB.
A Relative bulk secretion titer (*i*) and expression (*ii*) of SptP-DH-2xFLAG-6xHis in ASTE13 WT and Δ*phoB* strains in LB-(KH_2PO_4/K_2HPO_4) with no added carbon source, 0.4% w/v glucose, or 0.4% w/v glycerol. Expression and secretion titer were measured via semi-quantitative western blotting and normalized to LB-L with no additives. Error bars represent one standard deviation. **B** Western blots are representative of three biological replicates.

2.3.7 An ideal medium formulation includes glycerol, phosphate, and sodium chloride

Of the media tested in this work, secretion titer was highest in TB and LB-L with 0.4% w/v glycerol and 89 mM KH_2PO_4/K_2HPO_4. The conductivity of LB-L supplemented with glycerol and 89 mM KH_2PO_4/K_2HPO_4 was 21 mS/cm, which is higher than the 14 mS/cm measured for the LB-(KH_2PO_4/K_2HPO_4) medium listed in Table 2.5. The results of Figure 2.3 suggest that the increased conductivity could be responsible for the higher secretion titer. To evaluate the impact of the ratio of potassium phosphate and sodium chloride, we chose two conductivity levels and added the appropriate amount of sodium chloride to 20, 40, or 80 mM potassium phosphate in a base of 10 g/L tryptone and 5 g/L yeast extract to achieve the specified conductivity. We compared expression and secretion of SptP-DH-2xFLAG-6xHis in those six media with no added carbon source, 0.4% w/v glucose, and 0.4% w/v glycerol (Figure 2.9).

Expression was highest in media with 40 mM or 80 mM KH_2PO_4/K_2HPO_4 and glucose. Secretion titer was consistent across all media with no added carbon source. Secretion titer in media with glucose varied—secretion titer was negligible in media with 20 mM KH_2PO_4/K_2HPO_4, equal to no added carbon source in media with 40 mM KH_2PO_4/K_2HPO_4, and increased fivefold in combination with 80 mM KH_2PO_4/K_2HPO_4. Glycerol provided a fivefold increase in all media, though variability was higher in media with lower concentrations of potassium phosphate. The correlation between the pH of the secreted fraction and secretion titer matched the results of Figure 2.2, Figure 2.4, and Table 2.6—when the pH of the secreted fraction dropped below 6.0, secretion titer decreased dramatically.



B

	[KH_2PO_4/ K_2HPO_4] (mM)	[NaCl] (mM)	pH of Secreted Fraction		
			None	Glucose	Glycerol
A	20	180	8.1±0.0	5.0±0.0	6.4±0.1
B	20	240	8.1±0.1	4.9±0.1	6.5±0.1
C	40	150	8.0±0.0	6.1±0.1	6.7±0.1
D	40	210	8.0±0.1	6.0±0.1	6.8±0.1
E	80	90	7.8±0.0	6.6±0.0	7.0±0.1
F	80	150	7.8±0.0	6.6±0.1	7.0±0.0

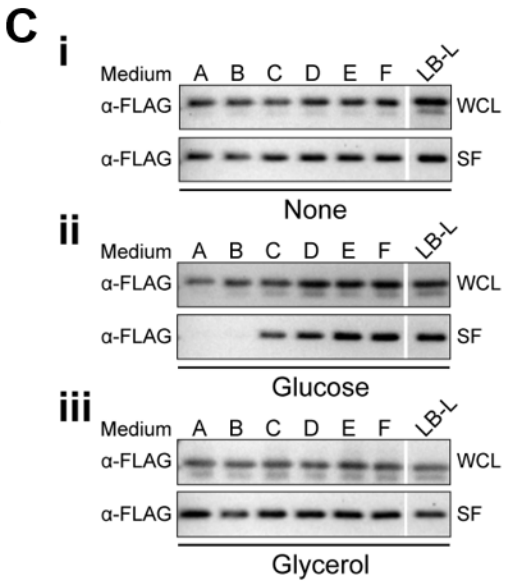


Figure 2.9 An ideal medium formulation includes glycerol, phosphate, and sodium chloride.
A Relative bulk expression (*i*) and secretion titer (*ii*) of SptP-DH-2xFLAG-6xHis media with 10 g/L tryptone, 5 g/L yeast extract, and the listed additives at pH 7.4. Glucose and glycerol were 0.4% w/v. Expression and secretion titer were normalized to LB-L with no additives using semi-quantitative western blotting. Error bars represent one standard deviation. **B** pH of the secreted fractions after harvest. **C** Western blots are representative of three biological replicates. Samples were diluted to fall within the linear range of the LB-L signal. Boxed bands are from the same blot but were rearranged for clarity.

2.3.8 Increases in secretion titer from optimized growth media and strain improvements are additive and general for diverse secreted proteins

We previously showed that knocking out a protein in the SPI-1 T3SS tip complex, SipD, increased secretion titer twofold in strains with *hilA* overexpression [52]. An ideal T3SS production platform would combine all features that increase secretion titer, so we evaluated if Δ*sipD* was additive with an optimized medium. To determine if the effect was general, we selected a variety of test proteins in addition to DH: magainin-1 (MAG1), an antimicrobial peptide; 14B7*, an scFv against the protective antigen of the anthrax toxin [159,160]; and recombinant human growth hormone (rhGH, mature somatropin). All proteins were cloned in the format SptP-POI-2xFLAG-6xHis and secreted in ASTE13 WT and Δ*sipD* strains in Medium "E" from Figure 2.9B (LB-ES for "enhanced secretion").

The Δ*sipD* strain improvement and the optimized medium were indeed additive for all proteins tested (Figure 2.10). Secretion titer increased by varying amounts, but the minimum increase provided by the combination of Δ*sipD* and LB-ES was six-fold above a WT strain in LB-L. Total protein expression showed a different pattern from secretion titer. SptP-MAG1-2xFLAG-6xHis and SptP-DH-2xFLAG-6xHis followed similar expression patterns, and the apparent effect of Δ*sipD* and LB-ES was to increase secretion efficiency for these proteins. SptP-rhGH-2xFLAG-6xHis and SptP-14B7*-2xFLAG-6xHis increased by a surprising eight- and fourteen-fold, however. SptP-rhGH-2xFLAG-6xHis expression increased by the same fraction as secretion titer in each condition, suggesting that secretion titer increased with expression while secretion efficiency was unchanged. SptP-14B7*-2xFLAG-6xHis secretion titer increased by a much smaller margin than expression in both WT and Δ*sipD* strains with LB-ES. We hypothesize that the discrepancy was caused by loss of expressed protein to insoluble aggregates, preventing secretion.

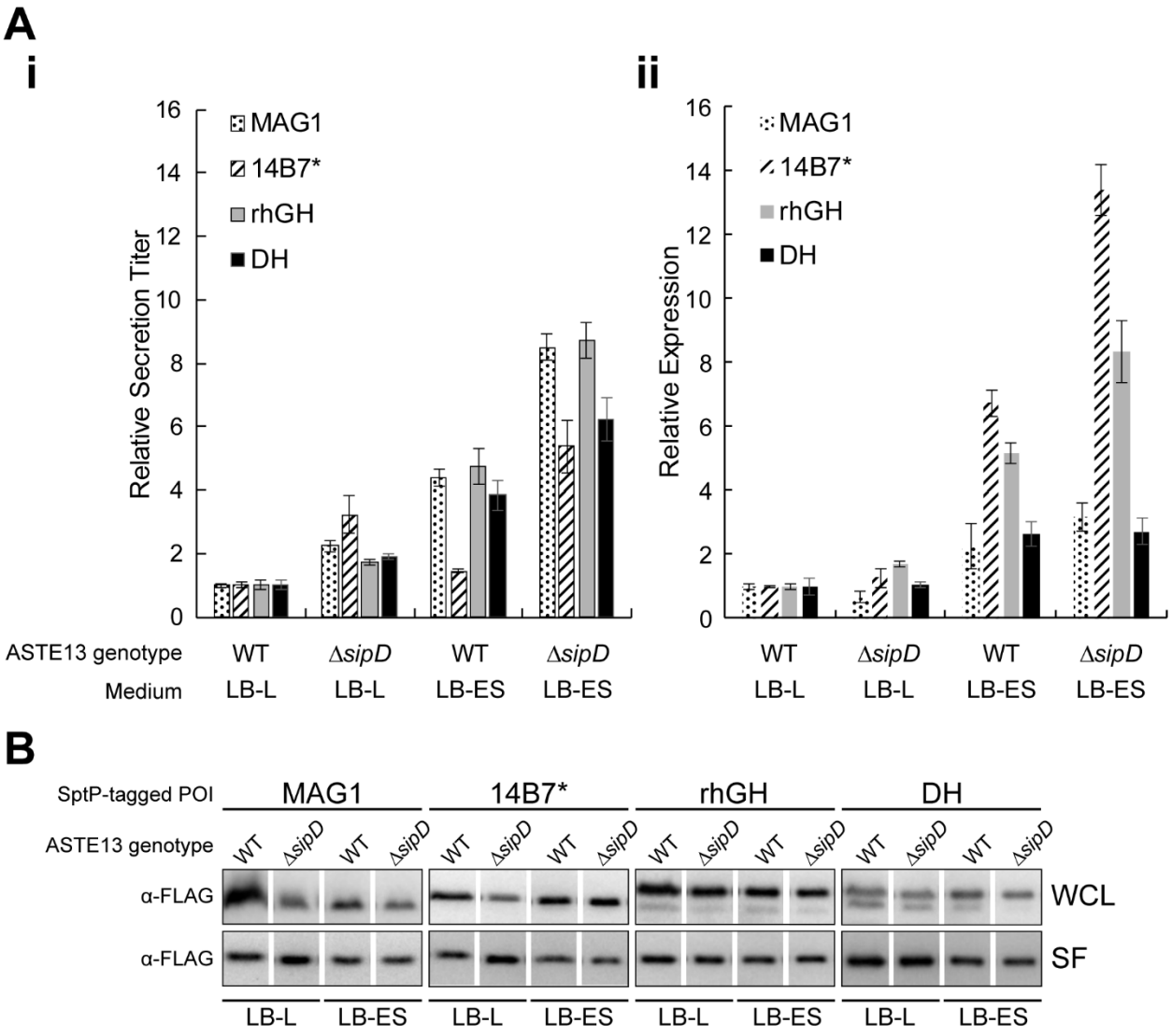


Figure 2.10 Increases in secretion titer from optimized media and strain modifications are additive. **A** Relative bulk secretion titer (*i*) and expression (*ii*) of test proteins from WT and Δ*sipD* strains in LB-L or LB-ES media. "LB-ES" is 10 g/L tryptone, 5 g/L yeast extract, 80 mM KH_2PO_4/K_2HPO_4 pH 7.4, 90 mM NaCl, and 0.4% w/v glycerol. Bulk relative expression and secretion titer were normalized to secretion from ASTE13 WT in LB-L for each protein using semi-quantitative western blotting. Error bars represent one standard deviation. **B** Western blots are representative of three biological replicates. Samples were diluted to fall within the linear range of the normalization signal. Boxed bands are from the same blot but were rearranged for clarity.

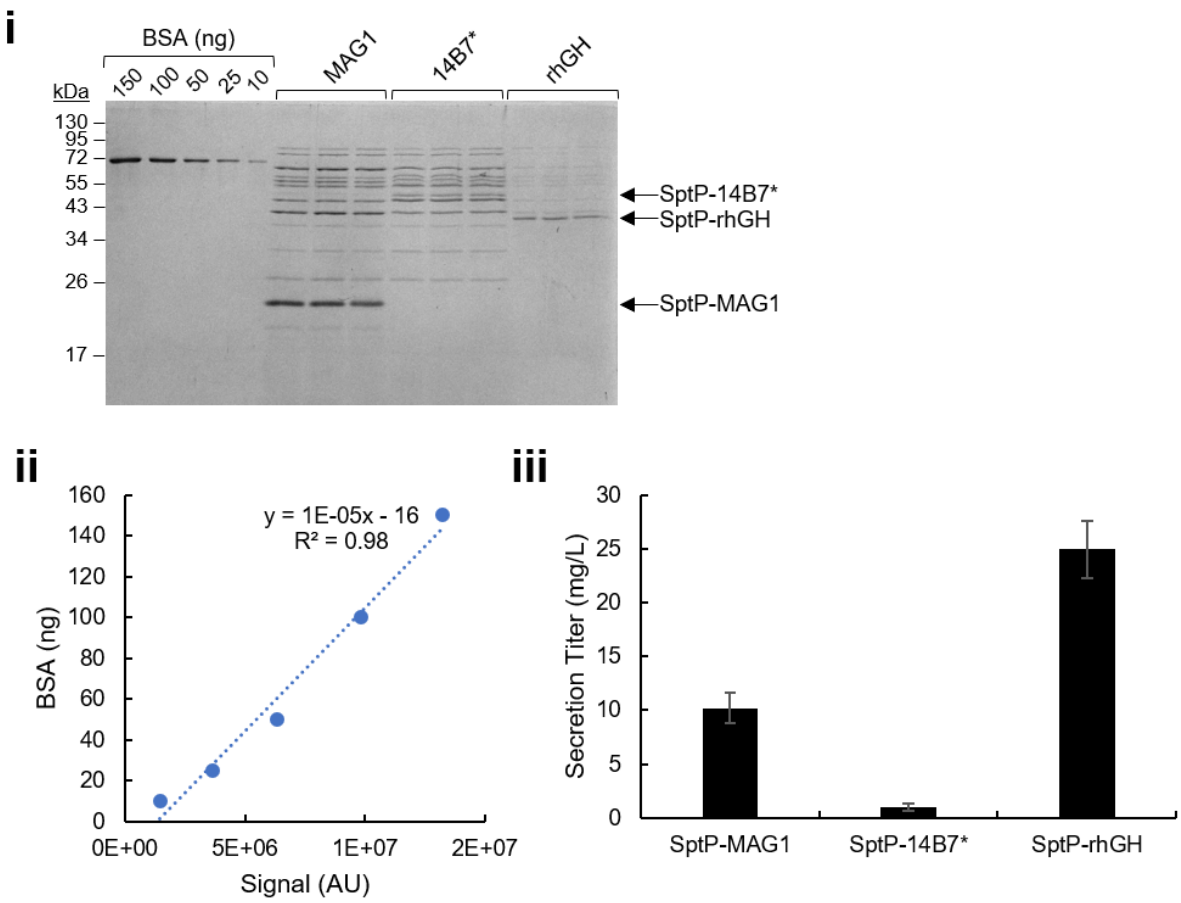


Figure 2.11 Titer of secreted proteins.
MAG1, 14B7*, and rhGH were secreted from an ASTE13 Δ*sipD* strain in LB-ES according to Methods. All proteins were in the format SptP-POI-2xFLAG-6xHis. Secretion titer (*iii*) was measured by performing densitometry relative to a BSA standard curve (*ii*) on a Coomassie-stained gel (*i*). The rhGH secreted fractions were diluted fivefold relative to those of MAG1 and 14B7*. Error bars represent one standard deviation.

2.4 Discussion

Heterologous protein secretion in bacteria retains the robust, low-cost potential of bacterial protein production while significantly decreasing the complexity of the product purification process. The secreted fraction contains many fewer biomolecules than lysate, which allows recovery of the purified product in a single step [27]. Here we expanded our efforts to engineer the *S. enterica* SPI-1 T3SS as a heterologous protein production platform by identifying and optimizing medium components that promoted high secretion titers. The combination of high ionic content, buffering capacity, and a glycolytic carbon source was critical for increased secretion titer. A mixture of phosphate buffer, sodium chloride, and glycerol fulfilled this requirement and provided titers of 10 ± 1 mg/L for the antimicrobial peptide magainin 1, 1 ± 0.3 mg/L for the scFv 14B7*, and 25 ± 3 mg/L for human growth hormone (Figure 2.11).

Transcriptional activity at the *inv*, *prg*, and *sic* loci trended with expression of the secreted protein but was less explanatory for secretion titer, especially in low-pH

conditions. Acidification of the medium to a pH below 6.0 appeared to cause an early arrest of transcriptional activity at all T3SS loci, but it had an even stronger negative effect on secretion titer. Secretion via the T3SS requires the proton motive force (PMF) [138], so it is possible that acidification of the extracellular environment deactivated the secretion apparatus.

Increased secretion titer correlated with increased expression, though secretion efficiency varied with medium composition. The data in Figure 2.10 implies that that there is an upper limit for the benefit of increased expression, however, so secretion efficiency is a more desirable optimization target. Maximizing secretion efficiency—finding the optimum between expression and secretion titer that minimizes product loss to insoluble aggregates and maximizes the fraction of expressed protein that is secreted—will optimize use of cellular resources. The T3SS is a ~4 MDa structure, so optimizing use of cellular resources is critical to create an efficient cellular reactor with selectivity for T3SS assembly, heterologous protein expression, and T3SS operation for heterologous protein secretion.

Though expression and secretion titer increased with medium conductivity, a proxy for ionic content, conductivity was insufficient to fully explain the effect of medium composition. TB and LB-(KH_2PO_4/K_2HPO_4) had the same conductivity and the same concentrations of glycerol and potassium phosphate, but TB had higher expression and secretion titer (Figure 2.12). Likewise, LB-L and 2X YT had the same conductivity, yet expression and secretion titer differed by threefold. Specific components present in tryptone or yeast extract could be responsible, but it is also important to consider osmolarity. Non-ionic carbon sources, tryptone, and yeast extract contribute primarily to osmolarity and minimally to conductivity. Direct measurement of osmolarity requires an osmometer, which is quite expensive, and we did not have access to one. We can assume, however, that 2X YT and TB have a higher osmolarity than LB-L given their increased concentrations of tryptone and yeast extract (Figure 2.1). Tryptone and yeast extract are complex mixtures, and it is difficult to isolate the effects of their components. A chemically defined medium is necessary to specifically evaluate the effects of the chemical identity, concentration, and contribution to osmolarity of the components of tryptone and yeast extract on expression and secretion titer.

Medium design is a critical aspect of any protein production process, but it is especially important for heterologous protein production via the SPI-1 T3SS because of the sensitivity of the T3SS to the extracellular environment. The regulation that governs that sensitivity is multilayered and interwoven [98–100], but here we identified a simple combination of glycolytic carbon source, buffer, and high osmolarity that increased secretion titer at least fourfold for a variety of secreted proteins. That combination can form the base of a fully chemically defined medium, which will expand on this work by allowing precise investigation of the effect of medium composition on secretion titer. A chemically defined medium will also facilitate precise investigation of other environmental effects on secretion titer, such as dissolved oxygen concentration and controlling growth via fed-batch or continuous culture strategies. Taking advantage of native T3SS regulation to define an extracellular environment that promotes robust, high secretion titers will identify new regulatory mechanisms and advance the development of the T3SS as a heterologous protein production platform.

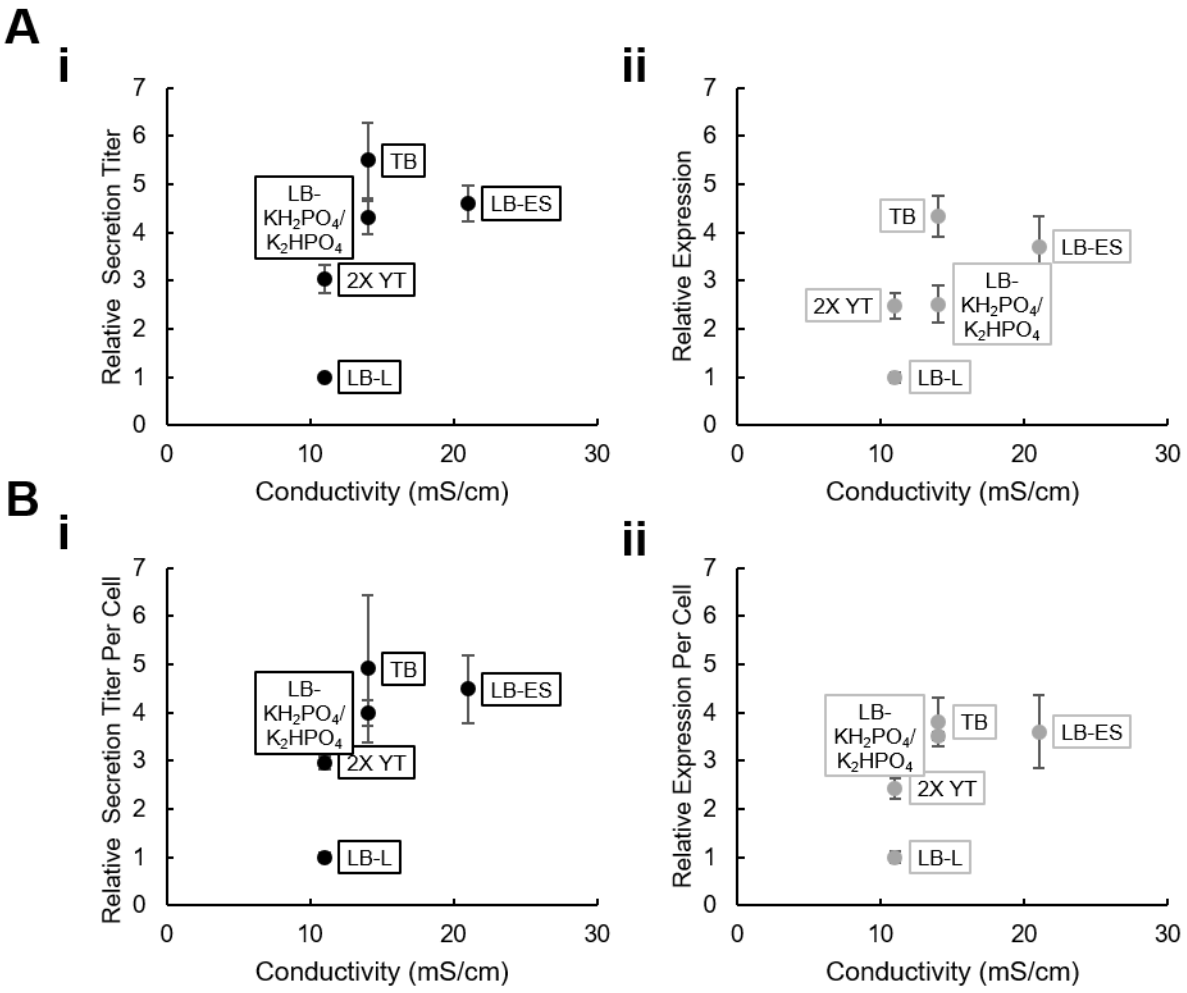


Figure 2.12 Undefined media components affects expression and secretion titer through mechanisms other than conductivity.
A Relative bulk secretion titer (*i*) and expression (*ii*) and **B** relative secretion titer (*i*) and expression (*ii*) per cell versus conductivity for secretion of SptP-DH-2xFLAG-6xHis from ASTE13 WT in media from Figure 2.1, Figure 2.4, and Figure 2.9.

Chapter 3: A Defined Medium for Increased Protein Secretion Titer

3.1 Introduction

The previous chapter outlined the importance of the physical and chemical environment for secretion system function and identified a combination of glycerol, phosphate, and sodium chloride at high osmolarity as critical for increased secretion titer. Those experiments were conducted with a base of undefined medium components, however, so interactions between the critical ingredients and nutrients in the undefined portion of the medium formulation were impossible to discern. Explicitly defining the chemical composition of the medium is necessary to precisely investigate environmental effects on secretion titer.

Bacterial growth media can be sorted into three categories: undefined media, in which the chemical composition is unknown except for one or two ingredients; semi-defined media, in which the chemical composition is known with the exception of one or two ingredients, and defined media, in which the chemical composition is known explicitly. Undefined and semi-defined media often produce more biomass than defined media, but it is difficult to isolate factors that influence growth, cellular physiology, or other experimental outcomes like protein expression because the chemical composition of those media is unknown. Chemically defined media alleviate this challenge and provide the additional benefit of reducing batch-to-batch variability.

In this chapter, I describe initial optimization of defined growth media for consistent and high titers of heterologous protein via the *S. enterica* T3SS. I evaluated three minimal growth media, NCE, M9, and PCN, and compared secretion titer to that in LB-L. I tested the effect of adding difference carbon sources to NCE, and I modified the formulas of NCE and PCN using the information learned in Chapter 2 to maximize secretion titer.

3.2 Methods

3.2.1 Strains and Growth Conditions

Strains and plasmids used are listed in Table 3.1 and Table 3.2. Defined media formulations were as listed in Appendix B. Carbon sources were added as specified. Secretion experiments were started by growing a single colony from a fresh streak of a frozen glycerol stock in the lysogeny broth Lennox formulation (LB-L: 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) with appropriate antibiotics (34 µg/mL chloramphenicol, 50 µg/mL kanamycin) for 12-16 hours overnight in an orbital shaker at 37°C and 225 rpm. For secretion experiments, overnight cultures were diluted 1:100 into the appropriate medium. All media were supplemented with 100 µg/mL isopropyl β-D-1-thiogalactopyranoside (IPTG) and appropriate antibiotics. All culturing steps were performed in 24-well deepwell plates (Axygen). Secretion was performed for the time specified at 37°C and 225 rpm in an orbital shaker. Whole culture lysate samples for total protein expression were prepared for SDS-PAGE by adding cell suspension to Laemmli buffer [150] in a 1:2 ratio prior to centrifugation. The secretion fraction was harvested by centrifuging cultures at 4000 x *g* for 10 minutes. SDS-PAGE samples for the secreted fraction were prepared by adding supernatant to Laemmli buffer in a 3:1 ratio [150]. All SDS-PAGE samples were boiled at 95°C for 5 minutes immediately after preparation.

For growth curves, 1 mL of overnight culture was pelleted at 5000 *x g* for 5 minutes. The supernatant was discarded, and cells were resuspended in 1 mL of the medium used for the growth curve. The resuspended cells were diluted to the specified OD in 5 mL of the appropriate medium supplemented with appropriate antibiotics. All culturing steps were performed in 24-well deepwell plates (Axygen). 100 µL samples were taken hourly for OD_{600} measurement and diluted in H_2O to fall within an OD_{600} range of 0.2-0.8. Growth rates were calculated by identifying the linear region (with a minimum of three time points) and calculating the slope via linear regression. Doubling time was calculated with $t_D = \log(2)\ /slope$.

Table 3.1 Strains used in Chapter 3.

Strain Name	Comment	Reference
ASTE13	LT2-derived lab strain similar to DW01	This study; DW01 [65]
ASTE13 Δ*prgI*	*prgI* knockout	[41]

Table 3.2 Plasmids used in Chapter 3.

Plasmid Name	ORFs under inducible control		ORI	ab^R	Reference
P_{sic} *bla*	*sicP*	*sptP-bla-2xFLAG-6xHis*	colE1	cam	[41]
P_{sic} *DH-GFP11*	*sicP*	*sptP-DH-GFP11-2xFLAG-6xHis*	colE1	cam	This study
P_{sic} *bla-GFP11*	*sicP*	*sptP-bla-GFP11-2xFLAG-6xHis*	colE1	cam	This study
P_{lacUV5} *hilA*	*hilA*		p15a	kan	[41]

Table 3.3 Primers used in Chapter 3.

Sequence	Amplicon	Used to Construct
A GGTCTC A **GCTT** gatatgttgaccccaactgaaag	*DH-GFP11*	*sptP-DH-GFP11-2xFLAG-6xHis*
A GGTCTC C **CGCT** *GGTGATGCCCGCCGCGTTCACGTATTCATGC AGCACCATATGATCCCGCGAACCACCCCCAG ATCCACCCCC* agagttctccttctcccgc	*DH-GFP11*	*sptP-DH-GFP11-2xFLAG-6xHis*
A GGTCTC A **GCTT** cacccagaaacgctggtga	*bla-GFP11*	*sptP-bla-GFP11-2xFLAG-6xHis*
A GGTCTC C **CGCT** *GGTGATGCCCGCCGCGTTCACGTATTCATGC AGCACCATATGATCCCGCGAACCACCCCCAG ATCCACCCCC* ccaatgcttaatcagtgagg	*bla-GFP11*	*sptP-bla-GFP11-2xFLAG-6xHis*

DNA Manipulations

Table 3.3PCR was performed with Phusion DNA polymerase using the primers listed in Table 3.3. Golden gate cloning was used to construct plasmids [151]. Genes for proteins to be secreted were inserted into a modified pPROTet.133 backbone vector (BD Clontech) under the control of the *sic* promoter [8]. All secretion plasmids expressed the SptP chaperone *sicP* and the *sptP* secretion signal (nucleotides 1 to 477). The SptP secretion tag was fused N-terminal to the protein of interest, and 2xFLAG and 6xHis tags

were fused C-terminal to the protein of interest. The GFP11 tag was added C-terminal to the protein of interest using 3' extensions of the primers to append the tag 3' to the gene of interest. All cloning was done in *E. coli* DH10B cells, and all DNA sequences were confirmed by Sanger sequencing (Quintara).

3.2.2 Protein Separation and Western Blotting

Samples were separated by SDS-PAGE and stained with Coomassie R-250 according to Studier [153] or transferred to a polyvinylidene fluoride membrane (PVDF, Millipore) for western blotting using the Mini-Transblot cell (Bio-Rad). For western blotting, membranes were probed with mouse anti-FLAG per manufacturer's instructions (Sigma Aldrich). To facilitate chemiluminescent detection, a secondary labeling step was performed with goat anti-mouse IgG (H+L) HRP conjugate according to manufacturer's instructions (Thermo Fisher). Bands were detected with the SuperSignal West Pico or Pico Plus substrate (Thermo Fisher) and a ChemiDoc XRS + imaging system (Bio-Rad). All relative protein quantities were calculated by performing densitometry using Image Lab software (Bio-Rad) and normalizing to the specified normalization condition. Error bars are standard deviation on three biological replicates unless otherwise specified.

3.2.3 Protein Purification

All proteins were purified from *S. enterica* Typhimurium ASTE13 Δ*prgI*. Cells were grown and induced as described in "Strains and Growth Conditions" except that the growth vessel was a 250 mL Erlenmeyer flask containing 50 mL of terrific broth (TB). Cells were harvested by pelleting at 5000 x *g* for 10 minutes. The supernatant was discarded, and the cell pellets were frozen at -80°C. Cell pellets were thawed and resuspended in 10 mL 20 mM sodium phosphate pH 7.4 with 500 mM NaCl and 20 mM imidazole. The resuspended cell solution was homogenized via sonication. The soluble and insoluble fractions were separated by centrifugation at 17,000 x *g* for 20 minutes at 4°C. The soluble fraction was filtered with a 0.2 µm syringe filter before being applied to a His GraviTrap column (GE Life Sciences) for purification. The eluted protein was buffer-exchanged into phosphate-buffered saline using a PD-10 desalting column and quantified compared to a BSA standard using densitometry on an SDS-PAGE gel stained with Coomassie R-250 [153].

3.2.4 Fragmentation Analysis and Zymograms

Proteins were purified as described. Secretion experiments were performed as described with ASTE13 WT lacking plasmids, ASTE13 WT expressing P_{lacUV5} *hilA*, and ASTE13 Δ*prgI* expressing both P_{lacUV5} *hilA* and the secretion vector containing the protein of interest. The secreted fraction was collected as specified in "Strains and Growth Conditions" and concentrated fivefold using Vivaspin 500 Centrifugal Concentrators with a 10 kDa cutoff (Sartorius). Equal volumes of the concentrated supernatant and the purified protein were mixed, and PBS was used as a control. The mixtures were held at 37°C for 22 or 24 hours and sampled as specified. Samples were prepared for regular SDS-PAGE by mixing the protein sample with Laemmli buffer and boiled at 95°C for 5 minutes.

Samples for the zymogram gel were prepared by mixing the concentrated supernatants by with Laemmli buffer lacking a reducing agent. 20 µL of the prepared sample was loaded on a Novex 10% Zymogram Plus Gel containing gelatin (Invitrogen

#ZY00100BOX). Electrophoresis and zymogram development were performed according to manufacturer's instructions. Briefly, following electrophoresis, the gel was placed in 50 mL Renaturing Buffer (Invitrogen) and rocked gently at room temperature for 30 minutes. The Renaturing Buffer was replaced with 1X Developing Buffer (Invitrogen) and rocked gently at room temperature for 30 minutes. The gel was placed in a fresh 50 mL of 1X Developing Buffer, the box was sealed, and the gel was placed in an orbital shaker set at 25 rpm and 37°C for overnight incubation. The gel was Coomassie-stained according to Studier [153] and imaged immediately to prevent significant destaining.

3.3 Results

3.3.1 An amino acid supplement in minimal media provides similar growth characteristics as LB-L

Minimal media formulations are generally undesirable for heterologous protein production because they provide, by definition, minimal nutrients, and protein production is a resource-intensive process [161,162]. This is especially true for protein secretion via the T3SS, as the ~4 MDa apparatus itself is composed of hundreds of proteins. A frequent method of relieving the biosynthetic burden imposed by growth in minimal media is to supplement with casamino acids, a tryptic digest of casein. This creates a semi-defined medium, however, which is undesirable for our studies. Fortunately, the commercially available EZ-Rich Defined Medium Kit (Teknova), based on PCN medium, includes a defined amino acid supplement (5X EZ, Appendix B).

Secretion via the T3SS in minimal media requires optimization beyond medium composition because T3SS induction and expression are growth-phase dependent. The ideal T3SS induction time is early exponential phase [41], which occurs at the time of subculture in rich media because there is no lag phase. Also, as observed in Chapter 2, transcriptional activity at T3SS loci declines significantly by early stationary phase in unoptimized LB media [41,163]. The biosynthetic burden imposed by minimal media introduces a lag phase in addition to decreasing growth rate relative to undefined media. Thus, it was important to determine the growth characteristics of each minimal medium to find the appropriate induction and harvest times. I selected NCE, M9, and PCN (Appendix B) as test media because NCE and PCN were specifically optimized for growth of enterobacteria [157,164], and M9 is the most commonly used minimal medium for bacterial growth. Iron is known to upregulate SPI-1 activity, so I added 50 µM ferric citrate to M9 as it lacks a direct iron source. Glycerol and succinate were compared as carbon sources in each medium.

Glycerol provided higher final cell densities and faster growth rates than succinate in all media except M9 and LB-L (Figure 3.1, Table 3.4). Glycerol also acidified the extracellular environment in PCN by 24 hours, which might cause decreased secretion titer according to the results of Chapter 2. Ferric citrate did not change the growth characteristics in M9 media. The final cell density in M9 was lower than expected, but inspection of the medium recipe used for the growth curve revealed that it included a tenfold lower ammonium chloride concentration than the recipe specified in Supplementary Table 3.1. Thus, cell cultures in M9 in this experiment were likely nitrogen-limited, which explains the low final OD. As expected, all minimal media had a slower growth rate and lower final cell densities than LB-L (Table 3.4). If these media were used

in these formulations (or with the correct ammonium chloride concentration in M9), induction of P_{lacUV5} *hilA* should occur two hours after subculture.

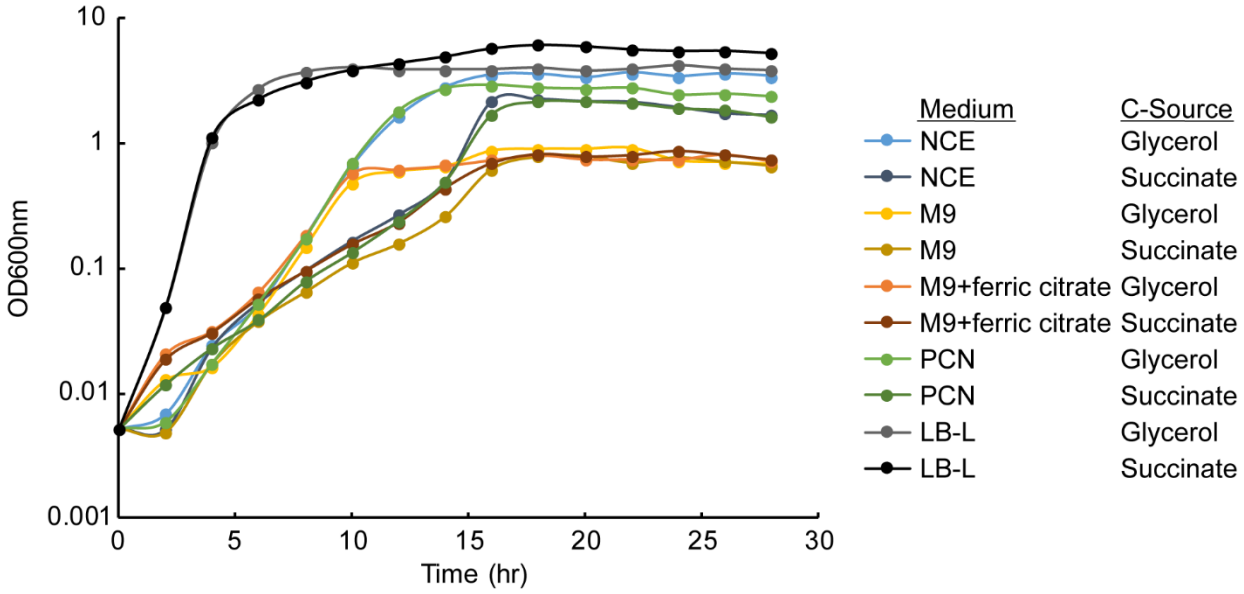


Figure 3.1 Growth curves in minimal media and LB-L with 0.4% glycerol or 0.4% succinate. Overnight cultures of ASTE13 containing P_{lacUV5} *hilA* and P_{sic} *sicP sptP-bla-2xFLAG-6xHis* were diluted to an OD_{600} of 0.5 in LB-L and subcultured 1:100 into the specified medium to achieve a starting OD_{600} of 0.005. Each media contained either 0.4% w/v glycerol or 0.4% w/v succinate as specified. OD_{600} was measured every two hours in a cuvette using a NanoDrop 2000 spectrophotometer. The data represent one biological replicate.

Table 3.4 Growth rates, end OD, and pH in defined media and LB-L.

Medium	Carbon Source	μ (hr^{-1})	t_D (hr)	Exponential Phase	OD at 24 hr	pH at 12 hr	pH at 24 hr
NCE	Glycerol	0.56	1.2	4hr-12hr	4.2	7.5	7.0
NCE	Succinate	0.39	1.8	4hr-14hr	2.3	7.5	8.0
M9	Glycerol	0.46	1.5	4hr-10hr	0.85	7.5	7.5
M9	Succinate	0.36	1.9	4hr-12hr	0.90	7.5	8.5
M9+ferric citrate	Glycerol	0.43	1.6	4hr-10hr	0.97	7.5	7.5
M9+ferric citrate	Succinate	0.27	2.5	4hr-14hr	1.0	7.5	8.5
PCN	Glycerol	0.61	1.2	4hr-10hr	3.2	7.0	5.0
PCN	Succinate	0.32	2.2	4hr-14hr	2.2	7.5	9.0
LB-L	Glycerol	1.3	0.5	1hr-4hr	4.9	5.5	5.0
LB-L	Succinate	1.35	0.5	1hr-4hr	6.5	8.0	9.0

I repeated the growth curves and supplemented each minimal medium with 1X EZ Supplement (Teknova) to provide amino acids and biosynthetic precursors. Sampling was hourly until 12 hours. Figure 3.2 shows the new growth curves overlaid with LB-L from the previous experiment. The 1X EZ Supplement produced growth rates and final cell densities equivalent to those in LB-L for all defined media (Table 3.5).

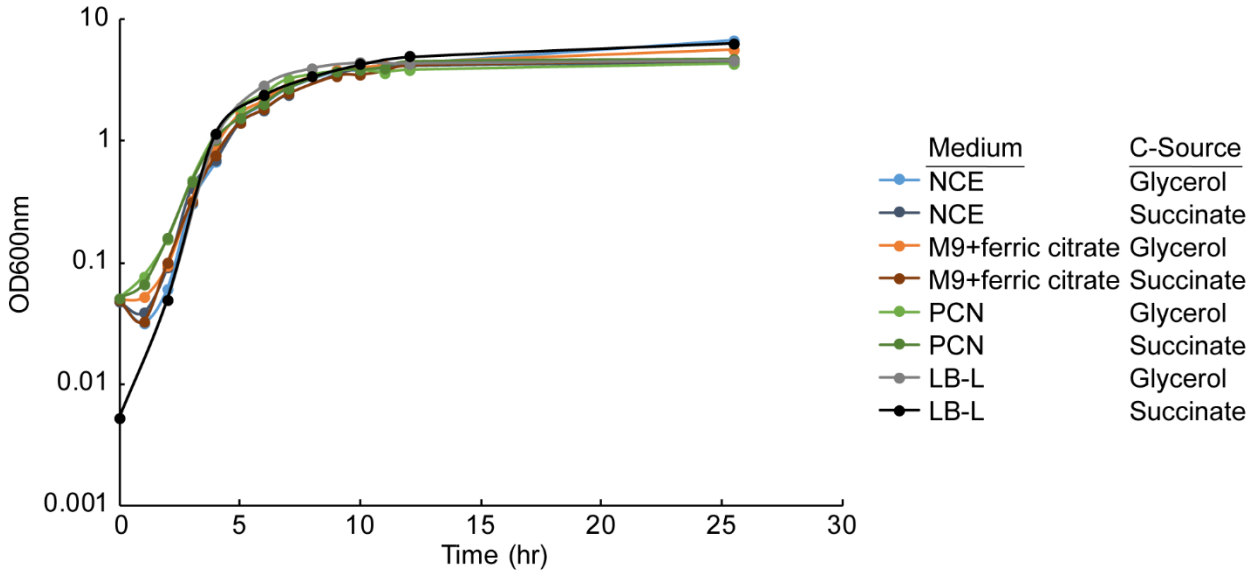


Figure 3.2 Growth curves for minimal media supplemented with 1X EZ Supplement and LB-L. Overnight cultures of ASTE13 containing P_{lacUV5} *hilA* and P_{sic} *sicP sptP-bla-2xFLAG-6xHis* were concentrated, resuspended in the media for the growth curve, and seeded at an OD_{600} of 0.05. Each media contained either 0.4% w/v glycerol or 0.4% w/v succinate as specified. OD_{600} was measured in a cuvette using a NanoDrop 2000 spectrophotometer. Sampling was hourly until 12 hours, and a final optical density was measured at 26 hours. The data represent a single biological replicate.

Table 3.5 Growth rates in defined media with 1X EZ Supplement compared to LB-L.

Medium	Carbon Source	μ (hr^{-1})	t_D (hr)	Exponential Phase	Max OD
NCE	Glycerol	0.96	0.72	1hr-5hr	6.8
NCE	Succinate	0.90	0.77	1hr-5hr	4.5
M9+ferric citrate	Glycerol	0.87	0.80	1hr-5hr	5.7
M9+ferric citrate	Succinate	0.94	0.74	1hr-5hr	4.6
PCN	Glycerol	0.80	0.86	1hr-5hr	4.3
PCN	Succinate	0.79	0.88	1hr-5hr	4.7
LB-L	Glycerol	1.3	0.52	1hr-4hr	4.9
LB-L	Succinate	1.4	0.51	1hr-4hr	6.5

3.3.2 Supplemented defined media produces secretion titers equivalent to LB-L

Adding the 1X EZ supplement to the defined media allowed secretion experiments to proceed on the same timeline as those in LB-L, so the next step was to compare secretion titer. Previously, secretion experiments ended at eight hours because T3SS transcriptional activity decreased in late exponential phase [41,163]. Chapter 2 showed that transcriptional activity was extended in the presence of buffers and carbon sources, however, so the experiment was extended to 24 hours with sampling at 8 and 12 hours. The data in Figure 3.3 show that secretion continues beyond eight hours in defined media and LB-L with 0.4% w/v succinate.

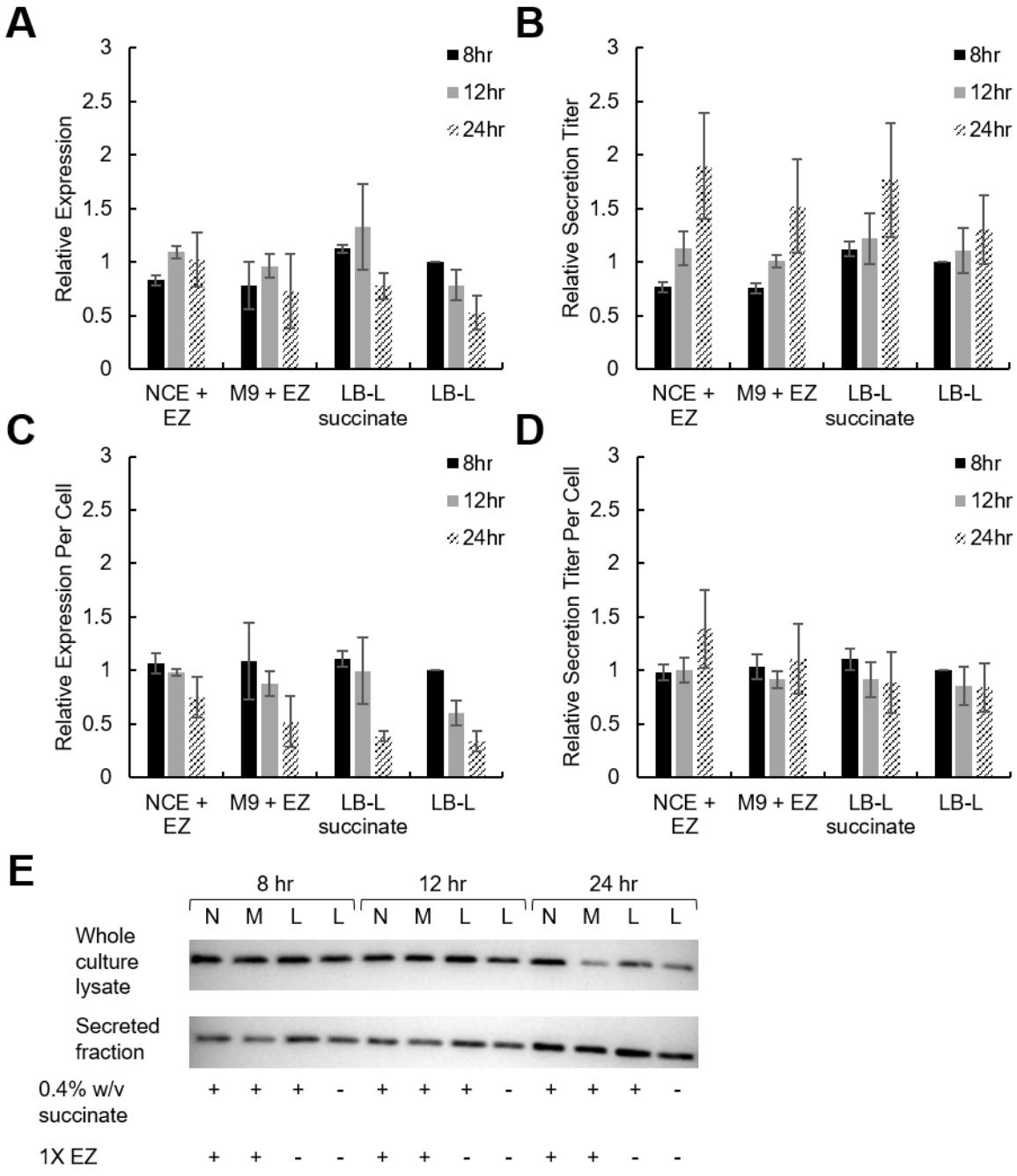


Figure 3.3 Relative expression and secretion titer in defined media.
Relative expression (**A,C**) and secretion titer (**B,D**) of SptP-Bla-2xFLAG-6xHis in NCE and M9 containing 1X EZ Supplement and 0.4% w/v succinate were compared to LB-L with 0.4% w/v succinate and un-supplemented LB-L. Values represent densitometry on semi-quantitative western blots normalized to LB-L at 8 hours. "Per cell" values are divided by OD_{600} at the specified time point. Induction of P_{lacUV5} *hilA* facilitated T3SS and SptP-Bla-2xFLAG-6xHis expression. Representative western blots are displayed in (**E**). "Whole culture lysate" represents total expressed protein. "N", "M", and "L" are NCE, M9, and LB-L.

The decrease in relative expression at 24 hours indicates that expression persisted until at least 12 hours in media that contained succinate but ceased sometime before 24 hours. Within error, all media reached an equivalent maximum secretion titer at 24 hours, but the relative titers at 8 and 12 hours suggest that the average rate of secretion was lower in NCE and M9 compared to LB-L with and without succinate.

3.3.3 Glycerol is an optimal carbon source in NCE

I selected NCE as a base medium to study the effect of a panel of carbon sources on secretion titer. NCE is used for growth and study of *S. enterica* Typhimurium [156,164], and it contains a high concentration of phosphate (Supplementary Table 3.1), which should maximize secretion titer according to the results outlined in Chapter 2.

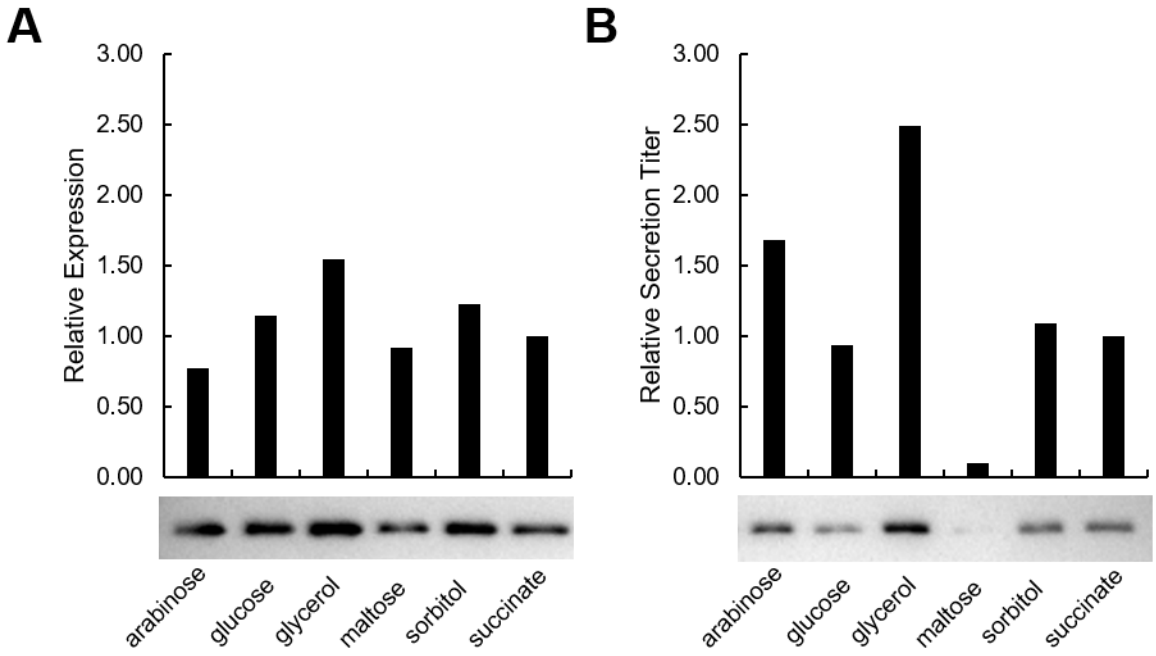


Figure 3.4 Relative expression and secretion titer in NCE with various carbon sources.
Expression (**A**) and secretion titer (**B**) of SptP-Bla-2xFLAG-6xHis at 8 hours were measured via semi-quantitative western blotting and normalized to succinate. 1X EZ Supplement was included in NCE. All carbon sources were 0.4% w/v. P_{lacUV5} *hilA* induction facilitated T3SS and SptP-Bla-2xFLAG-6xHis expression. Data represents one biological replicate.

Table 3.6 Optical density at 8 hours in NCE with various carbon sources.

Carbon Source	8hr OD_{600}
arabinose	3.4
glucose	3.1
glycerol	2.9
maltose	1.7
sorbitol	3.4
succinate	2.4

Glycerol provided the highest secretion titer, and the identity of the carbon source had a stronger impact on secretion titer than expression (Figure 3.4). Maltose was a poor carbon source—it supported a low cell density and negligible secretion titer (Table 3.6). Glucose and succinate provided equivalent bulk secretion titers, though secretion per cell was higher with added succinate according to OD_{600} values at 8 hours. The relative expression and secretion titers with added glucose, glycerol and succinate matched the trends observed in Chapter 2.

3.3.4 A high phosphate concentration is critical for increased secretion titer in defined media

To build on the results of sections 3.3.2 and 3.3.3, I supplemented PCN medium with concentrations of K_2HPO_4 and KH_2PO_4 equivalent to those in NCE to create "PCN+P" medium (Appendix B). I performed secretion for 24 hours, sampling at eight hours, in NCE, PCN, and PCN+P with 0.4% w/v glucose, glycerol, or succinate. The model protein was SptP-Bla-GFP11-2xFLAG-6xHis, as I was optimizing a split GFP assay to measure secretion titer in parallel. I discuss those efforts in Chapter 4.

Expression was similar across all media at 8 and 24 hours, which was expected based on the results of Chapter 2 and the composition of NCE, PCN, and PCN+P (Figure 3.5). PCN with glucose and glycerol showed a slight decrease in relative expression from 8 to 24 hours. The final cell density was lower in these media than other conditions tested (Table 3.7). The end pH in Table 3.4 from the growth curves suggests that the extracellular environment was acidified, though pH was not measured directly. The decrease in relative expression per cell for all media at 24 hours indicated that expression ceased at some point between 8 and 24 hours.

Secretion titer was more variable than expression. The highest secretion titers were recorded in NCE with glycerol and PCN+P with any carbon source at 24 hours (Figure 3.6). PCN with glucose had low secretion titer at eight hours, likely due to acidification of the extracellular environment. Secretion titer in PCN with glycerol was similar to PCN with succinate at eight hours, but secretion at 24 hours was lower than secretion at eight hours, indicating that the secreted protein had degraded.

Analyzing the fold change in secretion from 8 to 24 hours suggests that secretion rate varied among the media tested (Figure 3.6C), as previously observed when comparing secretion in NCE and M9 to LB-L (Figure 3.3). Secretion rate was highest in PCN+P with glucose or glycerol—secretion titer in those media had nearly reached their maximum values at eight hours. NCE with glycerol, PCN+P with succinate, and PCN+P with glucose or glycerol achieved similar secretion titers at 24 hours, but secretion rate was about twofold slower in NCE with glycerol and PCN+P with succinate. NCE with glucose had the lowest secretion rate of media that provided increased secretion titer.

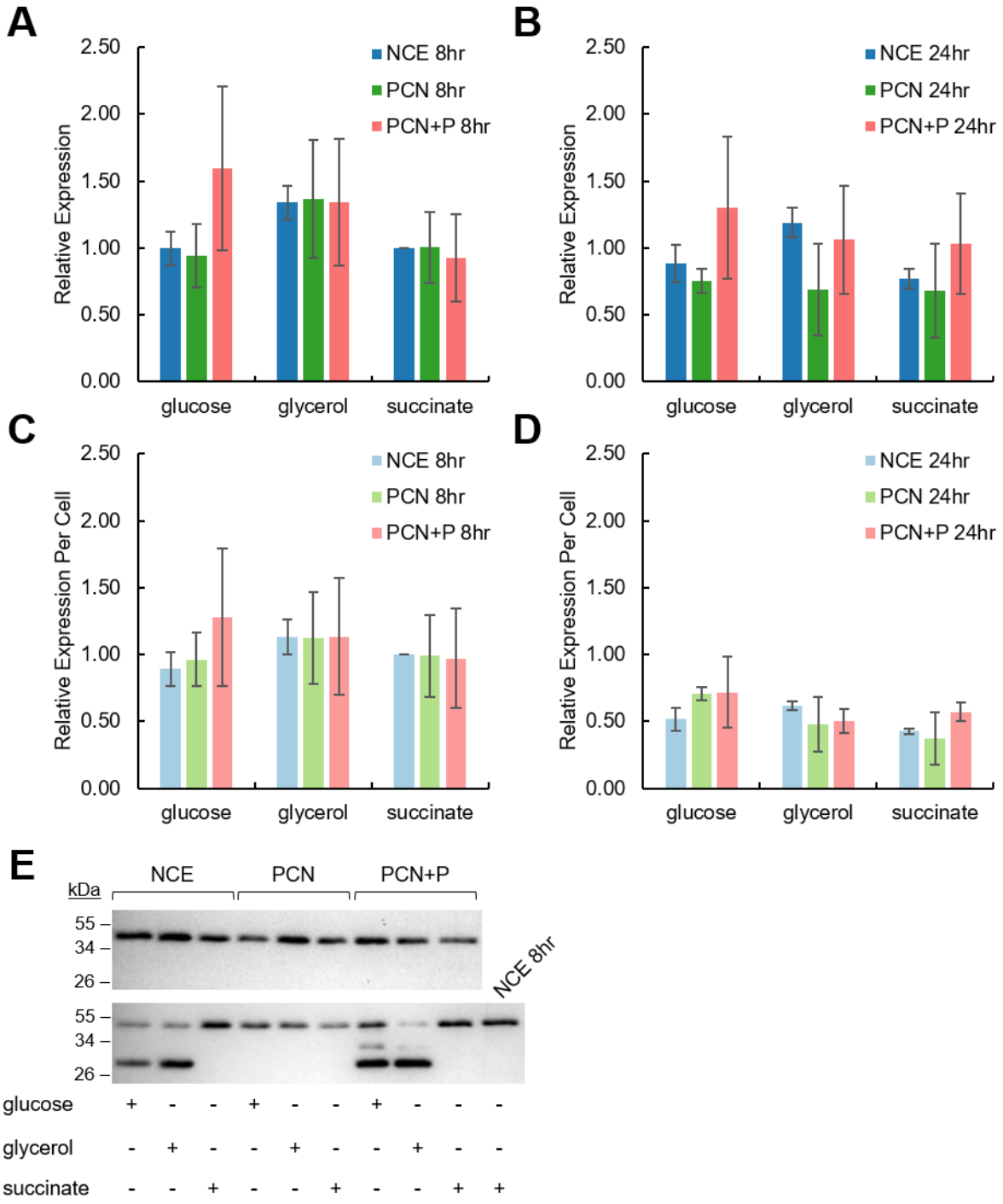


Figure 3.5 Expression of SptP-Bla-GFP11-2xFLAG-6xHis in defined media with glucose, glycerol, or succinate.
Expression of SptP-Bla-2xFLAG-6xHis was measured in NCE, PCN, and PCN+P at 8 (**A,C**) and 24 (**B,D**) hours via semi-quantitative western blotting. Data was normalized to NCE with 0.4% w/v succinate at 8 hours. "Per cell" values are bulk measurements divided by OD_{600} values at the appropriate time point. The T3SS was induced via P_{lacUV5} *hilA* expression. Representative western blots are shown in (**E**).

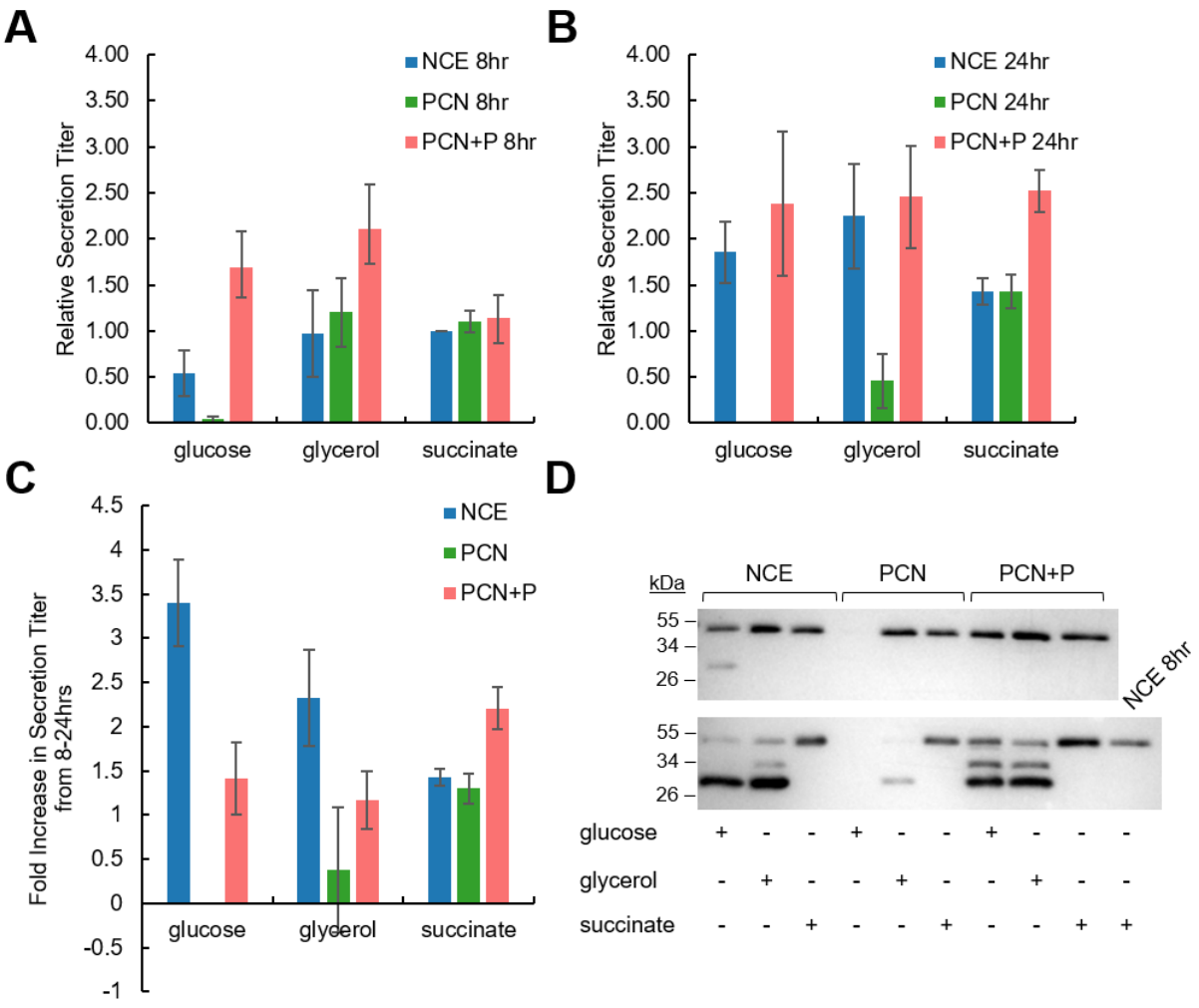


Figure 3.6 Relative secretion titer of SptP-Bla-GFP11-2xFLAG-6xHis in defined media with glucose, glycerol, or succinate.
Secretion titer of SptP-Bla-2xFLAG-6xHis was measured in NCE, PCN, and PCN+P at 8 (**A**) and 24 (**B**) hours via semi-quantitative western blotting. All secretion titers were normalized to NCE with 0.4% w/v succinate at 8 hours. The T3SS was induced via P_{lacUV5} *hilA* expression. Fold change from 8 to 24 hours was calculated by taking the ratio of relative secretion titers at each time point and propagating error appropriately (**C**). Representative western blots are shown in (**D**).

Table 3.7 Optical density at 8 and 24 hours in NCE, PCN, and PCN+P with glucose, glycerol, or succinate.

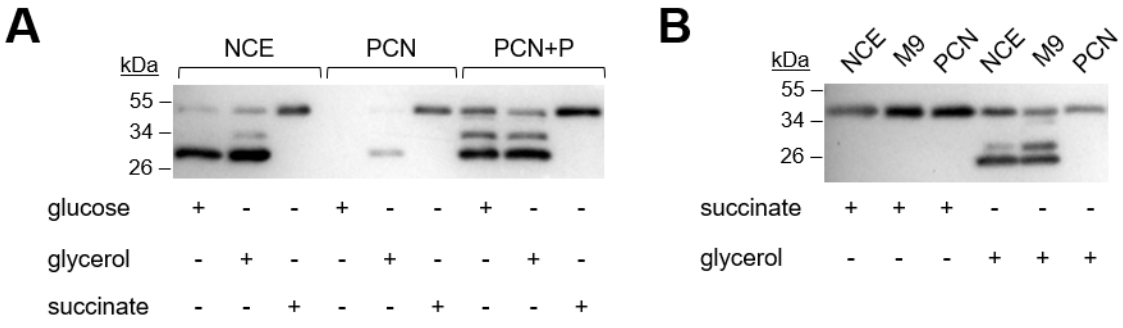
Medium	Carbon Source	OD at 8 hr	OD at 24 hr
NCE	glucose	3.1	4.5
PCN	glucose	2.6	2.9
PCN+P	glucose	3.4	4.8
NCE	glycerol	3.3	5.2
PCN	glycerol	3.5	3.8
PCN+P	glycerol	3.4	5.6
NCE	succinate	2.7	5.0
PCN	succinate	2.7	5.1
PCN+P	succinate	2.6	5.0

3.3.5 Fragments appear in the secreted fraction after eight hours in the presence of glucose and glycerol

When I developed the western blots for the experiments described in the previous section, I was surprised to discover fragmentation in the secreted fractions from media containing glucose and glycerol (the bar graphs above represent total protein quantification). I also noticed similar fragmentation in NCE, M9 and PCN with glycerol at 24 hours with SptP-DH-GFP11-2xFLAG-6xHis as the secreted protein (Figure 3.7). The fragments were different molecular weights for SptP-Bla-GFP11-2xFLAG-6xHis and SptP-DH-GFP11-2xFLAG-6xHis, but they were offset by the same difference in molecular weight between DH and beta-lactamase. This detail and the fact that the fragments appeared on an anti-FLAG western blot suggested that cleavage was N-terminal to the FLAG tags and occurring within the SptP tag. A quick analysis with ExPasY PeptideCutter [165] suggested that cleavage was occurring near amino acids 80 and 125 in the SptP tag, and that a trypsin-like or chymotrypsin-like serine protease was responsible.

I wanted to confirm that cleavage was occurring outside of the cell and identify the conditions under which the protease(s) might appear in the secreted fraction. If the cleavage was a result of proteolysis, the protease(s) could appear in the secreted fraction via several avenues. First, they could appear as a result of normal cellular growth, in which case they would be present in the secreted fraction of ASTE13 WT harboring no plasmids. The protease could also appear as a result of T3SS activation, and then it would be secreted upon overexpression of P_{lacUV5} *hilA*. Finally, the protease could be secreted by the T3SS itself, in which case it would be absent from the secreted fraction in a Δ*prgI* strain.

I performed secretion experiments in the strains listed in Table 3.8. The “LMW/Full-length” column is the ratio of the densitometry values for the low molecular weight (LMW) and full-length bands. The SDS-PAGE gel in Figure 3.8 is labeled according to the lane numbers specified in Table 3.8.



C

	Full-length (kDa)	LMW 1 (kDa)	LMW 2 (kDa)
SptP-BlaGFP11--2xFLAG-6xHis	52	34	31
SptP-DH-GFP11-2xFLAG-6xHis	47	29	26

Figure 3.7 Analysis of cleavage product size after secretion for 24 hours.
SptP-Bla-GFP11-2xFLAG-6xHis (**A**) and SptP-DH-GFP11-2xFLAG-6xHis (**B**) were secreted for 24 hours in the specified media following induction of P_{lacUV5} *hilA*. All carbon sources were 0.4% w/v. Fragment size for each protein was estimated from the molecular weight marker and listed in (**C**). "LMW" is "low molecular weight".

Table 3.8 Strains used to generate the secreted fraction and analysis of cleavage of purified SptP-tagged protein.

Lane	Strain	Plasmid(s)	Carbon Source	LMW/Full	Full/PBS
1	ASTE13 WT	None	Glycerol	0.15	0.87
2	ASTE13 WT	P_{lacUV5} *hilA*	Glycerol	0.19	0.81
3	ASTE13 *ΔprgI*	P_{lacUV5} *hilA* P_{sic} *sicP sptP-bla-GFP11-2xFLAG-6xHis*	Glycerol	0.18	0.84
4	ASTE13 WT	None	Succinate	0.05	1.00
5	ASTE13 WT	P_{lacUV5} *hilA*	Succinate	0.03	1.02
6	ASTE13 *ΔprgI*	P_{lacUV5} *hilA* P_{sic} *sicP sptP-bla-GFP11-2xFLAG-6xHis*	Succinate	0.03	0.99
7	PBS		N/A	0.03	1.00
8	ASTE13 WT	None	Glycerol	0.13	0.99
9	ASTE13 WT	P_{lacUV5} *hilA*	Glycerol	0.26	0.98
10	ASTE13 *ΔprgI*	P_{lacUV5} *hilA* P_{sic} *sicP sptP-DH-GFP11-2xFLAG-6xHis*	Glycerol	0.28	0.92
11	ASTE13 WT	None	Succinate	0.03	1.11
12	ASTE13 WT	P_{lacUV5} *hilA*	Succinate	0.07	1.11
13	ASTE13 *ΔprgI*	P_{lacUV5} *hilA* P_{sic} *sicP sptP-DH-GFP11-2xFLAG-6xHis*	Succinate	0.06	1.13
14	PBS	N/A	N/A	0.01	1.00

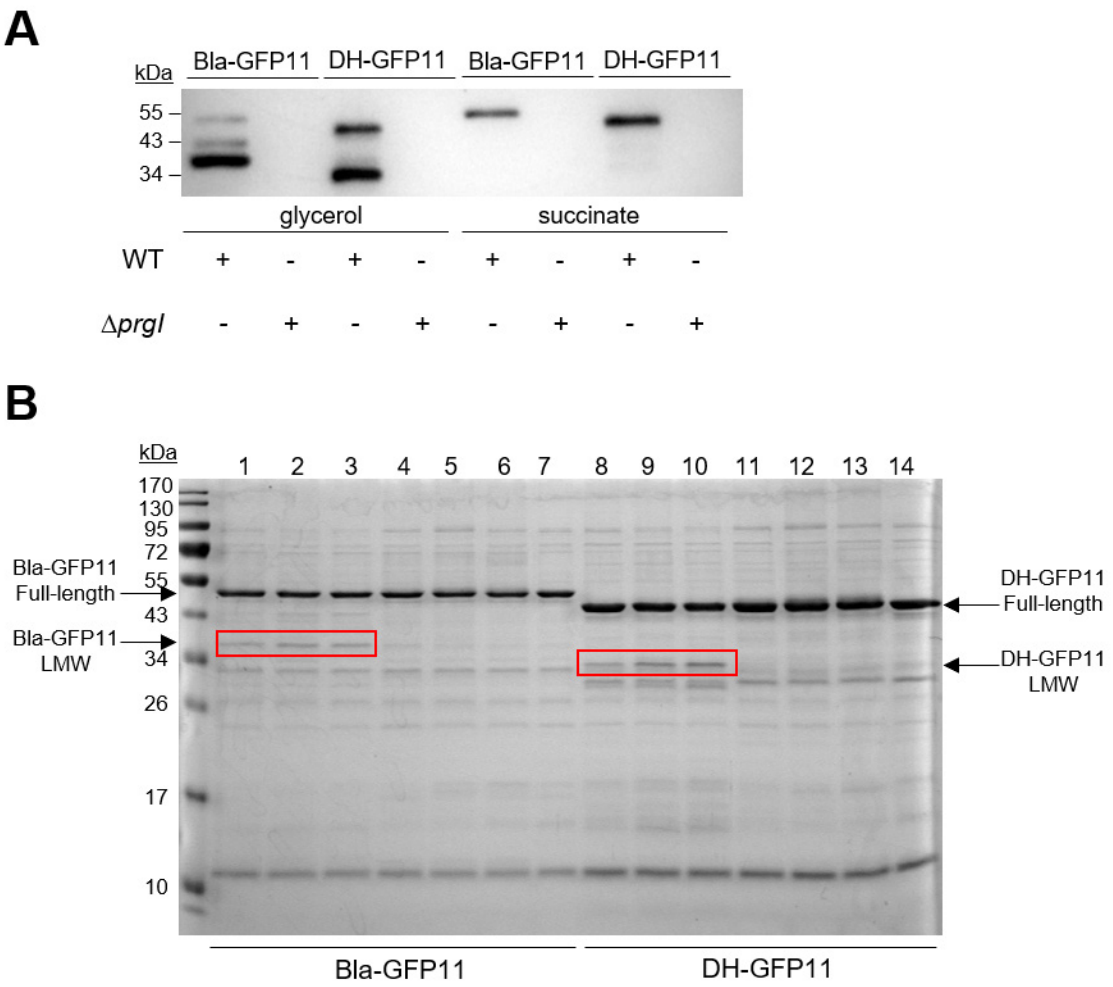


Figure 3.8 Cleavage of SptP-tagged protein in the presence of the secreted fraction.
SptP-Bla-GFP11-2xFLAG-6xHis and SptP-DH-GFP11-2xFLAG-6xHis were expressed in ASTE13 WT and ASTE13 Δ*prgI* strains in NCE with 1X EZ and 0.4% w/v glycerol or 0.4% w/v succinate, and the secreted fraction was analyzed via western blotting (**A**). SptP-Bla-GFP11-2xFLAG-6xHis and SptP-DH-GFP11-2xFLAG-6xHis were purified and mixed with fivefold-concentrated secreted fractions generated according to the conditions listed in Table 3.8. The mixtures were incubated at 37°C for 22 hours and Coomassie stained after SDS-PAGE (**B**).

Bands appeared at the same molecular weight as the main degradation product in all mixtures that included secretion fractions produced in the presence of glycerol. The ratio of the LMW fragment to the full length band and the ratio of the full-length band to the PBS control suggest that the full-length band decreased by a similar amount as the LMW fragment increased (Table 3.8). Degradation was not apparent in secreted fractions produced by cultures containing succinate.

To confirm that the cleavage was mediated by a protease, I purified a new batch of SptP-DH-GFP11-2xFLAG-6xHis and repeated the degradation experiment for 24 hours at 37°C with and without a protease inhibitor cocktail (PI). I also added LB-IM with glycerol as another condition. General degradation occurred in the samples mixed with concentrated secreted fractions from LB-IM with glycerol, likely due to the low pH of the

secreted fraction (Figure 3.9). The previously observed cleavage was apparent in samples containing the secreted fraction from NCE with glycerol, and no degradation was apparent in the samples containing the secreted fraction from NCE with succinate. The protease inhibitor cocktail slowed degradation, but the amount used was insufficient to completely inhibit cleavage.

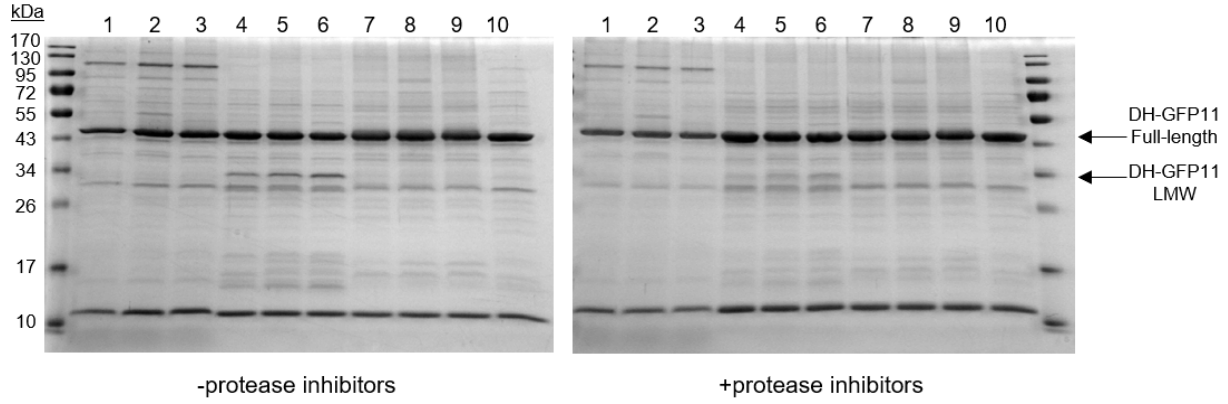


Figure 3.9 Cleavage of purified SptP-DH-2xFLAG-6xHis with and without protease inhibitors. SptP-DH-GFP11-2xFLAG-6xHis was purified and mixed with fivefold-concentrated secreted fractions generated according to the conditions listed in Table 3.9. Two identical sets of mixtures were generated, and a protease inhibitor cocktail (Sigma) was added to one set. The mixtures were incubated at 37°C for 24 hours and Coomassie stained after SDS-PAGE.

Table 3.9 Strains and fragmentation analysis.

Lane	Strain	Plasmid(s)	Medium	Carbon Source (0.4% w/v)	LMW/ Full -PI	LMW/ Full +PI
1	ASTE13 WT	None	LB-IM	Glycerol	0.03	0.03
2	ASTE13 WT	P_{lacUV5} *hilA*	LB-IM	Glycerol	0.04	0.02
3	ASTE13 *ΔprgI*	P_{lacUV5} *hilA* P_{sic} *sicP sptP-DH-GFP11-2xFLAG-6xHis*	LB-IM	Glycerol	0.05	0.02
4	ASTE13 WT	None	NCE	Glycerol	0.17	0.04
5	ASTE13 WT	P_{lacUV5} *hilA*	NCE	Glycerol	0.23	0.11
6	ASTE13 *ΔprgI*	P_{lacUV5} *hilA* P_{sic} *sicP sptP-DH-GFP11-2xFLAG-6xHis*	NCE	Glycerol	0.28	0.13
7	ASTE13 WT	None	NCE	Succinate	0.03	0.02
8	ASTE13 WT	P_{lacUV5} *hilA*	NCE	Succinate	0.05	0.01
9	ASTE13 *ΔprgI*	P_{lacUV5} *hilA* P_{sic} *sicP sptP-DH-GFP11-2xFLAG-6xHis*	NCE	Succinate	0.05	0.02
10	PBS	N/A	N/A	N/A	0.04	0.02

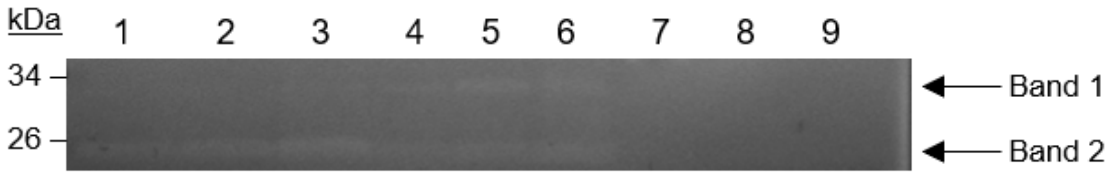


Figure 3.10 Zymogram of concentrated secreted fractions.
The fivefold-concentrated secretion fractions listed in Table 3.9 were run on a Novex Zymogram Plus gel and developed according to the kit instructions.

Zymogram gels, or PAGE gels with protein embedded in the gel matrix, are used to identify proteases. Samples are prepared in Laemmli buffer without reducing agent, and SDS-PAGE proceeds as usual. The proteases are renatured in a 2.7% Triton X-100 solution, and "developed" in a proprietary buffer from Invitrogen at 37°C to enable proteolytic activity. Active enzymes will cleave the embedded protein present at their location in the gel, and the absence of proteins at that location creates "bands" that do not stain with Coomassie. Coomassie will stain the rest of the gel, as it contains intact protein.

The concentrated secretion fractions of Table 3.9 created very faint bands on the zymogram (Figure 3.10), but a double banding pattern was present in the NCE with glycerol secreted fractions. A single band appeared in secreted fractions from LB-IM with glycerol. No bands were visible in the secreted fractions from NCE with succinate. These results suggested that more than one protease is present in the extracellular environment when glycerol (and likely glucose) is included as a carbon source.

3.4 Discussion

Growth medium optimization is a critical component of any recombinant protein production process. Many large-scale recombinant protein production processes in bacteria occur in semi-defined growth media with yeast extract as the undefined component [166]. Semi-defined media do not allow identification of specific factors that affect product titer, however. The T3SS is particularly sensitive to environmental cues, so creating and optimizing a defined medium is critical for increased and robust protein secretion titer. The results of this chapter showed that unoptimized defined media supplemented with biosynthetic precursors such as amino acids supported similar growth characteristics and secretion titers as LB-L. Including a high concentration of phosphate in combination with glucose or glycerol as a carbon source provided at least a twofold increase in secretion titer, though secretion rate varied with medium formulation.

Despite the increase in secretion titer provided by glucose or glycerol, the use of those carbon sources is complicated by the fragmentation observed in the secreted fraction after eight hours. The fragments are likely a result of proteolytic cleavage in SptP. Though cleavage in SptP does not affect the function of the secreted protein, creating multiple species complicates secreted protein quantification and downstream processing. Identifying the proteases and knocking them out would solve this issue, but that is infeasible if the proteases are essential for cell viability. Other solutions include using a medium with an increased secretion rate, such as PCN+P, to allow a harvest at eight hours or using PCN+P and succinate with a harvest at 24 hours. Glycerol is a superior

carbon source to succinate in terms of both energetics [161] and raw material costs, however, so finding an engineering solution that allows use of glycerol is optimal.

Finding an engineering solution to prevent uncontrolled proteolytic cleavage and performing a broader and more systematic optimization of medium formulation will require evaluating many conditions, but the low throughput of western blotting does not support such experiments. The roughly optimized PCN+P medium demonstrates that defined media have potential to high secretion titers, but a design-of-experiments (DoE) project to address other components such as divalent cations and more systematically optimize the combination of medium components will surely yield higher secretion titers. Developing an assay that is at least medium throughput for screening secretion titer, then, is essential to advance medium optimization and maximize secretion titer.

Chapter 4: Methods for Measuring Protein Secretion Titer

4.1 Introduction

Protein quantification is a critical aspect of engineering a system for protein production. Experimental scope and throughput are intrinsically tied to the capabilities of the analytical methods available for measuring protein concentration. Engineering projects to increase protein production capacity are often conducted with a model protein reporter that facilitates accurate and high-throughput quantification. In choosing a protein reporter, one must consider several factors, starting with the compatibility of the reporter with the target system. Second, the expected protein amount must be well above the limit of detection of the assay used to measure the reporter. Finally, experimental parameters or solution conditions might affect reporter activity—it is important to ensure that the reporter is measuring only protein quantity across all experimental conditions unless reporter performance is deliberately included in the experimental design.

The T3SS presents several unique challenges for protein quantification. First, we observe secretion titers on the order of 1-10 mg/L for most proteins, which is near or below the limit of detection for many techniques, especially in the complex milieu of rich bacterial growth media. Second, β-barrel fluorescent proteins such as GFP are not secreted [176]. Third, secreted proteins fold in the extracellular space, which complicates the use of assays that report on a correctly folded protein such as ELISA and enzyme activity [38]. Changing the extracellular solution environment as described in Chapters 2 and 3 changes the folding environment, and then it is impossible to dissociate protein folding and absolute protein quantity. Finally, the secreted protein is not 100% pure, and the ratio of natively secreted proteins to a heterologous secreted protein is not characterized. This renders general quantification methods such as BCA, Bradford, and UV absorption unreliable even in the case where titer exceed the limit of detection.

Our secreted protein titer measurements were previously limited to western blotting, enzyme activity assays, and ELISAs. ELISAs and enzyme activity assays report on the folded state of the protein and thus were incompatible with the projects outlined in Chapters 2 and 3. Western blotting is universally applicable, but it is too variable and labor-intensive for high-throughput characterization. In this chapter, I will describe several quantification methods that I evaluated and optimized for measuring secreted protein titer in a high-throughput manner.

4.2 Methods

4.2.1 Strains and Growth Conditions

Strains and plasmids used are listed in Table 4.1 and Table 4.2. Secretion experiments were started by growing a single colony from a fresh streak of a frozen glycerol stock in the lysogeny broth Lennox formulation (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) with appropriate antibiotics (34 μg/mL chloramphenicol, 50 μg/mL kanamycin) for 12-16 hours overnight in an orbital shaker at 37°C and 225 rpm. Overnight cultures were diluted 1:100 into the appropriate medium supplemented with 100 μg/mL isopropyl β-D-1-thiogalactopyranoside (IPTG) with appropriate antibiotics. All culturing steps were performed in 24-well deepwell plates (Axygen). Secretion was performed at 37°C and 225 rpm in an orbital shaker for the specified time. The secreted fraction was

harvested by centrifuging cultures at 4000 *x g* for 10 minutes. SDS-PAGE samples for the secretion fraction were prepared by adding supernatant to Laemmli buffer [150] in a 3:1 ratio unless otherwise specified. SDS-PAGE samples for whole culture lysate were prepared by mixing cell suspension with Laemmli buffer in a 1:2 ratio prior to centrifugation. All SDS-PAGE samples were boiled at 95°C for 5 minutes immediately after preparation.

Table 4.1 Strains used in Chapter 4.

Strain Name	Comment	Reference
ASTE13	LT2-derived lab strain similar to DW01	This study; DW01 [65]
ASTE13 Δ*prgI*	*prgI* knockout	[41]

Table 4.2 Plasmids used in Chapter 4.

Plasmid Name	ORFs under inducible control		ORI	ab^R	Reference
P_{sic} *DH*	*sicP*	*sptP-DH-2xFLAG-6xHis*	colE1	cam	[41]
P_{sic} *bla*	*sicP*	*sptP-bla-2xFLAG-6xHis*	colE1	cam	[41]
P_{sic} *AP*	*sicP*	*sptP-phoA-2xFLAG-6xHis*	colE1	cam	[38]
P_{sic} *bla-3xTC-tags*	*sicP*	*sptP-bla-3xTC-2xFLAG-6xHis*	colE1	cam	This study
P_{sic} *bla-FLAG-TC*	*sicP*	*sptP-bla-2xFLAG-1xTC-6xHis*	colE1	cam	This study
P_{sic} *bla-TC*	*sicP*	*sptP-bla-2xFLAG-6xHis-1xTC*	colE1	cam	This study
P_{sic} *DH-GFP11*	*sicP*	*sptP-DH-GFP11-2xFLAG-6xHis*	colE1	cam	This study
P_{sic} *AP-GFP11*	*sicP*	*sptP-AP-GFP11-2xFLAG-6xHis*	colE1	cam	This study
P_{lacUV5} *hilA*	*hilA*		p15a	kan	[41]

4.2.2 DNA Manipulations

PCR was performed with Phusion DNA polymerase using the primers listed in Table 4.3. Golden gate cloning was used to construct plasmids [151]. Genes for proteins to be secreted were inserted into a modified pPROTet.133 backbone vector (BD Clontech) under the control of the sic promoter [8]. All secretion plasmids expressed the SptP chaperone *sicP* and the *sptP* secretion signal (nucleotides 1 to 477). The SptP secretion tag was fused N-terminal to the protein of interest, and 2xFLAG and 6xHis tags were fused C-terminal to the protein of interest. The GFP11, 1xTC, and 3xTC tags were added C-terminal to the protein of interest using 3' extensions of the primers to append the tag 3' to the gene of interest. The GFP(1-10) gene was ordered as a gBlock with golden gate-compatible cloning arms from IDT and cloned into a pET28b+ vector (Novagen). All cloning was done in *E. coli* DH10B cells, and all DNA sequences were confirmed by Sanger sequencing (Quintara).

Table 4.3 Primers used in Chapter 4.

Sequence	Amplicon	Used to Construct
A GGTCTC A **GCTT** cacccagaaacgctggtga	*bla*	*sptP-bla-3xTC-2xFLAG-6xHis* *sptP-bla-2xFLAG-1xTC-6xHis* *sptP-bla-2xFLAG-6xHis-1xTC*
A GGTCTC A **CAGA** ccaatgcttaatcagtgaggcacc	*bla*	*sptP-bla-3xTC-2xFLAG-6xHis*
A GGTCTC A **TCTG** *GGGGATCGTTTCTTAACTGTTGTCCAGGG* *TGTTGCATGGAACCGGGTAGTTTTCT* gaattgttgtccgggatgct	*3xTC*	*sptP-bla-3xTC-2xFLAG-6xHis*
A GGTCTC A **CGCT** *CGGCTCCATGCAGCAACCAGGGCAGCAAT* *TCAGGAAAGATCCCGGTTCCATAC* agcatcccggacaacaattc	*3xTC*	*sptP-bla-3xTC-2xFLAG-6xHis*
A GGTCTC A **TCTA** gaggcatcaaataaaacgaaaggctcagtc	*pPROTet.133 backbone*	*sptP-bla-2xFLAG-1xTC-6xHis* *sptP-bla-2xFLAG-6xHis-1xTC*
A GGTCTC A **AAGC** ttactttctgctccaacatcgttatttttttc	*pPROTet.133 backbone*	*sptP-bla-2xFLAG-1xTC-6xHis* *sptP-bla-2xFLAG-6xHis-1xTC*
A GGTCTC A **CGCT** ccaatgcttaatcagtgaggcacc	*bla*	*sptP-bla-2xFLAG-1xTC-6xHis*
A GGTCTC A **TAGA** *TTAACTACCCGGTTCCATGCAACACCCTG* *GACAACAGTTAAGAAACGATCC* gtggtgatggtgatgatgcttg	*bla*	*sptP-bla-2xFLAG-6xHis-1xTC*
A GGTCTC A **GCTT** gatatgttgaccccaactgaaag	*DH-GFP11*	*sptP-DH-GFP11-2xFLAG-6xHis*
A GGTCTC C **CGCT** *GGTGATGCCCGCCGCGTTCACGTATTCAT* *GCAGCACCATATGATCCCGCGAACCACCC* *CCAGATCCACCCCC* agagttctccttctcccgc	*DH-GFP11*	*sptP-DH-GFP11-2xFLAG-6xHis*
A GGTCTC A **GCTT** cacccagaaacgctggtga	*bla-GFP11*	*sptP-bla-GFP11-2xFLAG-6xHis*
A GGTCTC C **CGCT** *GGTGATGCCCGCCGCGTTCACGTATTCAT* *GCAGCACCATATGATCCCGCGAACCACCC* *CCAGATCCACCCCC* ccaatgcttaatcagtgagg	*bla-GFP11*	*sptP-bla-GFP11-2xFLAG-6xHis*
A GGTCTC C **GCTT** cggacaccagaaatgcctgttctgg	*phoA-GFP11*	*sptP-phoA-GFP11-2xFLAG-6xHis*
A GGTCTC C **CGCT** *GGTGATGCCCGCCGCGTTCACGTATTCAT* *GCAGCACCATATGATCCCGCGAACCACCC* *CCAGATCCACCCCC* tttcagccccagagcggc	*phoA-GFP11*	*sptP-phoA-GFP11-2xFLAG-6xHis*

4.2.3 Protein Separation and Western Blotting

Samples were separated by SDS-PAGE and transferred to a polyvinylidene fluoride membrane (PVDF, Millipore) for western blotting using the Bio-Rad Criterion blotter. Samples were diluted as specified in the text before loading on SDS-PAGE. Membranes were probed with mouse anti-FLAG per manufacturer's instructions (Sigma Aldrich). To facilitate chemiluminescent detection, a secondary labeling step was performed with goat anti-mouse IgG (H+L) HRP conjugate according to manufacturer's instructions (Thermo Fisher). Bands were detected with the SuperSignal West Pico or Pico Plus substrate (Thermo Fisher) and a ChemiDoc XRS + imaging system (Bio-Rad).

4.2.4 Protein Quantification via Stained PAGE Gels

Absolute secretion titers were measured by performing SDS-PAGE and staining with Coomassie according to Studier [153] or SYPRO Ruby (Invitrogen) according to the manufacturer's instructions. Densitometry was performed in Image Lab software (Bio-Rad) and used to estimate titer relative to a bovine serum albumin (BSA) standard curve (Thermo). Error bars are standard deviation on three biological replicates unless otherwise specified.

4.2.5 Protein Quantification via Western Blotting

All relative protein quantities were calculated by performing densitometry using Image Lab software (Bio-Rad) and normalizing as described. Samples were diluted as specified, and relative protein amounts were corrected for dilution if appropriate. For calculation of absolute secretion titer, unknown samples were compared to a standard curve of purified SptP-DH-2xFLAG-6xHis. The unknown samples were diluted to fall within the range of the standard curve. Error bars are standard deviation on three biological replicates unless otherwise specified.

4.2.6 Protein Purification

All proteins were purified from *S. enterica* Typhimurium ASTE13 Δ*prgI* except for the SptP-DH-2xFLAG-6xHis standard used for the western blotting standard curve. For the proteins purified from ASTE13 Δ*prgI*, cells were grown and induced as described in "Strains and Growth Conditions" except that the growth vessel was a 250 mL Erlenmeyer flask containing 50 mL of LB-L medium. Cells were harvested by pelleting at 5000 *x g* for 10 minutes. The supernatant was discarded, and the cell pellets were frozen at -80°C. Cell pellets were thawed and resuspended in 10 mL 20 mM sodium phosphate pH 7.4 with 500 mM NaCl and 20 mM imidazole. The resuspended cell solution was homogenized via sonication. The soluble and insoluble fractions were separated by centrifugation at 17,000 *x g* for 20 minutes at 4°C.

The SptP-DH-2xFLAG-6xHis standard was grown and induced as described in "Strains and Growth Conditions", except that the overnight culture was diluted 1:100 into 50 mL terrific broth (TB) in a 250 mL Erlenmeyer flask with the appropriate volumes of IPTG and antibiotics. Secretion was performed for 8 hours. The supernatant was harvested by pelleting cells at 5000 *x g* for 10 minutes and filtering the supernatant through a 0.2 μm syringe filter. The supernatant was concentrated and buffer-exchanged into 20 mM sodium phosphate pH 7.4 with 500 mM NaCl and 20 mM imidazole using Corning Spin-X UF 20 concentrators.

The soluble fraction or buffer-exchanged secreted fraction was filtered with a 0.2 µm syringe filter before being applied to a His GraviTrap column (GE Life Sciences) for purification. The eluted protein was buffer-exchanged into phosphate-buffered saline using a PD-10 desalting column and quantified compared to a BSA standard using densitometry on an SDS-PAGE gel stained with Coomassie [153].

4.2.7 Split GFP Assay

The split GFP assay was performed as described in Cabantous and Waldo [167]. Briefly, $GFP(1\text{-}10)_{opt}$ was expressed, purified, and frozen in ~50 mg pellets according to Cabantous and Waldo [167]. The $GFP(1\text{-}10)_{opt}$ pellet was resuspended in 500 µL 8 M urea with fresh 5 mM DTT and incubated at 37°C for five minutes to facilitate full dissolution of the pellet. The dissolved inclusion body was added dropwise to 10 mL TNG buffer (100 mM Tris pH 7.4, 150 mM NaCl, 10% glycerol v/v) to facilitate refolding and gently inverted to prepare the $GFP(1\text{-}10)_{opt}$ working solution. The GFP11-tagged protein of interest was secreted or purified as described above, and either 20 µL of the secreted fraction or 20 µL of purified protein was mixed with 180 µL of the $GFP(1\text{-}10)_{opt}$ working solution in a black opaque 96-well plate (Costar). Fluorescence was measured using a BioTek Synergy HTX plate reader with a 485/20 nm excitation filter and a 516/20 nm emission filter.

4.2.8 Tetracysteine Tag Assay

The tetracysteine (TC) tag assay was performed as described in Haitjema *et al.* [55]. Briefly, 50 µL of 4 mM dithiothreitol (DTT, Alfa Aesar) was added to 50 µL of protein sample in a black opaque 96-well plate (Costar). The 2 mM stock solution of FlAsH-EDT_2 or ReAsH-EDT_2 (Thermo Scientific) was diluted to a 2 µM working stock in PBS, and 100 µL of the 2 working stock was mixed with the protein sample and reducing agent. The labeling proceeded as specified and fluorescence was measured using a BioTek Synergy HTX plate reader with a 485/20 nm excitation filter and a 516/20 nm emission filter for FlAsH-EDT_2 and a 560/40 nm excitation filter and a 620/15 nm emission filter for ReAsH-EDT_2.

For in-gel visualization, 10X Bond-Breaker™ TCEP Solution (0.5 M, Thermo Scientific Pierce) was diluted to 1X in 4X Laemmli buffer lacking a reducing agent. A 30 µL aliquot of protein sample was mixed with 10 µL of 4X Laemmli + TCEP and boiled at 95°C for 5 minutes. A 50 µM working solution of FlAsH-EDT_2 was prepared by diluting the 2 mM frozen stock in PBS. The boiled samples were cooled to room temperature, and 1 µL of the 50 µM FlAsH-EDT_2 working solution was added. Labeling proceeded at room temperature for 20 minutes, after which the samples were loaded on an SDS-PAGE gel. The gel was washed for 5 minutes in ddH_2O and imaged using a ChemiDoc XRS+ on its "SYPRO Ruby" preset protocol. The gels were post-stained with SYPRO Ruby (Invitrogen) according to manufacturer's instructions and imaged again with a ChemiDoc XRS+.

4.2.9 Alkaline Phosphatase Activity Assay

Alkaline phosphatase activity was measured by monitoring *p*-nitrophenol phosphate (pNPP) cleavage. A stock solution of 0.1 M pNPP (Sigma Aldrich) prepared in 1 M Tris pH 8.0 and frozen at -20°C was thawed and diluted tenfold in 1 M Tris pH 8.0 to create a 0.01 M pNPP working solution. If necessary, the pNPP working solution was

diluted further to achieve the specified final concentration after adding 40 µL to the assay mixture. The specified volume of protein solution was added to 140 µL 1 M Tris pH 8.0 and 40 µL of working pNPP solution in a clear flat-bottom 96-well plate (Grenier Bio-One) to a final volume of 200 µL. The mixture was incubated at 37°C as described and absorbance at 405 nm was recorded by a BioTek Synergy HTX plate reader.

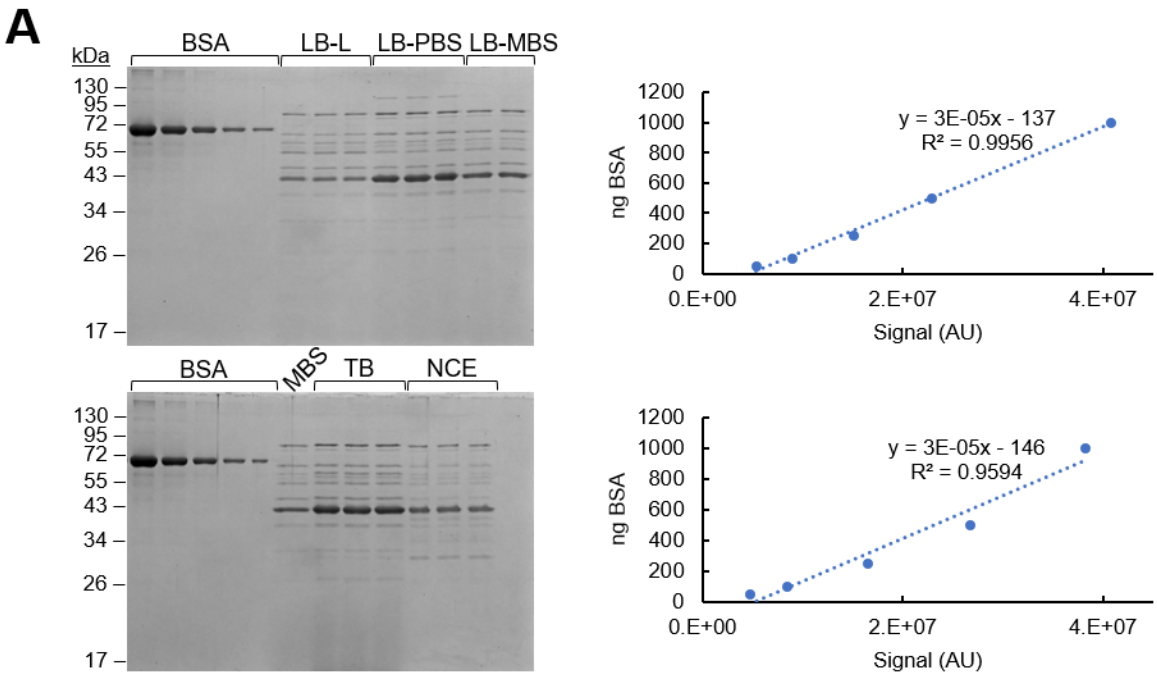
4.3 Results

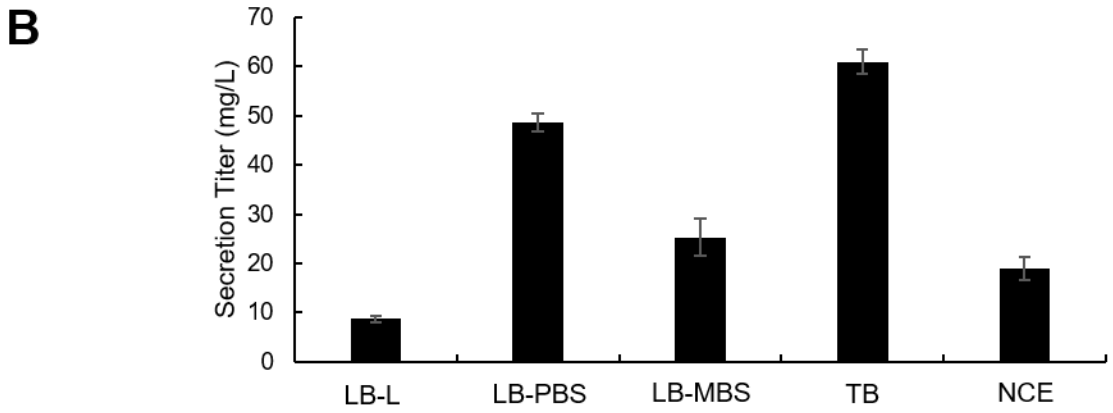
4.3.1 Western blots and directly stained SDS-PAGE gels are universal analysis techniques that require careful optimization

Western blotting and directly stained SDS-PAGE gels are central to any scientific study involving protein and are nearly universally applicable. Conventional direct staining of SDS-PAGE gels is non-specific; all proteins present are stained. The sensitivity and linearity of directly stained SDS-PAGE gels is dependent on the stain used [168]. Western blotting is specific and sensitive to femtomolar amounts of protein, but the blotting process is laborious and introduces many sources of variability. The linear range of analysis depends on the relative amounts of samples present (resulting from the amount of protein loaded on the SDS-PAGE gel and the electrophoretic transfer conditions) and the substrate used to visualize labeled bands [169–171].

Prior to the improvements in growth media described in Chapters 2 and 3, western blotting was the only available technique for measuring the secretion titer of most proteins. In fact, the most sensitive available substrate compatible with our western blotting workflow (SuperSignal West Femto, Thermo Scientific Pierce) was required to detect several proteins even after concentrating the secreted fraction. The increase in titer provided by HilA overexpression allowed detection of the model protein SptP-DH-2xFLAG-6xHis secreted in LB-L with the lower sensitivity SuperSignal West Pico. The signal did not saturate for several minutes, however, so when I began the medium optimization experiments outlined in Chapters 2 and 3 it did not occur to me that the medium improvements might cause titer to exceed the linear range for detection relative to LB-L.

The relative secretion titers calculated from western blotting were often highly variable, but I attributed that to the many steps at which error accrues during the blotting process. As I was optimizing the ideal formulation of potassium phosphate, sodium chloride, and glycerol described in Chapter 2, I estimated secretion titer of SptP-DH-2xFLAG-6xHis relative to a BSA standard curve on a Coomassie-stained SDS-PAGE gel (Figure 4.1). Titer estimates from the direct staining indicated that western blotting was underestimating the relative increases in secretion titer observed in TB and the optimized LB media. For example, semi-quantitative western blotting showed that TB provided a threefold increase in secretion titer relative to LB-L (Figure 4.2A), while the titers calculated from quantitative Coomassie gels showed a six-fold increase in TB relative to LB-L (Figure 4.1).





	LB-L	LB-PBS	LB-MBS	TB	NCE
0.4% w/v glycerol	-	+	+	+	+
1X EZ Supplement	-	-	-	-	+
K_2HPO_4 (mM)	0	32	0	72	34
KH_2PO_4 (mM)	0	8	0	17	29
MOPS (mM)	0	0	90	0	0
NaCl (mM)	86	150	199	0	0

Figure 4.1 Coomassie quantification of SptP-DH-2xFLAG-6xHis secretion titer.
Secretion of SptP-DH-2xFLAG-6xHis was performed following P_{lacUV5} *hilA* induction in the specified media for 8 hours. SDS-PAGE samples were prepared by mixing 90 µL of secreted fraction with 30 µL of 4X Laemmli buffer and boiling at 95°C for 5 minutes. BSA standards were 100, 50, 25, 10, and 5 ng/µL. 10 µL of each standard and unknown sample were loaded on a 12% tris-glycine gel. Titers in (**B**) were calculated from the gels shown in (**A**).

Including a calibration curve on each blot would cause a precipitous decrease in the already low throughput, so I compromised by diluting samples to achieve a signal within 50% of the normalization condition. I also ran triplicate of the normalization condition and normalized unknowns to the average of those values to compensate for biological and technical variation. These modifications significantly decreased variability and generated western blotting results that better matched the fold changes in titer observed from the quantitative Coomassie gel (Figure 4.2).

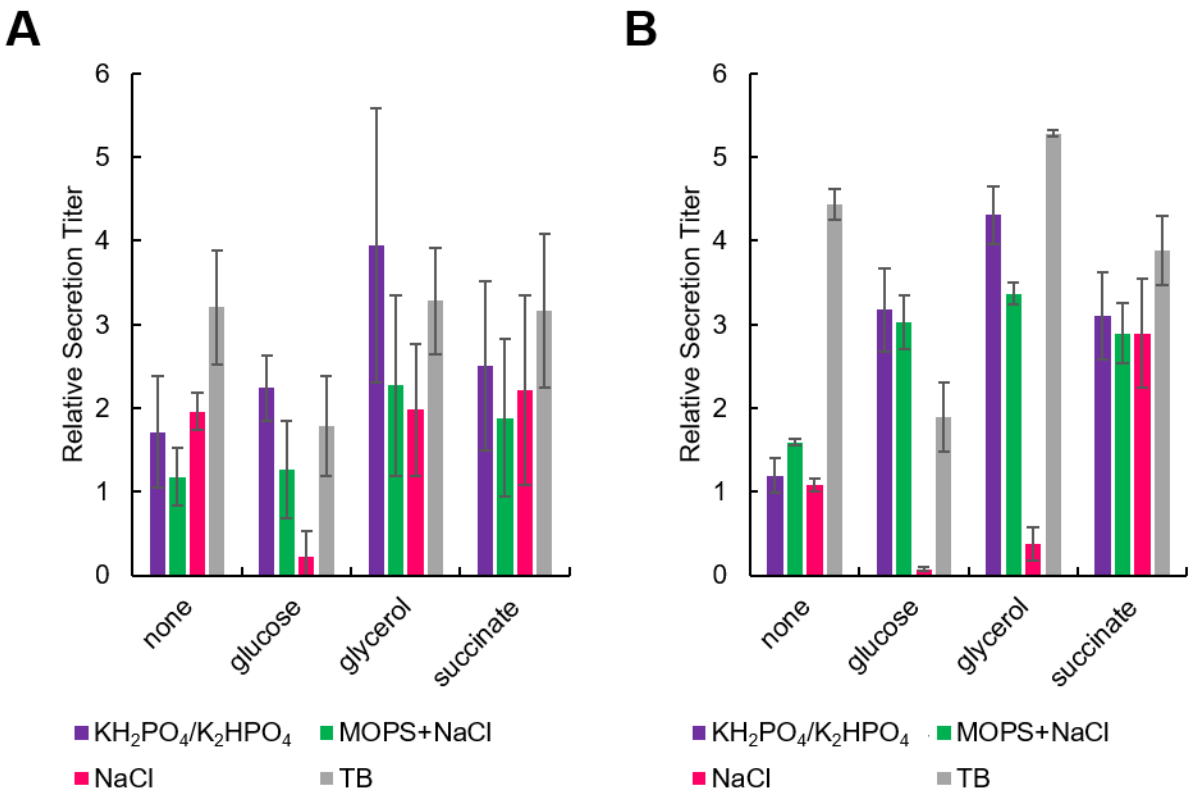


Figure 4.2 Relative secretion titer calculated from western blotting with different sample dilutions. Secretion of SptP-DH-2xFLAG-6xHis was performed following P_{lacUV5} *hilA* induction for 8 hours. Bars were generated from densitometry performed on the western blots represented in Figure 4.3. Samples analyzed in (**A**) were normalized to a single LB-L sample on each gel. Samples for (**B**) were normalized to the average of the LB-L biological replicates on each gel. All carbon sources were added to 0.4% w/v. [KH_2PO_4/K_2HPO_4] and [MOPS] were 90 mM and pH 7.4, [NaCl] in "NaCl" was 234 mM, and [NaCl] in "MOPS+NaCl" was 199 mM. All media except TB included 10 g/L tryptone and 5 g/L yeast extract.

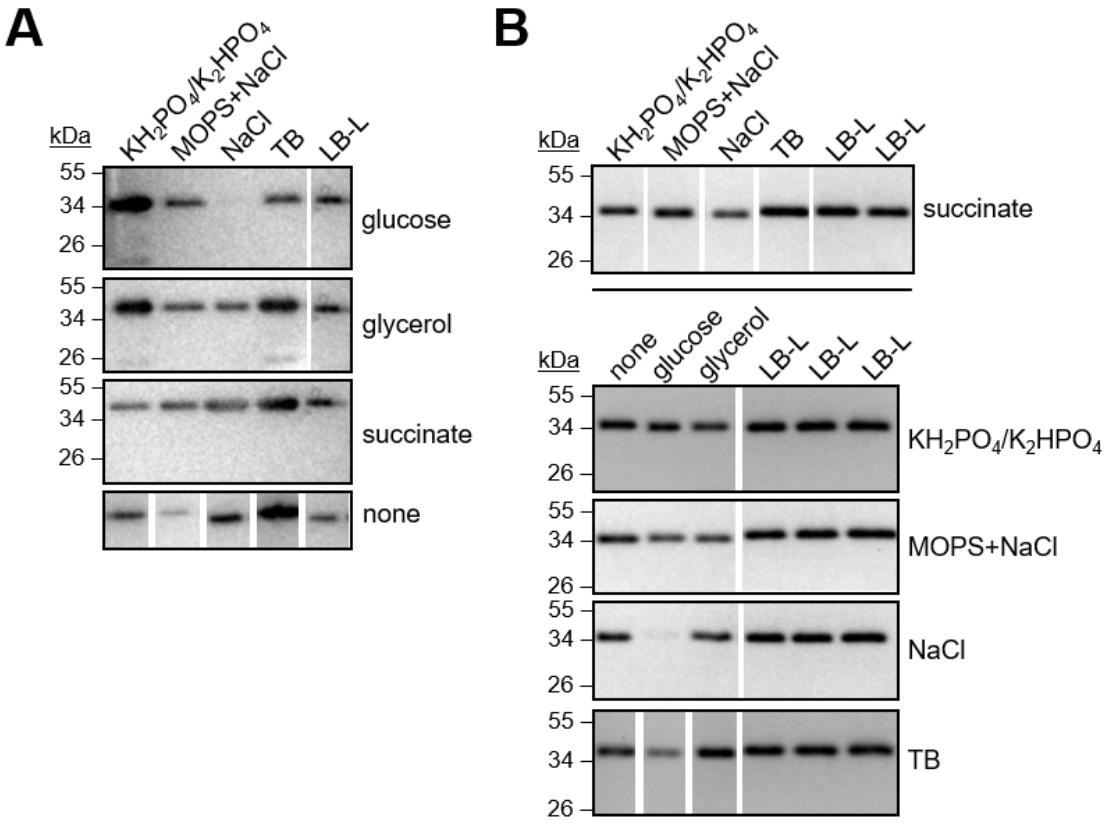
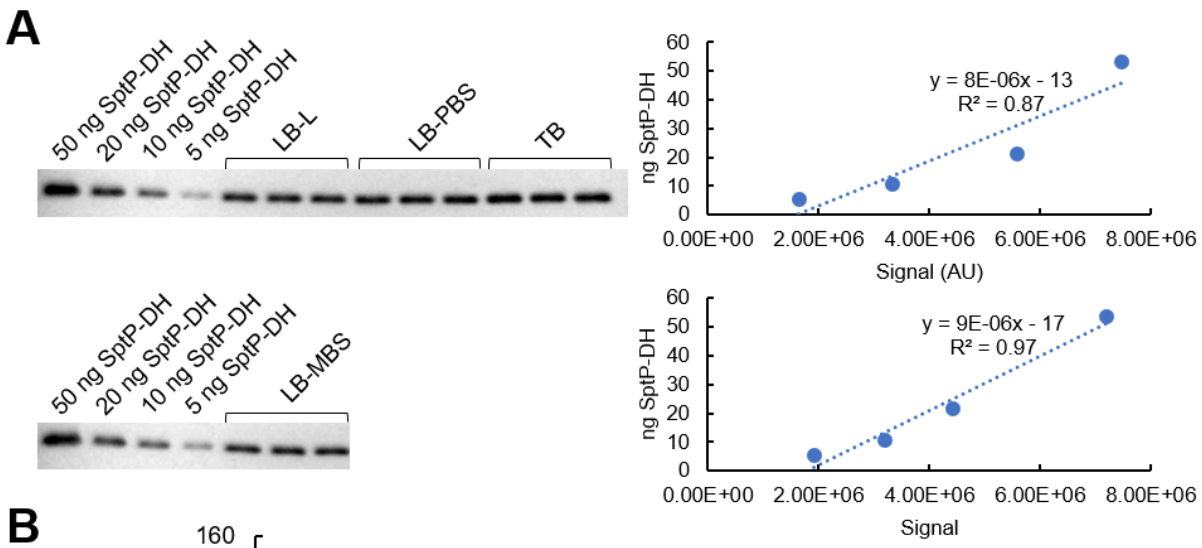
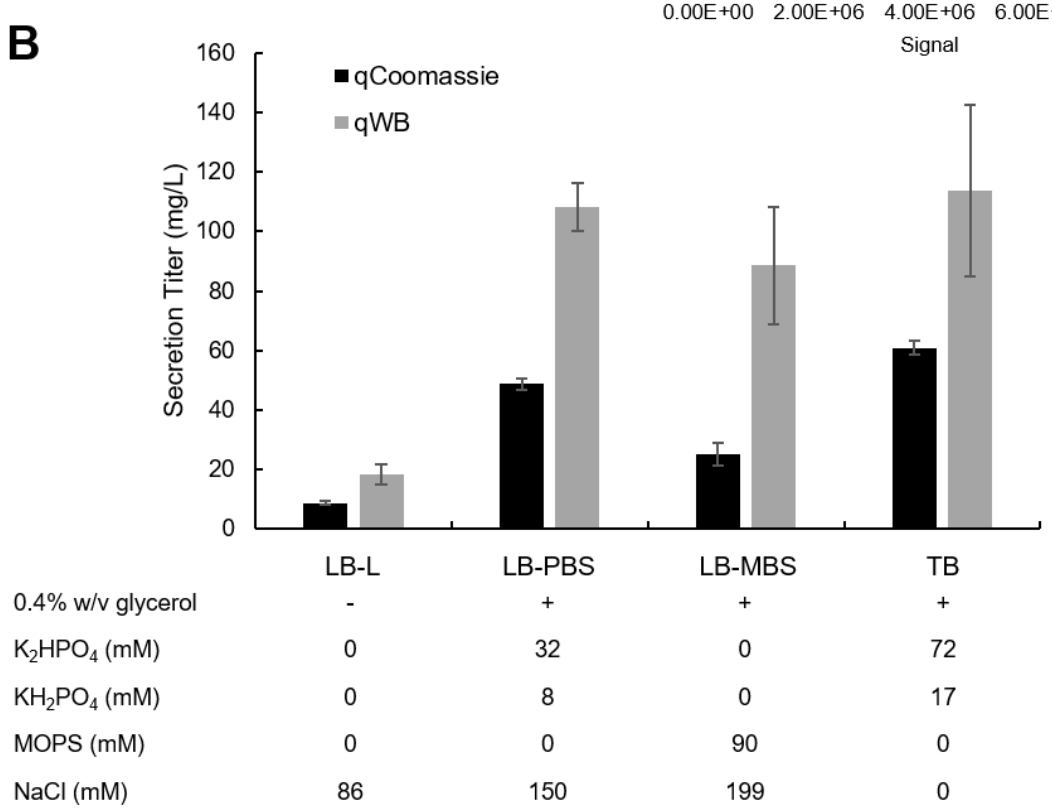


Figure 4.3 Representative western blots for Figure 4.2.
Samples for (**A**) were prepared by combining 40 µL of the secreted fraction with 16 µL 4X Laemmli buffer. 10 µL of each sample was loaded on the gel. Samples for (**B**) were prepared by combining 20 µL of the secreted fraction with 100 µL 1.25X Laemmli. Samples from TB with any carbon source, KH_2PO_4/K_2HPO_4 with glucose, glycerol or succinate, MOPS+NaCl with glucose, glycerol, or succinate, and NaCl with succinate were diluted an additional fourfold in 1X Laemmli buffer. Gel loading was: 12 µL of LB-L, NaCl with glucose, and NaCl with glycerol; 6 µL of NaCl, KH_2PO_4/K_2HPO_4, and MOPS+NaCl with no added carbon source ("none"); and 8 µL of 0.25X additionally diluted samples. Boxed bands are from the same gel but were rearranged for clarity.

Estimating secretion titer via quantitative Coomassie gels also cast doubt on the secretion titers previously reported from quantitative western blotting. When I compared SptP-DH-2xFLAG-6xHis secretion titer calculated from a quantitative western blot with a purified SptP-DH-2xFLAG-6xHis as a standard to the titers calculated in Figure 4.1, I discovered that the values differed at least twofold (Figure 4.4).

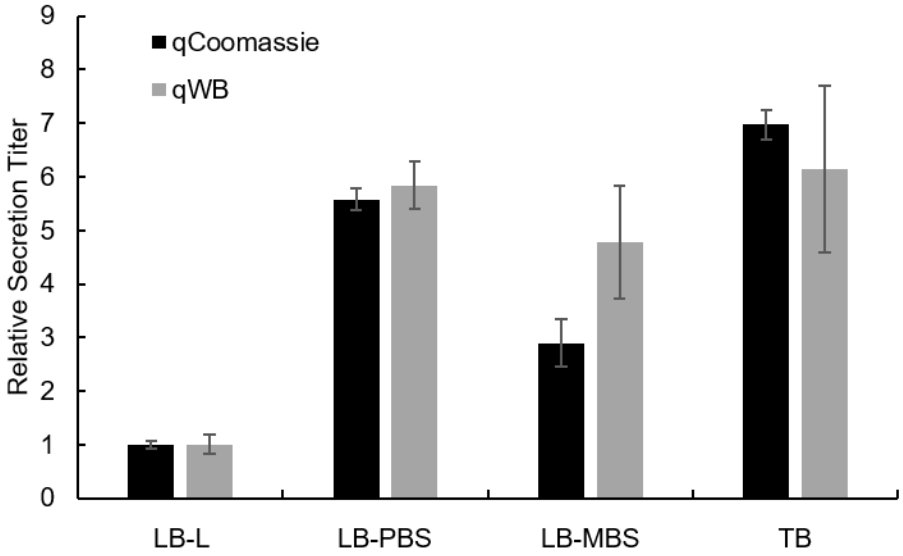




	LB-L	LB-PBS	LB-MBS	TB
0.4% w/v glycerol	-	+	+	+
K_2HPO_4 (mM)	0	32	0	72
KH_2PO_4 (mM)	0	8	0	17
MOPS (mM)	0	0	90	0
NaCl (mM)	86	150	199	0

Figure 4.4 Secretion titers calculated from a quantitative Coomassie gel and a quantitative western blot.
Secretion of SptP-DH-2xFLAG-6xHis was performed for 8 hours following P_{lacUV5} *hilA* induction. Medium formulations were as specified. The purified SptP-DH-2xFLAG-6xHis was quantified compared to a BSA standard curve on a SYPRO Ruby-stained gel. Samples for western blotting were prepared by combining 20 µL of the secreted fraction with 1.25X Laemmli buffer and boiling at 95°C for 5 minutes. All samples except LB-L were diluted fourfold further in 1X Laemmli buffer, and 10 µL of each sample was loaded. Titers from western blotting ("qWB") shown in (**B**) were calculated from the blots in (**A**) and "qCoomassie" titers are from Figure 4.1.

The discrepancy between secretion titers calculated from Coomassie gels and western blots is difficult to explain, especially because the concentration of the purified SptP-DH-2xFLAG-6xHis standard was also estimated from a BSA standard curve. The standard curves are linear, and all unknown samples fall within the range of the standard curve. If one calculates secretion titer relative to LB-L with the titers from each method, however, the ratios match each other and the results of Chapter 2 (Figure 4.5). Direct gel staining is generally less error-prone than quantitative western blotting, likely as a result of variability generated from the additional steps required for blotting. Thus, direct gel staining is preferable to western blotting for quantitative secretion titer measurements.



	LB-L	LB-PBS	LB-MBS	TB
0.4% w/v glycerol	-	+	+	+
K_2HPO_4 (mM)	0	32	0	72
KH_2PO_4 (mM)	0	8	0	17
MOPS (mM)	0	0	90	0
NaCl (mM)	86	150	199	0

Figure 4.5 Secretion titer relative to LB-L calculated from estimated titers.
Secretion titers of SptP-DH-2xFLAG-6xHis estimated from quantitative Coomassie gels and quantitative western blots (Figure 4.4) were divided by the average LB-L titer to generate relative secretion titers. Error bars represent standard deviation on three biological replicates.

4.3.2 The tetracysteine tag is not ideal for T3SS secretion titer measurement

The 14-amino acid tetracysteine (TC) tag developed by Tsein and colleagues contains a -Cys-Cys-Pro-Gly-Cys-Cys- motif that binds a biarsenical dye compound to generate red (ReAsH-EDT_2) or green (FlAsH-EDT_2) fluorescence [172]. The TC tag can theoretically be appended to any protein, and its short length should have a minimal impact on the fusion partner's function. Additionally, fluorescence produced from the TC tag should be directly proportional to the amount of secreted protein in multiple growth media, as the output is not dependent on a correctly folded protein. I endeavored to adapt

the TC tag assay described in Haitjema *et al.* [55] as a high-throughput alternative to our previously available methods.

Rich growth media provide high background fluorescence in the green spectrum, so I chose to optimize this assay in the amino-acid supplemented NCE media described in Chapter 3. I created three different constructs containing the TC tag fused to beta-lactamase (Bla) because number and orientation of TC tags can affect fluorescence [173]. The first construct had three repeats of the TC tag N-terminal to the 2xFLAG and 6xHis tags, the second had a single TC tag repeat between the 2xFLAG and 6xHis tags, and the third had a single TC tag at the C-terminus of the fusion protein.

Haitjema *et al.* used 1 µM FlAsH-EDT_2 reagent to label secreted proteins, but they did not report secretion titer explicitly [55]. We reported a secretion titer on the order of 1 mg/L for SptP-Bla-2xFLAG-6xHis with P_{lacUV5} *hilA* overexpression in LB-L [41], but for safety I assumed a secretion titer of 10 mg/L. If one assumes a minimum 1:1 molar ratio of reagent to TC sites, calculating the molar concentration of available TC sites will yield the minimum concentration of reagent required. The molar concentration of TC sites for SptP-Bla-3xTC-2xFLAG-6xHis secreted at 10 mg/L is:

$$\mathrm{MW}_{\mathrm{SptP-Bla-3xTC-2xFLAG-6xHis}} = 55404\ \mathrm{Da}\ [=] \frac{\mathrm{g}}{\mathrm{mol}}$$

$$\left(10 \frac{\mathrm{mg}}{\mathrm{L}}\ \text{secretion titer}\right) \cdot \frac{\mathrm{mol}}{55404\mathrm{g}} \cdot (3\ \text{TC sites}) = 0.54\ \mu\mathrm{M}\ \text{TC sites}$$

Secretion titer is likely lower than 10 mg/L based on our previously reported titer, so using 2 µM FlAsH-EDT_2 or ReAsH-EDT_2 should strike an appropriate balance between conserving the expensive reagent and providing a sufficient molar excess to label all secreted protein.

To test the various TC-tagged constructs, I used the secreted fractions of SptP-Bla-2xFLAG-6xHis expressed in ASTE13 WT and the various TC tag fusions expressed in ASTE13 Δ*prgI* strains to provide background fluorescence conditions. Signal-to-noise ratios were calculated by dividing the fluorescence of TC tag variants by the fluorescence of either SptP-Bla without a TC tag or ASTE13 Δ*prgI* strains carrying the same TC tag variant (Figure 4.6). A ratio greater than one would indicate that there was a positive signal resulting from the presence of the TC tag. A ratio equal to or less than one would indicate that the TC tag was unable to produce a positive signal above non-specific background levels. Unfortunately, none of the secreted constructs showed a positive difference in fluorescence compared to the corresponding Δ*prgI* strains and only a small positive difference in fluorescence compared to SptP-Bla-2xFLAG-6xHis.

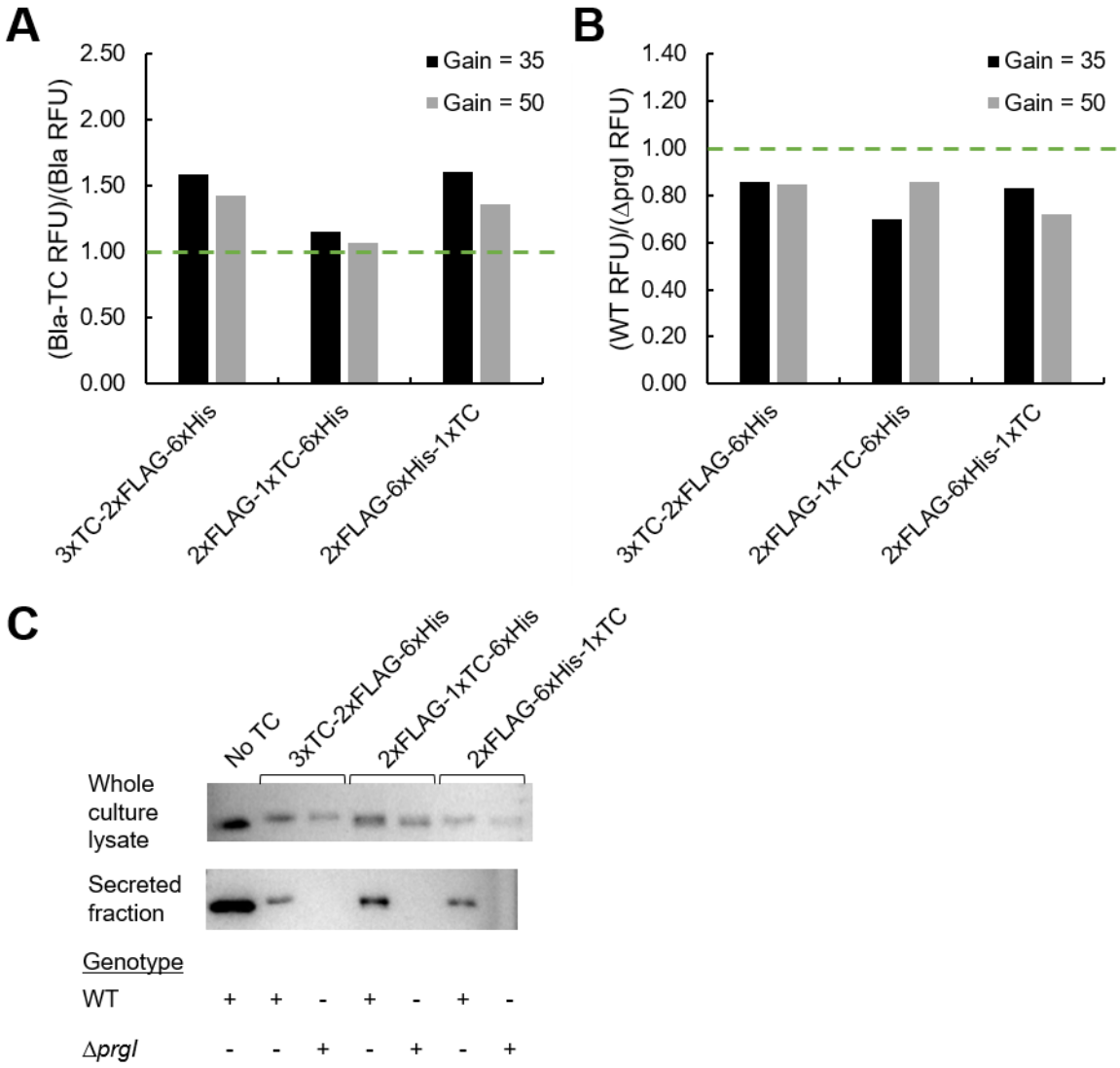


Figure 4.6 Signal-to-noise ratios for various C-terminal TC tag arrangements.
Fluorescence from the various SptP-Bla and TC tag arrangements was compared to SptP-Bla-2xFLAG-6xHis secreted from ASTE13 WT (**A**), and to the secreted fraction from ASTE13 Δ*prgI* strains expressing the same construct (**B**). Secretion was performed for 24 hours in NCE supplemented with 1X EZ and 0.4% w/v succinate, and P_{lacUV5} *hilA* induction facilitated T3SS expression (**C**). The secreted fraction was reduced with fresh 2 mM DTT and labeled with 2 μM FlAsH-EDT_2 for one hour at 37°C. Fluorescence was measured with gains of 35 and 50 on a BioTek Synergy HTX plate reader. Western blots of expressed and secreted protein are in (**C**).

I repeated the secretion experiment with the same constructs and labeled the secreted fraction with ReAsH-EDT_2. Smaller volumes and a shorter labeling time produced larger positive differences in fluorescence between ASTE13 WT and ASTE13 Δ*prgI* strains carrying the same TC tag variant (Figure 4.7). SptP-Bla-2xFLAG-1xTC-6xHis showed the largest signal-to-noise ratio.

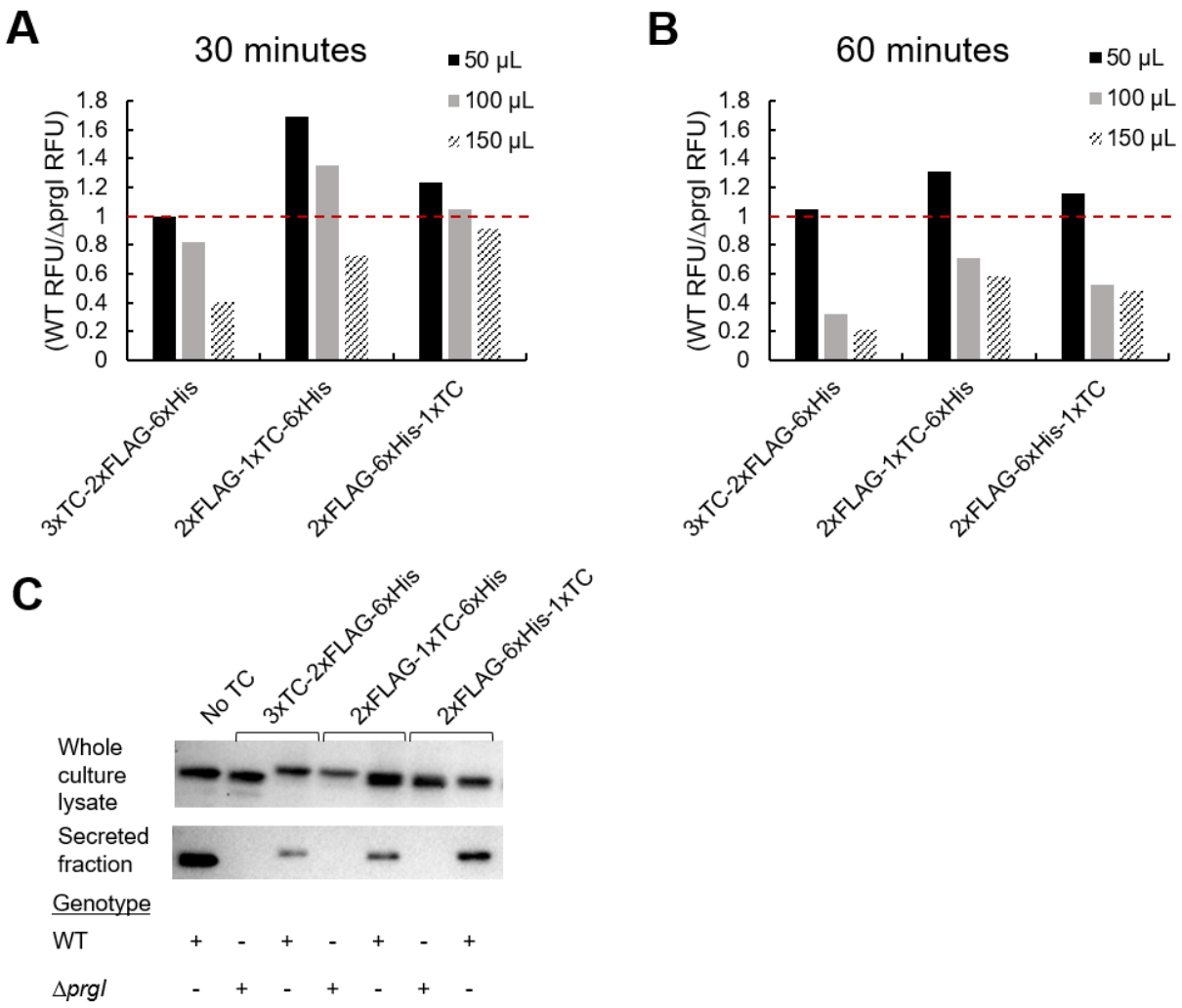


Figure 4.7 Optimization of labeling time and amount of sample for labeling of various TC-tagged constructs with ReAsH.
The ratio of the secreted fractions from ASTE13 WT and ASTE13 Δ*prgI* strains expressing SptP-Bla with the various TC tag arrangements was compared for 30 minutes (**A**) or 60 minutes (**B**) of labeling with 2 µM ReAsH-EDT_2 at 37°C after reduction with 2 mM DTT for 15 minutes. Secretion was performed for 24 hours in NCE supplemented with 1X EZ and 0.4% w/v succinate, and P_{lacUV5} *hilA* induction facilitated T3SS expression. Fluorescence was measured with a gain of 50 on a BioTek Synergy HTX plate reader. Western blots of expressed and secreted protein are in (**C**).

I purified each of the TC-tagged constructs and labeled a dilution series with both FlAsH-EDT_2 and ReAsH-EDT_2 to determine the linear range and limit of detection for the assay (Figure 4.8). SptP-Bla-2xFLAG-1xTC-6xHis was superior to the other two constructs for ReAsH-EDT_2 labeling and was the only construct to show any change in fluorescence when labeled with FlAsH-EDT_2. The assay was linear to 35 mg/L for ReAsH-EDT_2 labeling and 180 mg/L for FlAsH-EDT_2 labeling. The limit of detection was 2 mg/L for ReAsH-EDT_2 and 10 mg/L for FlAsH-EDT_2 under the conditions tested.

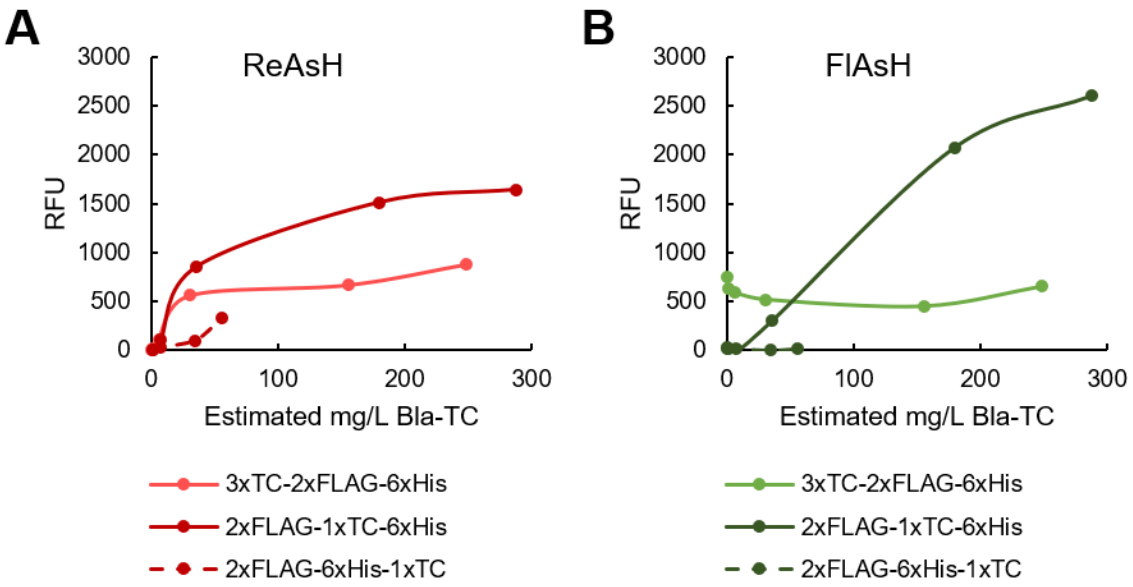


Figure 4.8 Limit of detection and linear range for various TC-tagged constructs labeled with ReAsH or FlAsH.
Each TC-tagged SptP-Bla was purified and diluted with PBS to provide a range of protein concentrations that spanned two orders of magnitude. Each concentration was reduced for 15 minutes with 2 mM DTT and then labeled with 2 µM ReAsH-EDT_2 (**A**) or 2 µM FlAsH-EDT_2 (**B**) for 30 minutes at 37°C. Fluorescence was measured with a gain of 50 (ReAsH) or a gain of 35 (FlAsH) on a BioTek Synergy HTX plate reader.

The excitation and emission maxima for FlAsH-EDT_2 are 508 nm and 528 nm, which is within the 516/20 nm emission filter but just outside the 485/20 nm excitation filter available on the plate reader used for the above experiments. To determine if the signal-to-noise ratio could be increased with a more appropriate filter set, I repeated secretion and labeling for each of the TC-tagged constructs and measured fluorescence at several excitation and emission wavelengths on a BioTek Synergy H1 plate reader with a monochromator. Figure 4.9 shows that altering the excitation and emission spectra gives only a marginal increase in the ratio between fluorescence of the secreted fractions of ASTE13 WT and ASTE13 Δ*prgI* strains carrying the same TC tag variant.

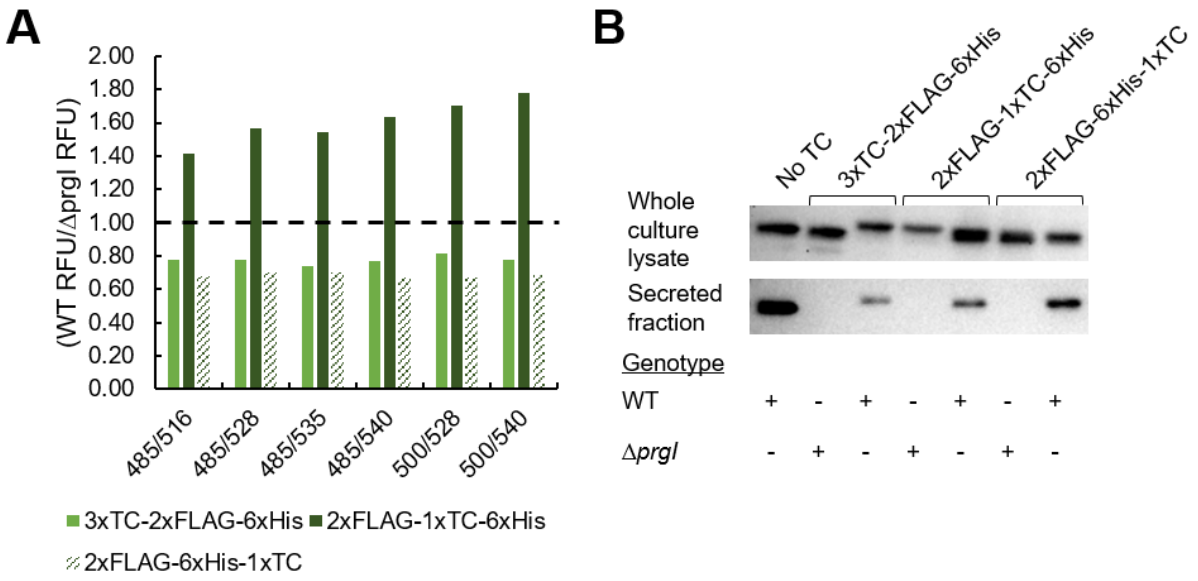


Figure 4.9 Signal-to-noise ratio for secreted TC-tagged constructs with various excitation and emission filters.
The ratio of the secreted fraction from ASTE13 WT and ASTE13 Δ*prgI* strains expressing the various TC tag constructs was calculated for the specified excitation and emission filter settings (**A**). Secretion was performed in NCE supplemented with 1X EZ and 0.4% w/v succinate, and P_{lacUV5} *hilA* induction facilitated T3SS and SptP-Bla expression (**B**). Secreted fractions were reduced for 15 minutes with 2mM DTT and then labeled with 2 μM FlAsH-EDT_2 for 30 minutes at 37°C. Fluorescence was measured with the specified excitation and emission filter settings on a BioTek Synergy H1 plate reader with a monochromator. Each filter setting had a 16 nm bandwidth, and the gain was 50 for all readings.

Though the optimization efforts described above provided a positive signal for SptP-Bla-2xFLAG-1xTC-6xHis relative to the Δ*prgI* background condition, I expected the difference in signal to be several-fold larger, especially in a defined medium. I became suspicious that FlAsH-EDT_2 and ReAsH-EDT_2 were labeling other proteins non-specifically, so I ran purified and secreted SptP-Bla-2xFLAG-6xHis and SptP-Bla-2xFLAG-1xTC-6xHis samples labeled with FlAsH-EDT_2 on SDS-PAGE. FlAsH-EDT_2 and ReAsH-EDT_2 remain bound to the TC motif in Laemmli buffer, so one can follow the usual labeling procedure, mix the sample with Laemmli buffer, and run SDS-PAGE as usual. I used FlAsH-EDT_2 because it could be detected with a UV transilluminator.

The gel analysis revealed that FlAsH-EDT_2 binds non-specifically to several proteins, and strongly to a ~30 kilodalton protein present in the purified samples (Figure 4.10). The purified samples were generated from ASTE13 Δ*prgI* strains, so it is possible that some of the proteins that labeled non-specifically in the purified samples are also present in the secreted fractions. The secreted fractions were not concentrated enough to provide evidence of non-specific binding on that gel, however. The potential for non-specific binding was enough to attempt a different assay—T3SS optimization projects can change the composition of the secreted fraction, and including a control for every new test condition would neutralize the increased throughput provided by this plate-based assay.

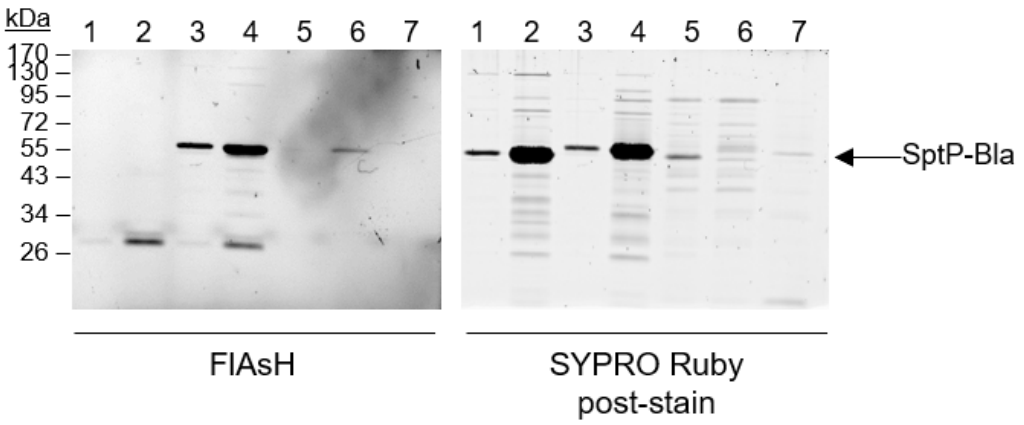


Figure 4.10 FlAsH in-gel labeling and SYPRO Ruby post-stain of purified and secreted SptP-Bla. Purified and secreted SptP-Bla-2xFLAG-6xHis and SptP-Bla-2xFLAG-1xTC-6xHis were labeled with 5 μM FlAsH-EDT_2 and run on SDS-PAGE. Lane labels correspond to the conditions listed in Table 4.4. Following SDS-PAGE, the gel was washed in ddH_2O and imaged using a ChemiDoc XRS+ UV transilluminator to visualize the FlAsH-EDT_2-labeled protein. The gel was post-stained with SYPRO Ruby and imaged again on a ChemiDoc XRS+.

Table 4.4 Strains and purified proteins for FlAsH in-gel labeling.

Lane	Strain	Expressed or Purified Protein	Expected Labeling?
1	Purified	50 ng SptP-Bla-2xFLAG-6xHis	No
2	Purified	500 ng SptP-Bla-2xFLAG-6xHis	No
3	Purified	50 ng SptP-Bla-2xFLAG-1xTC-6xHis	Yes
4	Purified	500 ng SptP-Bla-2xFLAG-1xTC-6xHis	Yes
5	ASTE13 WT	SptP-Bla-2xFLAG-6xHis	No
6	ASTE13 WT	SptP-Bla-2xFLAG-1xTC-6xHis	Yes
7	ASTE13 *Δprgl*	SptP-Bla-2xFLAG-1xTC-6xHis	No

4.3.3 The split GFP assay has low sensitivity and is affected by medium composition

Because the TC tag assay was incompatible with our system in its current format, I turned to split GFP as an alternative option [167]. The split GFP assay involves fusing the eleventh beta strand of superfolder GFP (sfGFP) to a protein of interest and combining the fusion with the remaining $GFP(1\text{-}10)_{opt}$ strands to complete the β-barrel and gain fluorescence. The $GFP(1\text{-}10)_{opt}$ fragment has little to no background fluorescence on its own. I fused the GFP11 fragment to DH and alkaline phosphatase to create the constructs SptP-DH-GFP11-2xFLAG-6xHis and SptP-AP-GFP11-2xFLAG-6xHis.

The established protocol for split GFP recommended at least a fourfold molar excess of $GFP(1\text{-}10)_{opt}$, so I calculated the amount of $GFP(1\text{-}10)_{opt}$ required to maintain a tenfold molar excess. Our highest reported secretion titer in LB-L was on the order of 30 mg/L for SptP-DH-2xFLAG-6xHis, so I used 50 mg/L to provide a cushion. The molar concentration of SptP-DH-GFP11-2xFLAG-6xHis for a secretion titer of 50 mg/L is:

$$\mathrm{MW}_{\mathrm{SptP-DH-GFP11-2xFLAG-6xHis}} = 46771\ \mathrm{Da}\ [=]\ \frac{\mathrm{g}}{\mathrm{mol}}$$

$$\left(50\,\frac{\text{mg secreted protein}}{\text{L}}\right)\cdot\frac{\text{mmol}}{46771\text{mg}} = 1.1\ \mu\text{M secreted protein}$$

$$(1.1\ \mu\text{M secreted protein})\cdot(20\ \mu\text{L to assay well}) = 21\ \text{pmol in assay well}$$

Thus, the concentration of GFP(1-10)$_{opt}$ required to achieve a tenfold molar excess is:

$$\frac{(210\ \text{pmol GFP}(1-10)_{\text{opt}})}{180\ \mu\text{L to assay well}} = 1.2\ \mu\text{M GFP}(1-10)_{\text{opt}}\ \text{working solution}$$

The estimated concentration of the GFP(1-10)$_{opt}$ working solution ranged from 48 µM to 112 µM, which is more than sufficient for a secretion titer of even 500 mg/L according to these calculations.

I purified SptP-AP-GFP11-2xFLAG-6xHis to determine the sensitivity and optimal incubation time for the split GFP assay. Purified SptP-AP-GFP11-2xFLAG-6xHis spanning two orders of magnitude was mixed with GFP(1-10)$_{opt}$ and fluorescence was recorded every 15 minutes for four hours. Concentrations of SptP-AP-GFP11-2xFLAG-6xHis below 20 mg/L failed to show an increase in signal after four hours (Figure 4.11). After an overnight incubation at room temperature, however, all standards showed an increase in fluorescence above background, and those values were used to construct a standard curve. The correlation between fluorescence and protein concentration was linear over the entire concentration range.

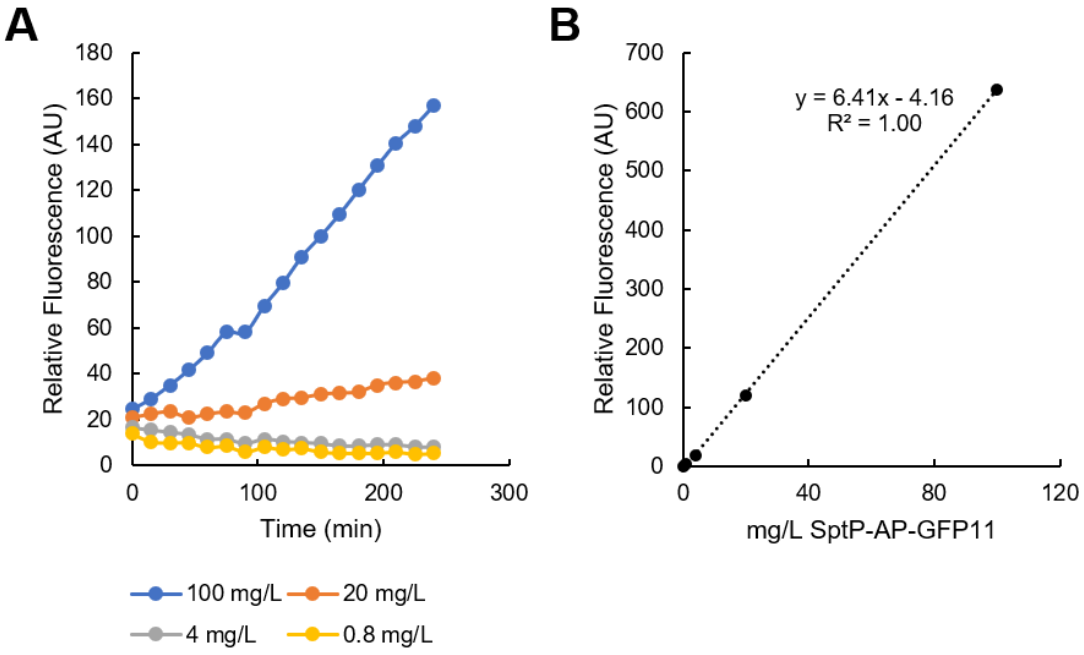


Figure 4.11 Relative fluorescence for purified SptP-AP-GFP11-2xFLAG-6xHis.
Purified SptP-AP-GFP11-2xFLAG-6xHis was mixed with a GFP(1-10) working solution and fluorescence was monitored at room temperature over four hours for several concentrations of purified SptP-AP-GFP11-2xFLAG-6xHis (**A**). Samples were held overnight at room temperature and read again to construct a standard curve (**B**). Fluorescence was recorded using a BioTek Synergy HTX plate reader with gain = 50.

To determine if the assay was sensitive to differences in secretion of SptP-AP-GFP11-2xFLAG-6xHis, I performed secretion in a range of media and subjected the secreted fractions to overnight incubation with the $GFP(1\text{-}10)_{opt}$ working solution. All media tested except LB-L with 0.4% w/v glycerol showed a difference between ASTE13 WT and ASTE13 Δ*prgI* strains, though the difference was small in LB-L with no additives (Figure 4.12). The low sensitivity could be a result of low secretion titer.

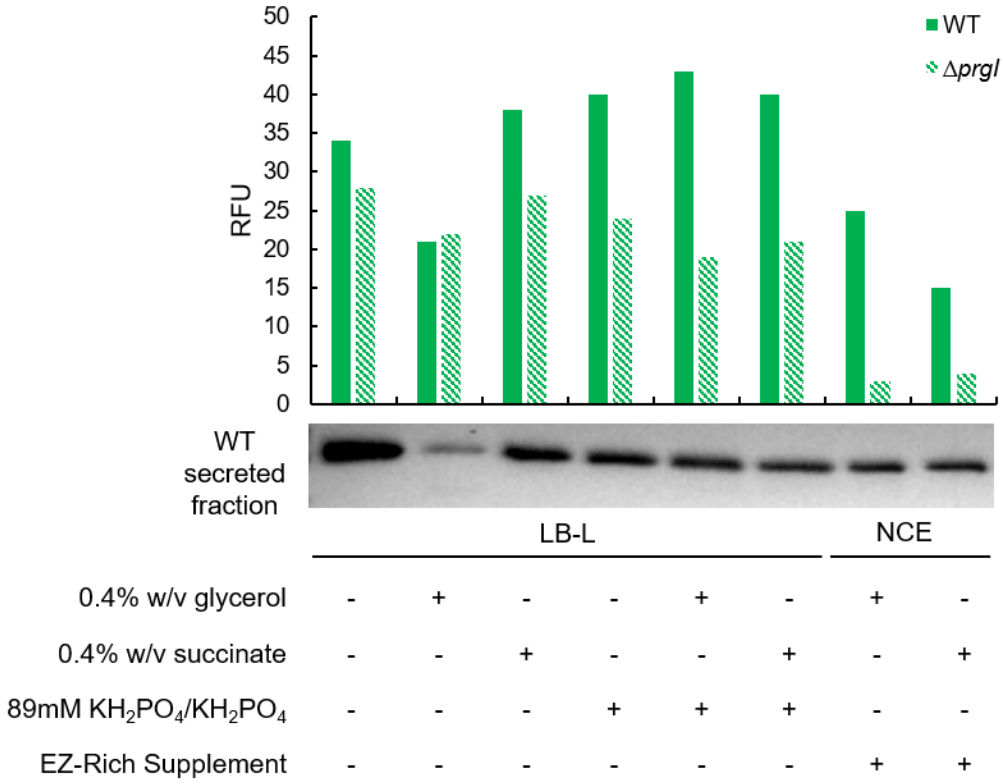


Figure 4.12 Fluorescence of SptP-AP-GFP11-2xFLAG-6xHis secreted in various media. SptP-AP-GFP11-2xFLAG-6xHis was secreted in the specified media for 8 hours. P_{lacUV5} *hilA* induction facilitated T3SS expression in all conditions. Following harvest, secreted fractions were mixed with $GFP(1\text{-}10)_{opt}$ working solution and held at room temperature overnight. Fluorescence was recorded on a BioTek Synergy HTX plate reader with a gain of 50. "RFU" refers to relative fluorescence units, and the data represents a single biological replicate.

SptP-DH secretes at the highest titer of all proteins tested for the T3SS, so I compared it to alkaline phosphatase as a GFP11 fusion partner with the hope that increasing titer would increase the signal-to-noise ratio for the split GFP assay. Purified SptP-DH-GFP11-2xFLAG-6xHis was linear over two orders of magnitude and produced a higher fluorescence signal than SptP-AP-GFP11-2xFLAG-6xHis at the same protein concentration (Figure 4.13). Despite a significant increase in secretion titer, secreted SptP-DH-GFP11-2xFLAG-6xHis offered only a small increase in signal-to-noise ratio compared to secreted SptP-AP-GFP11-2xFLAG-6xHis. Signal-to-noise ratio increased with secretion titer, and relative secretion titer calculated from relative fluorescence units

corresponded well to relative secretion titer calculated from the standard densitometry analysis.

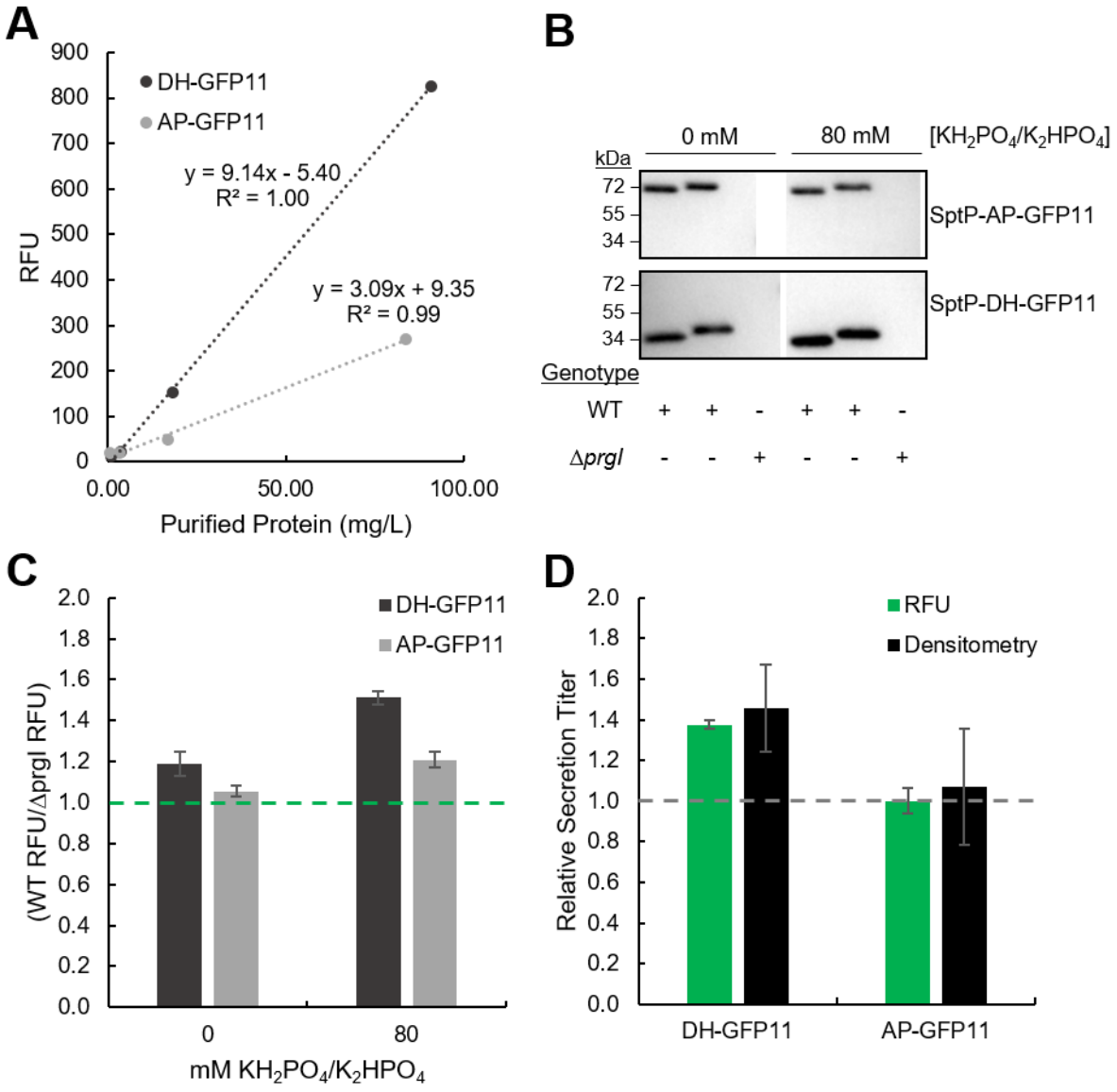
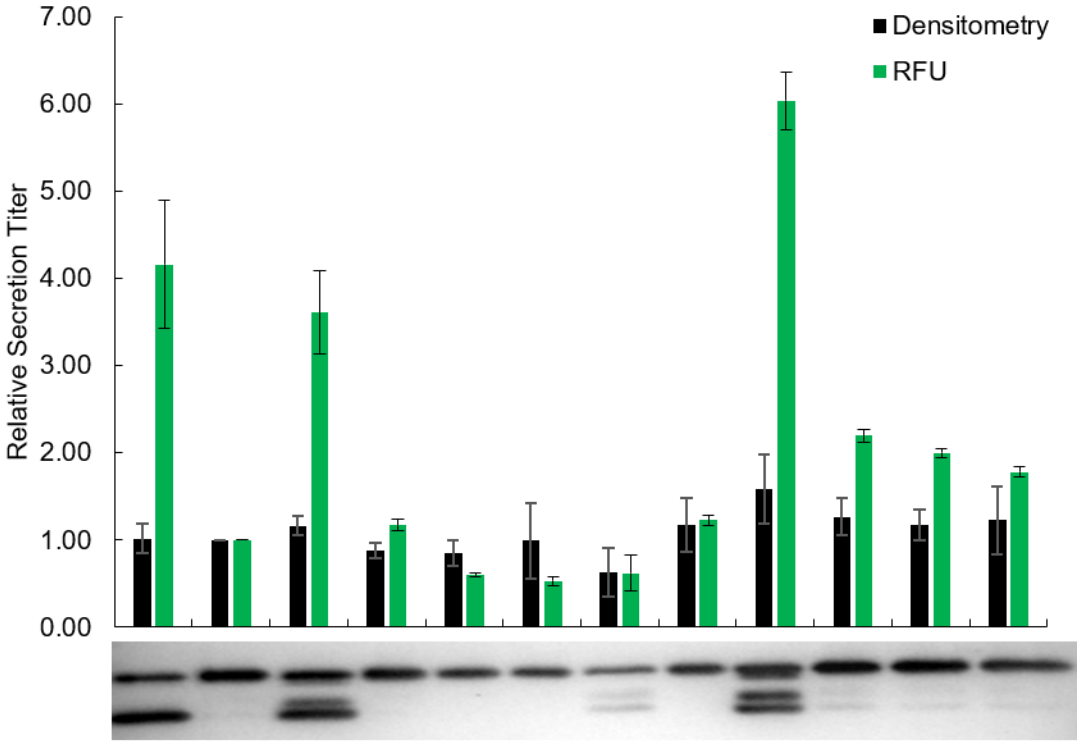


Figure 4.13 Comparison of SptP-AP-GFP11 and SptP-DH-GFP11.
Purified SptP-AP-GFP11-2xFLAG-6xHis and SptP-DH-GFP11-2xFLAG-6xHis diluted over two order of magnitude were mixed with GFP(1-10)$_{opt}$ working solution. Following incubation at room temperature overnight, fluorescence was recorded to construct a standard curve (**A**). Signal-to-noise ratio (**C**) and a comparison between fluorescence and densitometry (**D**) was measured for SptP-AP-GFP11-2xFLAG-6xHis and SptP-DH-GFP11-2xFLAG-6xHis. Proteins were secreted in LB-L and LB-L with 80 mM KH_2PO_4/K_2HPO_4 following P_{lacUV5} *hilA* induction for 8 hours. After harvest, the secreted fraction was mixed with GFP(1-10)$_{opt}$ working solution and held at room temperature overnight. Fluorescence was recorded with a BioTek Synergy HTX plate reader with a gain of 50. Densitometry was performed on an anti-FLAG western blot as described in Methods. For the comparison between fluorescence and densitometry, values were normalized to LB-L without phosphate for each secreted protein. Error bars represent standard deviation on two biological replicates. All bands from the representative western blots in (**B**) are from the same blot but were rearranged for clarity.

The positive signal-to-noise ratio observed with SptP-DH-GFP11-2xFLAG-6xHis and the agreement between relative fluorescence and densitometry were promising results, suggesting that the split GFP assay could increase throughput for secretion titer analysis significantly. Assay development was proceeding in parallel to the defined medium development described in Chapter 3, so I continued to compare relative fluorescence and densitometry as I evaluated the effect of altering medium composition on secretion titer. The proteolytic cleavage observed in the secreted fraction from defined media with glycerol as a carbon source produced substantially different results for relative secretion titer calculated from fluorescence and semi-quantitative western blotting, however. In an experiment comparing the effects of the phosphate counter ion and sodium chloride concentration, secretion titer calculated from relative fluorescence was two- to six-fold higher in samples containing fragments (Figure 4.14). The fragmentation observed in the secreted fraction after 24 hours in the presence of glucose and glycerol would complicate any analysis, but the several-fold difference observed in samples with fragmentation made this assay unsuitable for a medium optimization project.



	NCE						PCN					
0.4% w/v glycerol	+	-	+	-	-	-	+	-	+	-	-	-
0.4% w/v succinate	-	+	-	+	+	+	-	+	-	+	+	+
Na_2HPO_4 (mM)	0	0	0	34	34	34	0	0	0	0	34	34
NaH_2PO_4 (mM)	0	0	0	0	29	29	0	0	0	0	0	29
K_2HPO_4 (mM)	34	34	34	0	0	0	1.3	1.3	34	34	0	0
KH_2PO_4 (mM)	29	29	29	29	0	0	0	0	29	29	29	0
NaCl (mM)	0	0	129	0	0	0	50	50	179	50	50	50
KCl (mM)	0	0	0	0	0	0.3	0	0	0	0	0	0

Figure 4.14 Relative secretion titer measured by fluorescence and densitometry for SptP-DH-GFP11-2xFLAG-6xHis.
Following P_{lacUV5} *hilA* induction, SptP-DH-GFP11-2xFLAG-6xHis was secreted for 24 hours in NCE and PCN with the specified modifications. After harvest, the secreted fraction was mixed with $GFP(1\text{-}10)_{opt}$ working solution and held at room temperature overnight. Fluorescence was recorded with a BioTek Synergy HTX plate reader with a gain of 50. Densitometry was performed on total protein (including fragments) from an anti-FLAG western blot as described in Methods. Fluorescence and densitometry values were normalized to NCE with succinate. Error bars represent standard deviation on two biological replicates.

4.3.4 An alkaline phosphatase activity assay increases throughput of secretion titer analysis

Enzyme activity assays are commonly used in molecular biology, and those described for "model" enzymes are well-characterized, often simple, and usually inexpensive. Activity assays are incompatible with medium optimization projects because they explicitly report on the folded state of the enzymes present in solution. For projects that do not involve changing medium composition and therefore the folding environment, however, enzyme activity is an attractive option. If the enzyme follows Michaelis-Menten kinetics, the initial rate of product formation is proportional to the amount of enzyme if the substrate is in excess of K_M:

$$V_0 = \frac{dP}{dt} = \frac{k_{cat}[E][S]}{[S] + K_M}$$

$$\text{for } [S] \gg K_M,$$

$$\frac{dP}{dt} = k_{cat}[E]$$

Thus, calculating the slope of the initial linear regime of $[P]$ vs. t for the activity of secreted enzymes allows relative comparisons of secreted enzyme titer. We showed that SptP-tagged beta-lactamase and alkaline phosphatase were active after secretion via the T3SS [38]. I chose to further optimize a high-throughput alkaline phosphatase activity assay because beta-lactamase is the resistance gene in the plasmid we use to facilitate genomic modification, and I planned to create a library using that technique.

To find the ideal 96-well plate assay conditions, I created a dilution series of purified SptP-AP-2xFLAG-6xHis and performed activity assays over a range of pNPP concentrations (Figure 4.16). Increasing absorbance at 405 nm represented product formation, and the relationship between absorbance and time was linear over the course of the experiment (Figure 4.15). The lower concentrations of pNPP provided a weak correlation between SptP-AP-2xFLAG-6xHis concentration and initial rate, which is unsurprising given that the K_M calculated for secreted SptP-AP-2xFLAG-6xHis was 100 ± 40 μM [38].

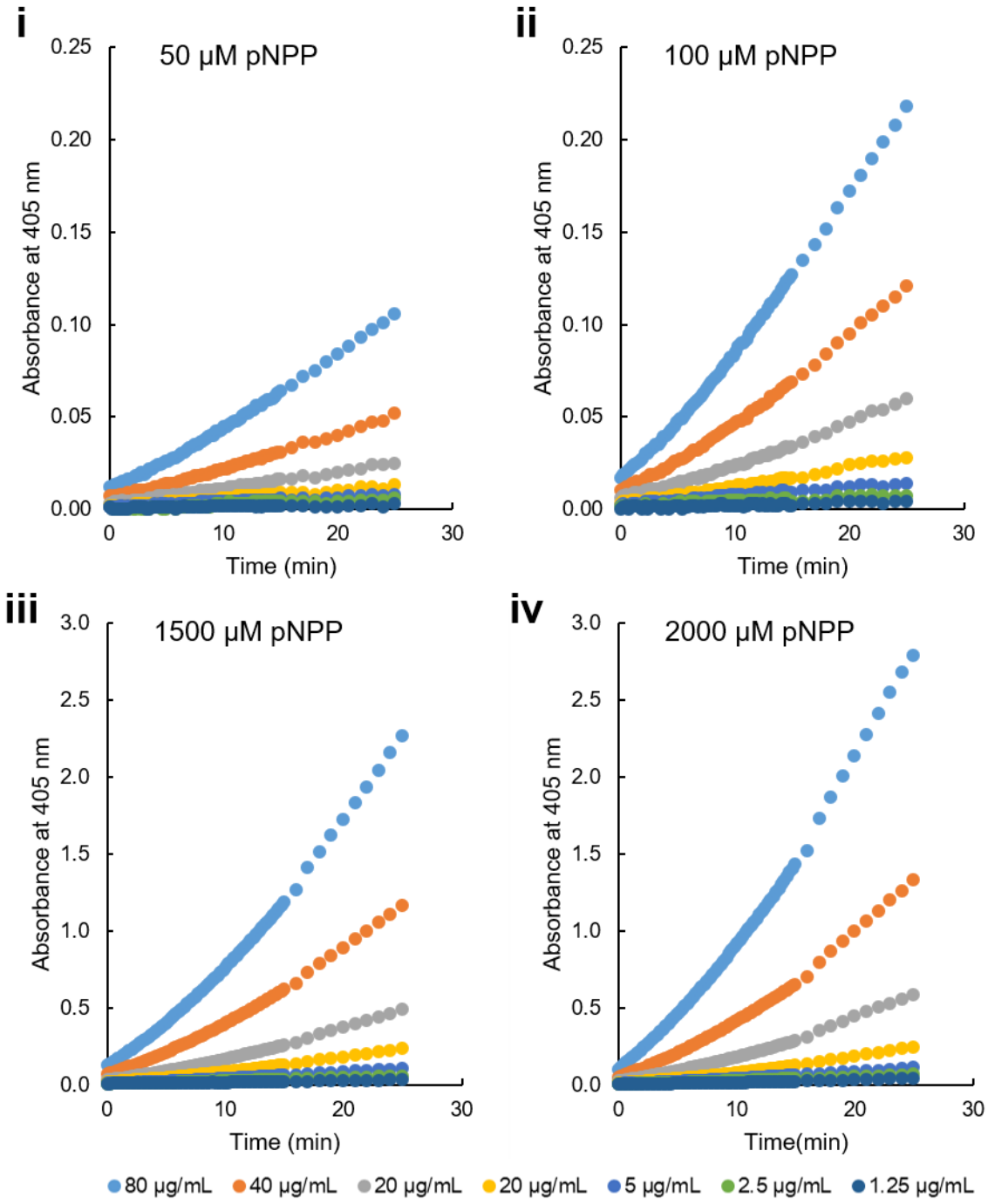


Figure 4.15 Absorbance over time for varying amounts of purified SptP-AP and pNPP.
Purified SptP-AP-2xFLAG-6xHis was diluted in PBS over a range of concentrations and mixed with the specified amount of pNPP. Product formation was monitored by recording absorbance at 405 nm every 15 seconds for 15 minutes and then every minute until 25 minutes on a BioTek Synergy HTX plate reader set to 37°C. Note that the scales for *i,ii* differ from *iii,iv*. Data represents a single assay replicate.

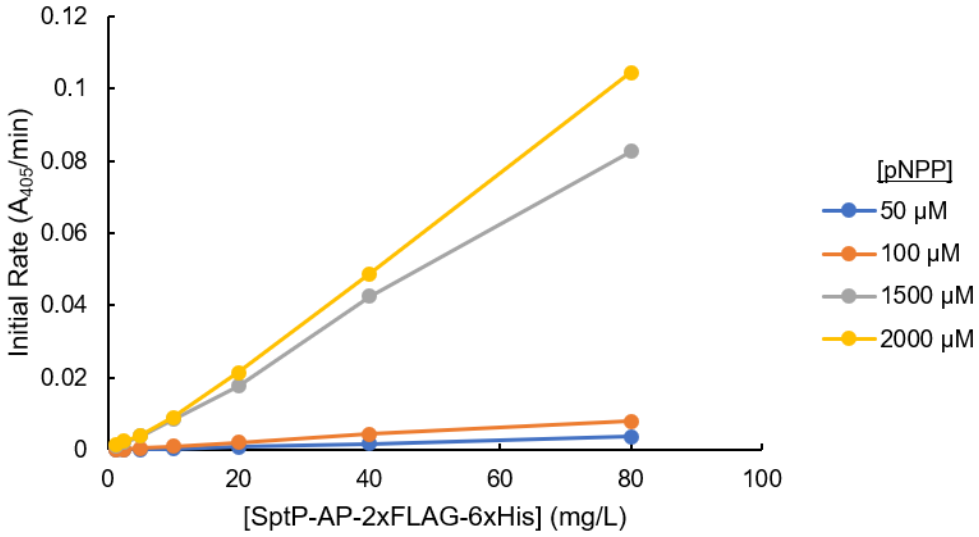


Figure 4.16 Initial rate of SptP-AP activity on pNPP.
Purified SptP-AP-2xFLAG-6xHis was diluted in PBS over a range of concentrations and mixed with the specified amount of pNPP. Product formation was monitored by recording absorbance at 405 nm every 15 seconds for 15 minutes and then every minute until 25 minutes on a BioTek Synergy HTX plate reader set to 37°C. Initial rate is represented by the slope of the line that relates absorbance and time (Figure 4.15). Data represents a single assay replicate.

I expected secretion titer of SptP-AP-2xFLAG-6xHis to be on the order of 1 mg/L, which was at the low end of the range of purified SptP-AP-2xFLAG-6xHis concentrations that I evaluated in Figure 4.16. I tested activity on 1.5 mM and 2 mM pNPP with secreted fractions from ASTE13 WT and ASTE13 Δ*prgI* strains expressing SptP-AP-2xFLAG-6xHis and found that alkaline phosphatase activity was detected in the secreted fraction of ASTE13 WT but not ASTE13 Δ*prgI* (Figure 4.17). The samples with 2 mM pNPP showed a larger difference between the secreted fraction containing SptP-AP-2xFLAG-6xHis (WT) and the secreted fraction lacking SptP-AP-2xFLAG-6xHis (Δ*prgI*), so that condition was selected for future projects.

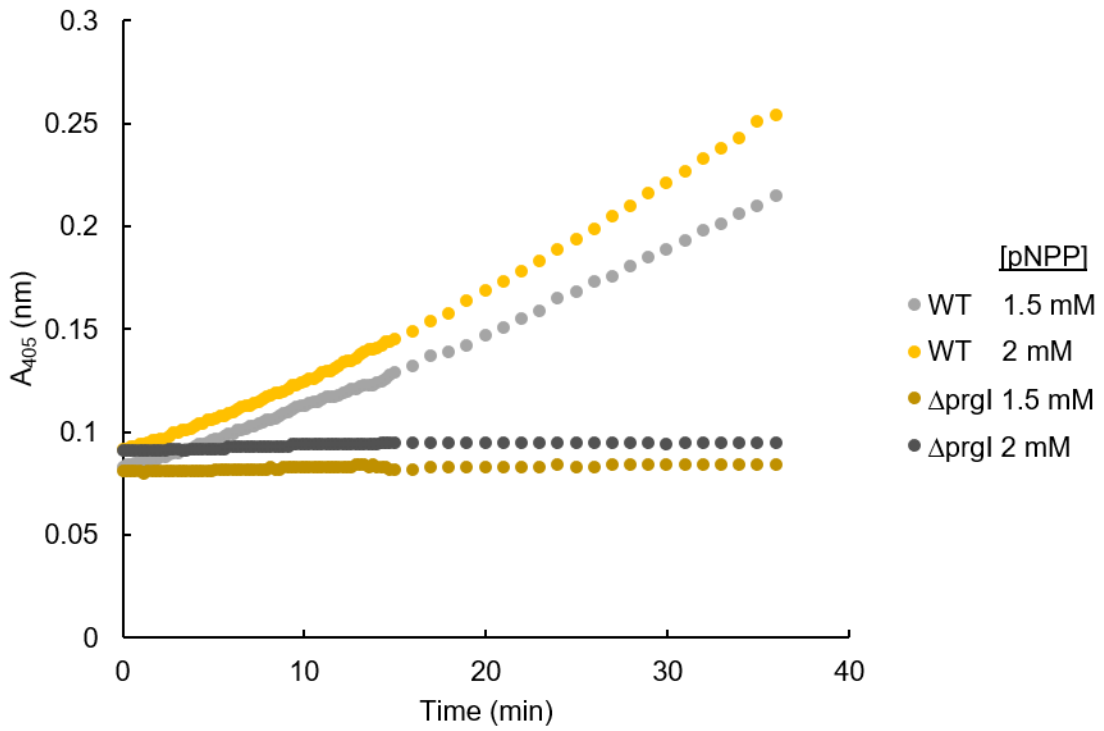


Figure 4.17 Product formation over time for secreted SptP-AP-2xFLAG-6xHis.
SptP-2xFLAG-6xHis was expressed in ASTE13 WT and ASTE13 Δ*prgI* for 24 hours following P_{lacUV5} *hilA* induction in NCE with 1X EZ Supplement and 0.4 % w/v succinate. The secreted fraction was mixed with assay buffer and the specified concentration of pNPP. Absorbance was monitored at 37°C every 15 seconds for the first 15 minutes, and then every minute for the next 20 minutes. Data represents a single biological replicate.

4.4 Discussion

No analysis technique is perfect—the ideal analysis technique matches the needs of the system and the experimental design while producing accurate and precise results. Even for the same system, different experimental goals may require multiple analysis techniques.

The universal applicability of SDS-PAGE with general protein stains and western blotting is a benefit for analyzing multiple proteins with a similar workflow. The linear range must be established for each new protein, however, and samples must be diluted to produce a densitometry signal within at most order of magnitude of the normalization condition, ideally within 30-50%. Using general protein stains such as Coomassie and SYPRO Ruby to estimate secretion titer instead of western blotting simplifies the workflow considerably, but one must compensate for other proteins stained at the same molecular weight by including and subtracting a “background” secreted fraction as described in Section 2.2.6. The discrepancy between secretion titer calculated from quantitative Coomassie gels and quantitative western blots should be resolved with an orthogonal analysis technique.

The TC tag was a promising technique for increasing secretion titer estimation in multiple solution conditions, but the non-specific binding described here and the probably

related low signal-to-noise ratio decreased its attractiveness for the goals of this work. It also requires a reducing environment, which would invalidate its use in combination with an analysis that required disulfide bonds, such as an alkaline phosphatase activity assay. The assay is quite simple, however, so the technology is worth revisiting for future projects with high secretion titers and unchanged growth media.

Like the TC tag, the split GFP assay showed promise for increasing throughput of relative secretion titer analysis, but the results of this chapter show that signal can vary depending on the size of the protein fused to the GFP11 fragment. The stability of the β-barrel fold reduces the dependence of this assay on solution conditions, but the output fluorescence still requires a "refolding" event in the extracellular space that may be affected by variation in media formulations. A significant drawback of the split GFP assay is that the $GFP(1\text{-}10)_{opt}$ detector fragment must be purified from inclusion bodies, and refolding must occur immediately before use. These features suggest that a calibration curve should be included in each experiment, which decreases overall throughput.

Enzyme assays are more sensitive than the TC tag or split GFP assay, making them attractive options for increasing throughput if the medium formulation remains constant. They require a kinetic measurement of product formation over time, which is a more complicated analysis than an endpoint fluorescence measurement, but the SptP-AP assay was the most reliable assay tested here for engineering projects that do not involve changing medium formulations.

The work described here expands the suite of analysis techniques available for estimating secretion titer. A reporter or analysis technique that combines the ease of endpoint fluorescence measurement and constant performance in multiple solution conditions across a broad range of secretion titers remains elusive, however. Secretion titers of many heterologous proteins are too low for most assays even with the medium improvements described in Chapters 2 and 3. As engineering efforts continue to increase titer and purity of the secreted protein, general analysis techniques such as HPLC will become more feasible options and accelerate further optimization of the T3SS as a protein production platform.

4.5 Acknowledgments

I would like to thank Leah Keiser for her work cloning SptP-AP-GFP11-2xFLAG-6xHis and starting the development of the split GFP assay. I would also like to thank the Jewett lab for the use of their BioTek H1 plate reader.

Chapter 5: Engineering the T3SS Needle Protein, PrgI, for Increased Secretion Titer

Portions of the following are in preparation (Burdette, Metcalf, Valdivia, Wong, Kennedy, Tullman-Ercek)

5.1 Introduction

Previous efforts to engineer the T3SS for heterologous protein production have focused on engineering the regulatory system and investigating the effect of the secretion signal [8,41,50]. Two recent studies showed that removing proteins from the tip complex of the T3SS apparatus increased secretion titer, suggesting that the structure of the T3SS can also be optimized for heterologous protein secretion [52,174]. The SPI-1 T3SS apparatus is a ~4 MDa protein complex that spans the inner and outer membrane of the cell [117,136,175]. A basal body anchors the apparatus in the inner and outer membrane, while a polymeric needle capped with a tip complex extends ~60 nm into the extracellular space. The SPI-1 T3SS needle is composed of ~100 copies of a single 80-amino acid protein, PrgI (Figure 5.1). Proteins secreted via the T3SS travel through the basal body and the polymeric needle in an unfolded state and spontaneously refold in the extracellular space [38,176].

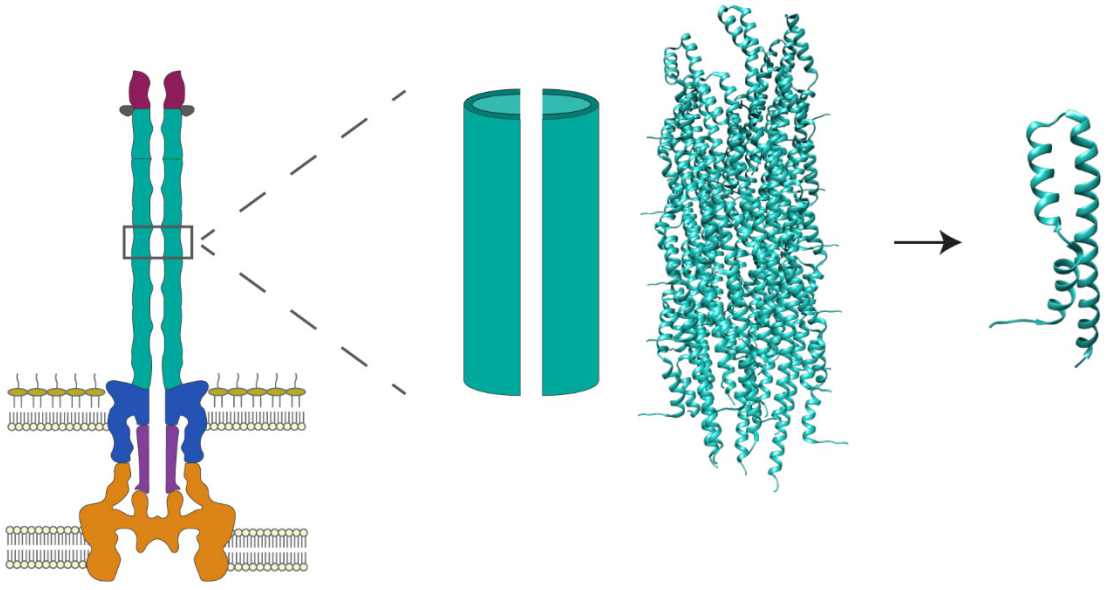


Figure 5.1 Cartoon of the SPI-1 T3SS needle complex highlighting PrgI.
The needle apparatus cartoon is adapted from Marlovits *et al.* [177]. The needle and monomer structures are PDB = 6dwb from Chimera [178].

Evidence suggests that the T3SS exists in active and passive secretion states. In its native context, the T3SS is activated by docking to a host cell membrane [134,179]. We and others have shown that the T3SS is active in the absence of host cells, however [27,41,163,180]. The rate of secretion via the SPI-1 T3SS in the active, docked

conformation exceeds the rate of secretion in the absence of host cells (i.e. the passive state) by tenfold [134,181]. Thus, secretion is proceeding sub-optimally in the absence of host cells even with the engineering improvements discussed in the preceding chapters.

A popular hypothesis is that secretion is activated by a conformational change in the T3SS apparatus upon docking to a host cell [133,136,182]. In this model, the tip complex inserts into the host cell membrane and initiates a structural change that acts as a signal transduction to the basal body. The T3SS needle would likely be affected by this structural change and thus might achieve an “activated” conformation that facilitates rapid protein secretion. Mutational studies in various T3SS needle proteins provide support for this theory—single amino acid changes in the needle protein had a significant effect on controlling secretion of native substrates [183–186].

Thus far, mutational studies of the T3SS needle protein have been alanine scans. A recently developed protein engineering strategy, SyMAPS, coupled comprehensive codon mutagenesis (CCM) and high throughput sequencing to comprehensively describe the mutational tolerance of the MS2 viral capsid [187]. SyMAPS presented a workflow that evaluated every amino acid substitution at every residue to provide a complete view of the relationship between mutation and function at a single residue in a multimeric protein structure. We endeavored to adapt the SyMAPS workflow to identify single amino acid changes in PrgI that confer increased secretion titer and to generate a complete view of the mutational tolerance of a secretion-competent PrgI needle.

This chapter outlines the significant optimization required to adapt the SyMAPS workflow to evaluate the “secretion fitness” of a CCM library of PrgI and reports initial results from the high throughput sequencing data analysis. I describe development of secretion and assay conditions for the use of a genomic copy of SptP-AP-2xFLAG-6xHis for library screening, a recombineering method to create a genomically integrated CCM library from a mixture of gene blocks, and a method of scoring variants to construct a secretion fitness landscape according to the secretion titers measured in the library screen. A surprising 2% of screened library members increased secretion titer more than twofold. A majority of substitutions that increased secretion titer were larger and more hydrophobic amino acids, including many at residues that face the interior of the needle and potentially interact with the secreted protein.

5.2 Methods

5.2.1 Strains and Growth Conditions for Secretion Experiments

Strains, plasmids, and primers used are listed in Table 5.1, Table 5.2, and Table 5.3. Secretion experiments were started by growing a single colony in the lysogeny broth Lennox formulation (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) with appropriate antibiotics (34 µg/mL chloramphenicol for P_{sic} vectors, 50 µg/mL kanamycin for P_{lacUV5} *hilA*) for 12-16 hours overnight in an orbital shaker at 37°C and 225 rpm unless otherwise specified. Overnight cultures were diluted 1:100 into the appropriate medium supplemented with appropriate antibiotics and 100 µg/mL isopropyl β-D-1-thiogalactopyranoside (IPTG) if the strain carried P_{lacUV5} *hilA*. All culturing steps were performed in 24-well deepwell plates (Axygen) unless otherwise specified. Secretion was performed at 37°C and 225 rpm in an orbital shaker for eight hours. The secreted fraction was harvested by centrifuging cultures at 4000 *x g* for 10 minutes. SDS-PAGE samples for the secretion fraction were prepared by adding supernatant to Laemmli buffer in a 3:1 ratio; SDS-PAGE samples for whole culture lysate were prepared by adding cell suspension to Laemmli buffer in a 1:2 ratio. All SDS-PAGE samples were boiled at 95°C for 5 minutes immediately after preparation.

Table 5.1 Strains used in Chapter 5.

Strain	Genotype	Reference
ASTE13	LT2-derived lab strain similar to DW01	This study; DW01 [65]
ASTE13 Δ*prgI*	Δ*prgI*	[41]
ASTE13 *prgI::catsacB*	*prgI::catsacB*	[181]
DTE509	ASTE13 *prgI::prgI*^{L9A}	[181]
DTE510	ASTE13 *prgI::prgI*^{Q48A}	[181]
DTE511	ASTE13 *prgI::prgI*^{Y54A}	[181]
DTE512	ASTE13 *prgI::prgI*^{D70A}	[181]
DTE513	ASTE13 *prgI::prgI*^{P41A}	[181]
DTE514	ASTE13 *prgI::prgI*^{P41C}	This study
DTE515	ASTE13 *prgI::prgI*^{P41D}	This study
DTE516	ASTE13 *prgI::prgI*^{P41E}	This study
DTE517	ASTE13 *prgI::prgI*^{P41F}	This study
DTE518	ASTE13 *prgI::prgI*^{P41G}	This study
DTE519	ASTE13 *prgI::prgI*^{P41H}	This study
DTE520	ASTE13 *prgI::prgI*^{P41I}	This study
DTE521	ASTE13 *prgI::prgI*^{P41K}	This study
DTE522	ASTE13 *prgI::prgI*^{P41L}	This study
DTE523	ASTE13 *prgI::prgI*^{P41M}	This study
DTE524	ASTE13 *prgI::prgI*^{P41N}	This study
DTE525	ASTE13 *prgI::prgI*^{P41Q}	This study

Strain	Genotype	Reference
DTE526	ASTE13 *prgI::prgI*^{P41R}	This study
DTE527	ASTE13 *prgI::prgI*^{P41S}	This study
DTE528	ASTE13 *prgI::prgI*^{P41T}	This study
DTE529	ASTE13 *prgI::prgI*^{P41V}	This study
DTE530	ASTE13 *prgI::prgI*^{P41W}	This study
DTE531	ASTE13 *prgI::prgI*^{P41Y}	This study
sLAB190	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis*	This study
sLAB191	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::catsacB*	This study
sLAB192	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41A}	This study
sLAB193	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41C}	This study
sLAB194	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41M}	This study
sLAB195	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41T}	This study
sLAB196	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41V}	This study
sLAB203	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41D}	This study
sLAB204	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41E}	This study
sLAB205	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41F}	This study
sLAB206	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41G}	This study
sLAB207	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41H}	This study
sLAB208	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41I}	This study
sLAB209	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41K}	This study
sLAB210	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41L}	This study
sLAB211	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41N}	This study
sLAB212	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41Q}	This study
sLAB213	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41R}	This study
sLAB214	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41S}	This study
sLAB215	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41W}	This study
sLAB216	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::prgI*^{P41Y}	This study
sLAB305	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis* ∆*invA*	This study

Table 5.2 Plasmids used in Chapter 5.

Plasmid Name	ORFs under inducible control		ORI	ab^{R}	Reference
P$_{sic}$ *DH*	*sicP*	*sptP-DH-2xFLAG-6xHis*	colE1	cam	[41]
P$_{sic}$ *AP*	*sicP*	*sptP-phoA-2xFLAG-6xHis*	colE1	cam	[38]
P$_{lacUV5}$ *hilA*	*hilA*		p15a	kan	[41]
pSIM6	*gam, beta, exo*		pSC101	cb	[152]
P$_{lacUV5}$ *P41A*	*prgI*^{P41A}		colE1	kan	[181]

Plasmid Name	ORFs under inducible control	ORI	ab^R	Reference
P_{lacUV5} *P41C*	*prgI*^{P41C}	colE1	kan	This study
P_{lacUV5} *P41D*	*prgI*^{P41D}	colE1	kan	This study
P_{lacUV5} *P41E*	*prgI*^{P41E}	colE1	kan	This study
P_{lacUV5} *P41F*	*prgI*^{P41F}	colE1	kan	This study
P_{lacUV5} *P41G*	*prgI*^{P41G}	colE1	kan	This study
P_{lacUV5} *P41H*	*prgI*^{P41H}	colE1	kan	This study
P_{lacUV5} *P41I*	*prgI*^{P41I}	colE1	kan	This study
P_{lacUV5} *P41K*	*prgI*^{P41K}	colE1	kan	This study
P_{lacUV5} *P41L*	*prgI*^{P41L}	colE1	kan	This study
P_{lacUV5} *P41M*	*prgI*^{P41M}	colE1	kan	This study
P_{lacUV5} *P41N*	*prgI*^{P41N}	colE1	kan	This study
P_{lacUV5} *P41Q*	*prgI*^{P41Q}	colE1	kan	This study
P_{lacUV5} *P41R*	*prgI*^{P41R}	colE1	kan	This study
P_{lacUV5} *P41S*	*prgI*^{P41S}	colE1	kan	This study
P_{lacUV5} *P41T*	*prgI*^{P41T}	colE1	kan	This study
P_{lacUV5} *P41V*	*prgI*^{P41V}	colE1	kan	This study
P_{lacUV5} *P41W*	*prgI*^{P41W}	colE1	kan	This study
P_{lacUV5} *P41Y*	*prgI*^{P41Y}	colE1	kan	This study

5.2.2 PCR and Cloning

Primers used in this study are listed in Table 5.3. PCR was performed with Phusion DNA polymerase for Quikchanges and constructing parts for recombineering. Saturation mutagenesis at position 41 was performed by introducing mutations to *prgI* carried on a P_{lacUV5}-inducible plasmid with a Quikchange protocol. Mutations were confirmed by Sanger sequencing. Double-stranded DNA fragments for recombineering contained the replacement gene(s) flanked 5' and 3' by 40 bp of homology to the genetic locus at which the replacement gene(s) should be inserted. The 40 bp of homology was included in oligos and attached via PCR using the primers listed in Table 5.3. The *cat-sacB* cassette was amplified from the purified genome of *E. coli* TUC01, PrgI position 41 mutants were amplified from the appropriate P_{lacUV5}-inducible plasmid, and *sptP-phoA-2xFLAG-6xHis* was amplified from a P_{sic} *sicP sptP-phoA-2xFLAG-6xHis* secretion plasmid. Colony PCR was performed by diluting a colony in a 50 µL PCR reaction containing the appropriate primers and amplifying with GoTaq polymerase. Correct sequences were confirmed by Sanger sequencing.

Table 5.3 Primers used in Chapter 5.

Sequence	Amplicon	Used to Construct
GCAGCAAAACCCTCCGATTGTGCGCTACTGGCGGCGTATC	*prgI*^{P41C} QC	P_{lacUV5} *P41C*
GATACGCCGCCAGTAGCGCACAATCGGAGGGTTTTGCTGC	*prgI*^{P41C} QC	P_{lacUV5} *P41C*
GCAGCAAAACCCTCCGATGATCGCTACTGGCGGCGTATC	*prgI*^{P41D} QC	P_{lacUV5} *P41D*
GATACGCCGCCAGTAGCGCATCATCGGAGGGTTTTGCTGC	*prgI*^{P41D} QC	P_{lacUV5} *P41D*
GCAGCAAAACCCTCCGATGAAGCGCTACTGGCGGCGTATC	*prgI*^{P41E} QC	P_{lacUV5} *P41E*
GATACGCCGCCAGTAGCGCTTCATCGGAGGGTTTTGCTGC	*prgI*^{P41E} QC	P_{lacUV5} *P41E*
CAAAACCCTCCGATTTTGCGCTACTGGCGGCG	*prgI*^{P41F} QC	P_{lacUV5} *P41F*
CGCCGCCAGTAGCGCAAAATCGGAGGGTTTTG	*prgI*^{P41F} QC	P_{lacUV5} *P41F*
CAAAACCCTCCGATGGAGCGCTACTGGCGGC	*prgI*^{P41G} QC	P_{lacUV5} *P41G*
GCCGCCAGTAGCGCTCCATCGGAGGGTTTTG	*prgI*^{P41G} QC	P_{lacUV5} *P41G*
CAGCAAAACCCTCCGATCATGCGCTACTGGCGGCGTATC	*prgI*^{P41H} QC	P_{lacUV5} *P41H*
GATACGCCGCCAGTAGCGCATGATCGGAGGGTTTTGCTG	*prgI*^{P41H} QC	P_{lacUV5} *P41H*
CAAAACCCTCCGATATAGCGCTACTGGCGGCG	*prgI*^{P41I} QC	P_{lacUV5} *P41I*
CGCCGCCAGTAGCGCTATATCGGAGGGTTTTG	*prgI*^{P41I} QC	P_{lacUV5} *P41I*
CAAAACCCTCCGATAAGGCGCTACTGGCG	*prgI*^{P41K} QC	P_{lacUV5} *P41K*
CGCCAGTAGCGCCTTATCGGAGGGTTTTG	*prgI*^{P41K} QC	P_{lacUV5} *P41K*
CAAAACCCTCCGATTTGGCGCTACTGGCG	*prgI*^{P41L} QC	P_{lacUV5} *P41L*
CGCCAGTAGCGCCAAATCGGAGGGTTTTG	*prgI*^{P41L} QC	P_{lacUV5} *P41L*
CAAAACCCTCCGATATGGCGCTACTGGCGGCG	*prgI*^{P41M} QC	P_{lacUV5} *P41M*
CGCCGCCAGTAGCGCCATATCGGAGGGTTTTG	*prgI*^{P41M} QC	P_{lacUV5} *P41M*
CAAAACCCTCCGATAATGCGCTACTGGCGGCG	*prgI*^{P41N} QC	P_{lacUV5} *P41N*

Sequence	Amplicon	Used to Construct
CGCCGCCAGTAGCGCATTATCGGAGG GTTTTG	*prgI*^{P41N} QC	P$_{lacUV5}$ *P41N*
CAGCAAAACCCTCCGATCAAGCGCTAC TGGCGGCGTATC	*prgI*^{P41Q} QC	P$_{lacUV5}$ *P41Q*
GATACGCCGCCAGTAGCGCTTGATCGG AGGGTTTTGCTG	*prgI*^{P41Q} QC	P$_{lacUV5}$ *P41Q*
GCAGCAAAACCCTCCGATAGAGCGCTA CTGGCGGCGTATC	*prgI*^{P41R} QC	P$_{lacUV5}$ *P41R*
GATACGCCGCCAGTAGCGCTCTATCGG AGGGTTTTGCTGC	*prgI*^{P41R} QC	P$_{lacUV5}$ *P41R*
GCAGCAAAACCCTCCGATAGTGCGCTA CTGGCGGCGTATC	*prgI*^{P41S} QC	P$_{lacUV5}$ *P41S*
GATACGCCGCCAGTAGCGCACTATCGG AGGGTTTTGCTGC	*prgI*^{P41S} QC	P$_{lacUV5}$ *P41S*
GCAGCAAAACCCTCCGATACAGCGCTA CTGGCGGCGTATC	*prgI*^{P41T} QC	P$_{lacUV5}$ *P41T*
GATACGCCGCCAGTAGCGCTGTATCGG AGGGTTTTGCTGC	*prgI*^{P41T} QC	P$_{lacUV5}$ *P41T*
GCAGCAAAACCCTCCGATGTAGCGCTA CTGGCGGCGTATC	*prgI*^{P41V} QC	P$_{lacUV5}$ *P41V*
GATACGCCGCCAGTAGCGCTACATCGG AGGGTTTTGCTGC	*prgI*^{P41V} QC	P$_{lacUV5}$ *P41V*
CAAAACCCTCCGATTGGGCGCTACTGG CGGCG	*prgI*^{P41W} QC	P$_{lacUV5}$ *P41W*
CGCCGCCAGTAGCGCCCAATCGGAGG GTTTTG	*prgI*^{P41W} QC	P$_{lacUV5}$ *P41W*
CAAAACCCTCCGATTATGCGCTACTGG CGGCG	*prgI*^{P41Y} QC	P$_{lacUV5}$ *P41Y*
CGCCGCCAGTAGCGCATAATCGGAGG GTTTTG	*prgI*^{P41Y} QC	P$_{lacUV5}$ *P41Y*
AACATACTGCAGGAATATGCTAAAGTAT GAGGAGAGAAAA tgtgacggaagatcacttcg	*cat-sacB*	ASTE13 *sptP::catsacB*
GCTTACTTTCAGATAGTTCTAAAAGTAA GCTATGTTTTTA atcaaagggaaaactgtccatat	*cat-sacB*	ASTE13 *sptP::catsacB*
CTTGAGTCATTTGTGAATCAGCAGGAA GCGCTCAAAAACATACTGCAGGAATAT GCTAAAGTATGAGGAGAGAAAA ttgaataatttaacgttgtcttcg	*sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis*	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis*
ACTTTCTATCGCGGCAAACAAATAATTA TACAGAAATAGCTTACTTTCAGATAGTT CTAAAAGTAAGCTATGTTTTTA ttagtggtgatggtgatgatgc	*sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis*	ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis*
CCCAAGCCCACTTTAATTTAACGTAAAT AAGGAAGTCATT atggcaacaccttggtcagg	*prgI*	All ASTE13 *prgI* variants

Sequence	Amplicon	Used to Construct
GGACAATAGTTGCAATCGACATAATCC ACCTTATAACTGA ttaacggaagttctgaataatggc	*prgI*	All ASTE13 *prgI* variants
CTATAGTGCTGCTTTCTCTACTTAACAG TGCTCGTTTACG tgtgacggaagatcacttcg	*cat-sacB*	ASTE13 *sptP::sptP$^{(1-159)}$-phoA-2xFLAG-6xHis* Δ*invA*
GCCCTTATATTGTTTTTATAACATTCACT GACTTGCTAT atcaaagggaaaactgtccat	*cat-sacB*	ASTE13 *sptP::sptP$^{(1-159)}$-phoA-2xFLAG-6xHis* Δ*invA*
TTATATTGTTTTTATAACATTCACTGACT TGCTATCGTAAACGAGCACTGTTAAGTA GAGAAAGCAGCAC	N/A	ASTE13 *sptP::sptP$^{(1-159)}$-phoA-2xFLAG-6xHis* Δ*invA*
TCGTCGGCAGCGTCAGATGTGTATAAG AGACAG atggcaacaccttggtcag	*prgI*	PrgI library for NGS – PCR step 1
GTCTCGTGGGCTCGGAGATGTGTATAA GAGACAG ttaacggaagttctgaataatggc	*prgI*	PrgI library for NGS – PCR step 1

5.2.3 Strain Construction for Single Genomic Modifications

Strain modifications were generated by λ Red recombineering as described by Thomason *et al.* [152]. Briefly, a colony of ASTE13 carrying the pSIM6 plasmid was inoculated in LB-L with 30 µg/mL carbenicillin and grown at 30°C and 225 rpm for 16-20 hours. The overnight culture was diluted 1:70 into 35 mL of LB-L and grown at 30°C until OD_{600} reach 0.4-0.6. The culture was washed twice with 30 mL ice-cold sterile ddH_2O and centrifugation at 4600 x g for 3 minutes to collect the cells. After the second wash, cells were resuspended in ~400 µL of ice-cold sterile ddH_2O. Aliquots of 50 µL resuspended cells were mixed with 200 ng of the appropriate PCR fragment and electroporated at 1800V for 5 milliseconds. A negative control containing no added DNA was also electroporated. Cells were mixed with 950 µL SOC medium immediately after electroporation and either recovered at 30°C for an hour for *cat-sacB* cassette introduction (first step of recombineering) or transferred to a test tube containing 9 mL of LB-L and grown at 37°C, 225 rpm for four hours for *cat-sacB* removal and replacement (second step of recombineering). Cells were diluted serially to 10^{-3} in sterile PBS. 200 µL of diluted cells was plated on 6% sucrose agar and grown at 37°C overnight. The second step of recombineering for the Δ*invA* knockout replaced the *cat-sacB* cassette by electroporating 1 µL of a 10 µM solution of a single 60 bp oligo containing the first and last 30 bp of the *invA* gene.

5.2.4 Library Construction

A library of gene blocks carrying all possible amino acid substitutions was synthesized and pooled by Twist Biosciences. Codons were fully randomized (“NNN”), but the library excluded wild-type residues and stop codons. Residues 1-6 and 76-80 were not modified. The lyophilized DNA from Twist Biosciences was reconstituted in ultrafiltered water to a concentration of 200 ng/µL. Recombineering was performed with an ASTE13 *sptP::sptP$^{(1-159)}$-phoA-2xFLAG-6xHis prgI::cat-sacB* as described above with the following modifications. 200 ng (4 µL of a 50 ng/µL resuspended solution) of the library was transformed into 100 µL of recombination-competent cells via electroporation at

1800V and 5 milliseconds. A negative control containing no added DNA was also electroporated. Cells were immediately mixed with 900 µL SOC medium and transferred to a 14 mL disposable culture tube (Fisherbrand) containing 2 mL of LB-L for a four-hour recovery at 37°C and 225 rpm. Recombination efficiency was assessed by plating 200 µL of cells diluted serially to 10^{-3} in sterile PBS from both the library and the negative control on 6% sucrose agar and allowing colonies to develop at room temperature for 24 hours. The remainder of the culture was mixed with 60% glycerol in a 1:3 ratio and aliquoted into three cryovials for storage at -80°C. Before storage, 2 x 33 µL aliquots of the glycerol mixture were diluted to facilitate plating single colonies for screening. The first aliquot was diluted in 1.2 mL sterile PBS, and the second aliquot was diluted in 1.2 mL PBS with 15% glycerol and frozen at -80°C. The 1.2 mL aliquot without glycerol was further split into 3 x 400 µL aliquots, and each was plated on a 15 cm agar plate with 6% sucrose LB-agar. Colonies developed for 24 hours at room temperature.

5.2.5 Library Screening

Single colonies were inoculated in 0.5 mL LB-L in a 2 mL square 96-well deepwell plate (Axygen) and grown overnight at 37°C, 350 rpm. ASTE13 *sptP::sptP$^{(1-159)}$-phoA-2xFLAG-6xHis*, ASTE13 *sptP::sptP$^{(1-159)}$-phoA-2xFLAG-6xHis prgI::catsacB*, and ASTE13 *sptP::sptP$^{(1-159)}$-phoA-2xFLAG-6xHis ΔinvA* were included in each deepwell plate as controls. Overnight cultures were stored for analysis and high-throughput sequencing by diluting 180 µL of overnight culture with 60 µL 60% glycerol in a sterile, round-bottom 96-well plate (Corning), sealing the plate, and storing it at -80°C. To facilitate secretion, overnight cultures were diluted 1:100 into 0.5 mL TB in a fresh 2 mL square 96-well deepwell plate and grown for 8 hours at 37°C, 350 rpm. The secretion fraction was harvested by pelleting cells in the deepwell plates at 4000 *x g* for 10 minutes, collecting 200 µL of the supernatant, and storing it in a sealed plate at 4°C.

5.2.6 Alkaline Phosphatase Activity

Alkaline phosphatase activity was measured by monitoring *p*-nitrophenol phosphate (pNPP, Sigma) cleavage. A stock solution of 0.1 M pNPP prepared in 1 M Tris pH 8.0 was thawed from -20°C and diluted to 0.01 M in 1 M Tris pH 8.0. 20 µL of the secretion fraction was added to 140 µL of 1 M Tris pH 8.0, and 40 µL of the 0.01 M pNPP solution was added to each well. AP activity was measured on a BioTek Synergy HTX plate reader by monitoring absorbance at 405 nm at 37°C for one hour, taking measurements each minute.

5.2.7 Sample Preparation for High Throughput Sequencing

The randomly arrayed glycerol stocks were inoculated in identical arrangements in fresh media in sterile, flat-bottom 96-well plates using a Tecan Fluent and grown with lids in DigiLab HiGro shaking stacks at 37°C, 200 rpm for 18 hours. Each clone was assigned a pool according to its relative secretion titer (Table 5.4). A Tecan Fluent was used to reformat clones and sort them into their assigned pools. A VBA macro assigned pools according to relative secretion titer and assigned clones to a new plate and well ID to provide instructions for the Tecan Fluent. 150 µL of the fresh cell suspension was mixed with 50 µL 60% glycerol in a fresh sterile, round-bottom 96-well plate (Corning). The Tecan Fluent failed after sorting pools F-J, so the remainder were done by hand over the course of a week. Each well of the newly sorted glycerol stocks was sampled and pooled

according to Table 5.4 for genomic DNA purification. Genomic DNA was purified from 1 mL of each pool and 0.5 mL of the naïve library using the GenElute Bacterial Genomic DNA kit (Sigma).

Table 5.4 Pools for high throughput sequencing according to relative secretion titer.

Pool	Relative Secretion Titer	Number of Clones	Sample Volume for Mixture (µL)	Nextera XT Primer i5	Nextera XT Primer i7
A	0.01-0.6	2015	5	N707	S502
B	0.6-0.8	575	10	N710	S502
C	0.8-1.0	691	10	N711	S502
D	1.0-1.2	480	20	N712	S502
E	1.2-1.4	236	20	N714	S502
F	1.4-1.6	151	20	N705	S503
G	1.6-1.8	77	30	N706	S503
H	1.8-2.0	75	30	N707	S503
I	2.0-2.5	70	30	N710	S503
J	>2.5	36	50	N711	S503

PCR for library preparation was conducted with Phusion polymerase. The purified genomes from the pool mixtures were amplified using the “Round 1” reaction recipe (Table 5.5) and cycling conditions (Table 5.6) with primers sLAB278 and sLAB279 (Table 5.3) to attach Illumina Nextera XT adapters. Reactions were purified using the Promega Wizard SV PCR cleanup kit. For each pool, 8 x 25 µL reactions were performed and pooled after PCR cleanup to minimize jackpot effects. A second round of PCR attached Nextera XT barcodes according to Table 5.4 using the “Round 2” reaction recipe and cycling conditions in Table 5.5 and Table 5.6 with the pooled Round 1 reactions as templates.

Table 5.5 PCR reactions for high throughput sequencing library preparation

Component	Round 1 (25 µL x 8 per pool)	Round 2 (50 µL x 1 per pool)
5X HF Buffer (NEB)	5 µL	10 µL
10 mM dNTPs (NEB)	0.5 µL	1 µL
10 µM FWD primer	1.25 µL	2.5 µL
10 µM REV primer	1.25 µL	2.5 µL
Template DNA	2.5 µL of 5 ng/µL gDNA	5 µL purified and combined Round 1 reaction
Phusion (NEB)	0.25 µL	0.5 µL
H_2O	14.25 µL	28.5 µL

Table 5.6 PCR cycling conditions for high throughput sequencing library preparation

	Round 1 (25 µL x 8 per pool)		Round 2 (50 µL x 1 per pool)	
Step	**T (°C)**	**Time (sec)**	**T (°C)**	**Time (sec)**
Initial Denaturation	98	60	98	30
Amplification	98	10	98	10
(Round 1 – 22 cycles)	62	15	61	15
(Round 2 – 8 cycles)	72	30	72	30
Elongation	72	300	72	300
Hold	4	indefinite	4	indefinite

5.2.8 High-Throughput Sequencing Data Processing

The code for data processing using the Linux command-line interface (bash) is given following each explanation. Data were trimmed using Trimmomatic64 with a 2-unit sliding quality window of 30 and a minimum length of 30. Sequences were cropped to 243 bp.

java -jar trimmomatic-0.36.jar SE input_forward_HTS001.fastq.gz HTS001_trimmed SLIDINGWINDOW:2:30 MINLEN:30 CROP:243

Reads were then aligned to the wild-type PrgI reference gene with Burrows–Wheeler Aligner (BWA-MEM)66 and piped into Samtools to convert to a bam file.

bwa mem -p Reference/ref.fasta HTS001_trimmed.fastq | samtools view -bT Reference/ref.fasta -o HTS001.bam

Reads that fully mapped to PrgI were kept for further analysis.

samtools view HTS001.bam | grep "243M" | sort | less -S>HTS001.txt

The trimmed reads were further processed to generate a secretion fitness landscape using code written in-house (Appendix C).

5.3 Results

5.3.1 Saturation mutagenesis at a single residue of PrgI reveals variable secretion titer

Kenjale *et al.* conducted an alanine scan of the homologous T3SS needle protein of *Shigella flexneri*, MxiH, and reported mutations that changed secretion behavior [185]. Previous Tullman-Ercek lab members mapped a subset of those mutations to PrgI and discovered that those mutations altered heterologous protein secretion titer [181]. One of the mapped mutations, $PrgI^{P41A}$, increased secretion titer twofold. The P41 residue is located in the hinge region of the PrgI monomer and has extensive contact with neighboring monomers [191], so it was interesting that it could be changed at all. To

further investigate the mutational tolerance of this residue, we performed saturation mutagenesis. In the course of mapping mutations from MxiH to PrgI, we also discovered that PrgI could not be complemented by a plasmid in our system [181], so all mutations were incorporated into the genomic copy via recombineering [152]. Secretion titer was measured using quantitative western blotting for the position 41 mutants and the remaining mutations mapped from MxiH (L9A, Q48A, Y54A, and D70A) (Figure 5.2).

We were surprised to discover that many substitutions at residue 41 were functional. Large residues and those with positive or negative chargers were generally not permitted, but alanine, cysteine, threonine, and valine all provided wild-type or higher secretion titers. The mutational tolerance at this residue and the sensitivity of secretion titer to single amino acid changes in PrgI showed promise that a CCM strategy modeled after SyMAPS would yield higher secreting variants. We first needed to create and optimize a workflow to construct and screen the library, however.

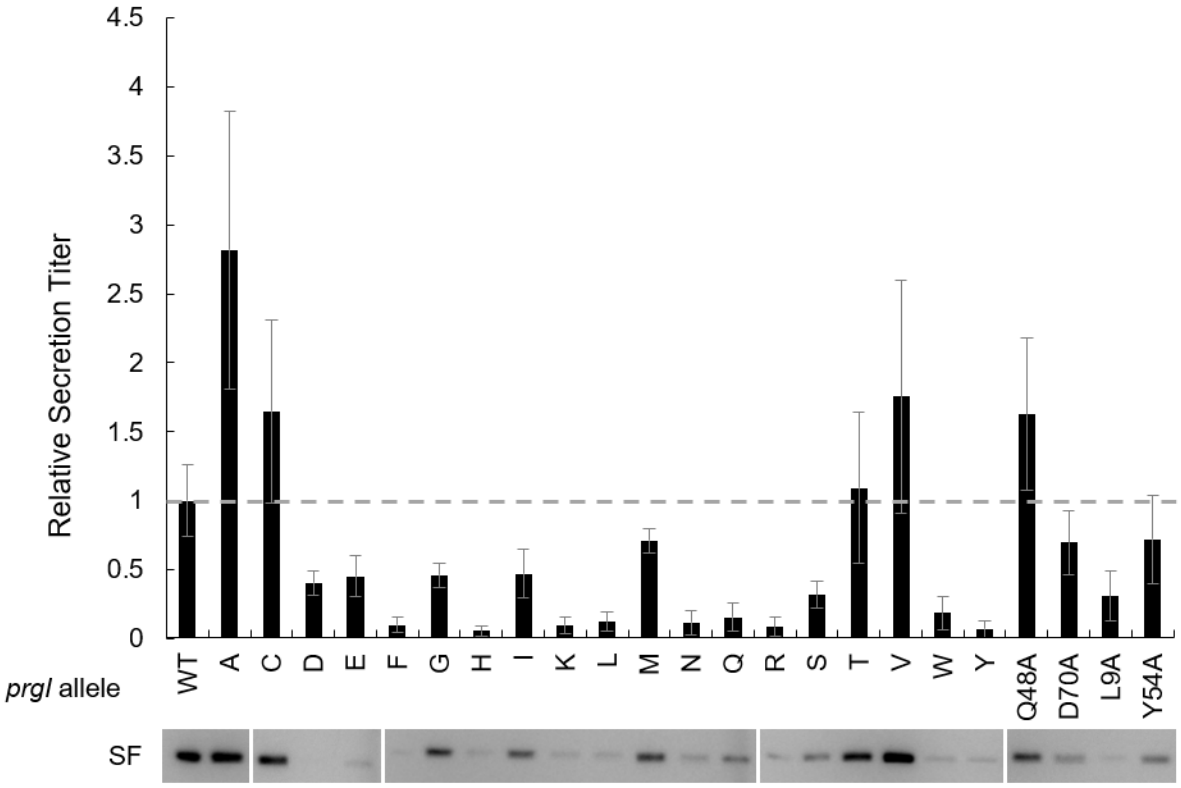


Figure 5.2 Relative secretion titer of SptP-DH-2xFLAG-6xHis from mutations at PrgI residue 41. SptP-DH-2xFLAG-6xHis was secreted in ASTE13 containing the indicated PrgI variant. P_{lacUV5} *hilA* overexpression facilitated T3SS and SptP-DH-2xFLAG-6xHis expression. Secretion titer was measured using densitometry from western blots and normalized to a $PrgI^{WT}$ strain. Error bars represent standard error. “SF” is “secreted fraction” and western blots are representative of three biological replicates.

5.3.2 Optimizing an alkaline phosphatase activity assay for high-throughput screening of secretion titer

Ideally, library screening is conducted in a high-throughput fashion via flow cytometry or by exerting a selection pressure to isolate positive clones in bulk. Screening for secretion titer separates genotype from phenotype, however, and these common techniques would then be indirect measurements. Indirect measurements would likely be sufficient for an initial screen or selection, but we have not yet succeeded in finding a flow cytometry or selection method that is sensitive and robust enough for our system. Thus, we were confined to measuring secretion titer of individual library members.

Section 4.3.4 showed that a plate-based alkaline phosphatase activity assay could replace western blotting for secretion titer measurement. I followed the established protocol and compared initial rate and densitometry for relative secretion titer of SptP-AP-2xFLAG-6xHis secreted from the PrgI position 41 mutants (Figure 5.3). The pattern of secretion titer relative to $PrgI^{WT}$ was similar to SptP-DH-2xFLAG-6xHis, though the absolute ratios differed. Importantly, the two methods of measuring secretion titer corresponded well, indicating that the pNPP assay was sufficiently sensitive for screening a PrgI mutant library.

Plasmid-based secreted protein and T3SS expression was problematic for screening a PrgI mutant library. Genomic integration of the library would preclude pre-transforming the strain used to construct the library, and the number of clones required to achieve adequate coverage would increase if we transformed the entire library after construction [188]. Thus, I integrated SptP-AP-2xFLAG-6xHis at the ASTE13 *sptP* locus, and I attempted to re-optimize the alkaline phosphatase activity assay for secretion titer measurement in ASTE13 *sptP::sptP*$^{(1\text{-}159)}$*-phoA-2xFLAG-6xHis* without a HilA overexpression plasmid.

A genomic copy of the secreted protein decreases expression and secretion titer by decreasing gene copy number relative to the high-copy P_{sic} plasmid, so I extended the activity assay timeline to 90 minutes. According to Figure 5.4, however, 90 minutes was still insufficient—absorbance (and therefore product formation) increased less than twofold relative to background and achieved a 20-fold lower value at 90 minutes than previously observed for plasmid-based SptP-AP-2xFLAG-6xHis expression at 35 minutes. I repeated the experiment, this time recording absorbance for 16 hours, and was able to recapitulate the pattern originally observed for SptP-DH-2xFLAG-6xHis secretion in the position 41 mutants with HilA overexpression (Figure 5.5).

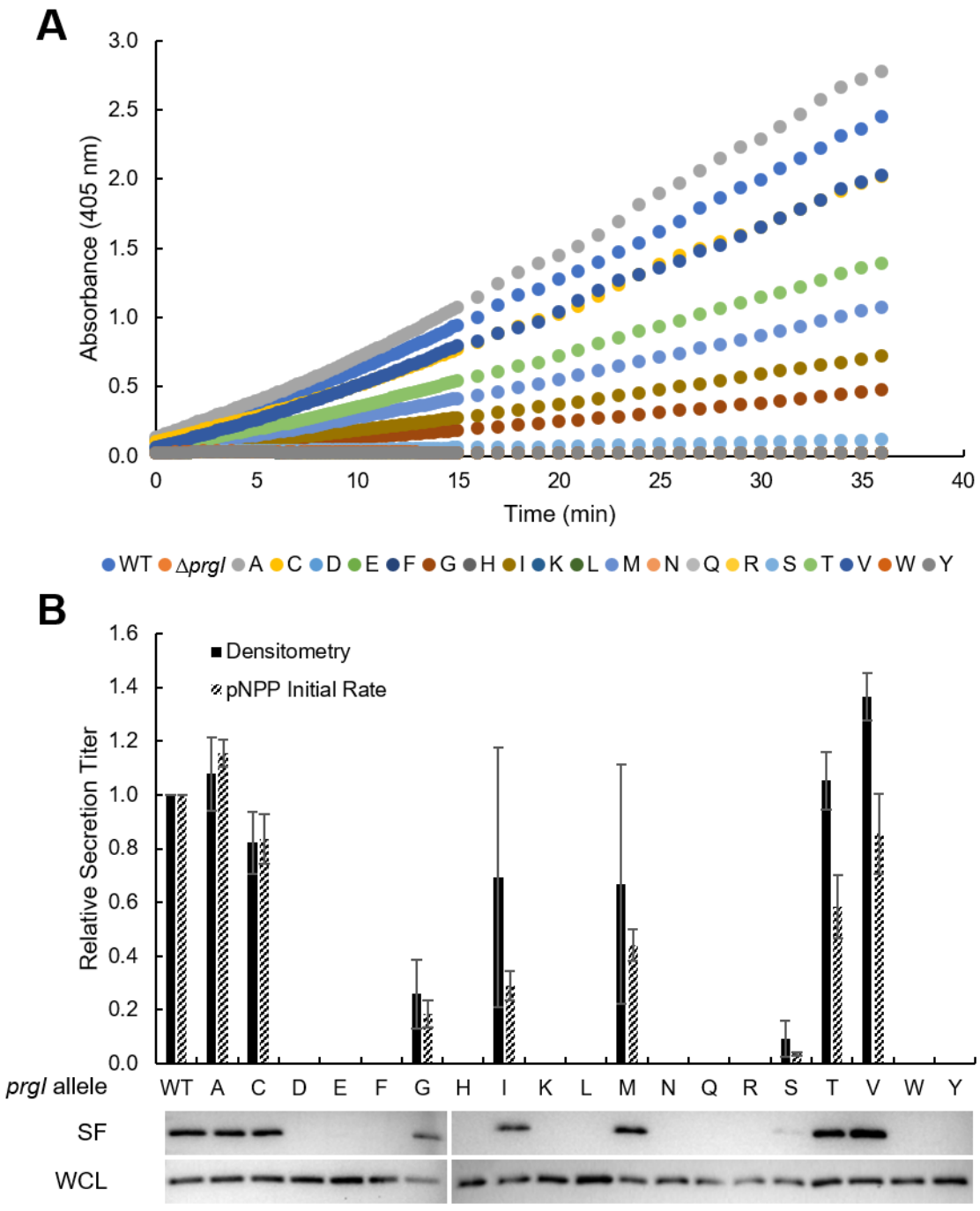


Figure 5.3 Absorbance over time and initial rate versus densitometry for secreted alkaline phosphatase expressed from a plasmid.
SptP-AP-2xFLAG-6xHis was secreted from the specified PrgI mutants in LB-L according to Methods. P_{lacUV5} *hilA* overexpression facilitated T3SS and SptP-AP-2xFLAG-6xHis expression. Densitometry was calculated from western blots and normalized to $PrgI^{WT}$. Initial rate was calculated by averaging the slope of A_{405nm} versus time for two biological replicates and normalizing to $PrgI^{WT}$ (**B**). Absorbance at 405 nm was recorded every 15 seconds for 15 minutes and then every minute until 36 minutes on a BioTek Synergy HTX plate reader at 37°C (**A**). Error bars represent standard error.

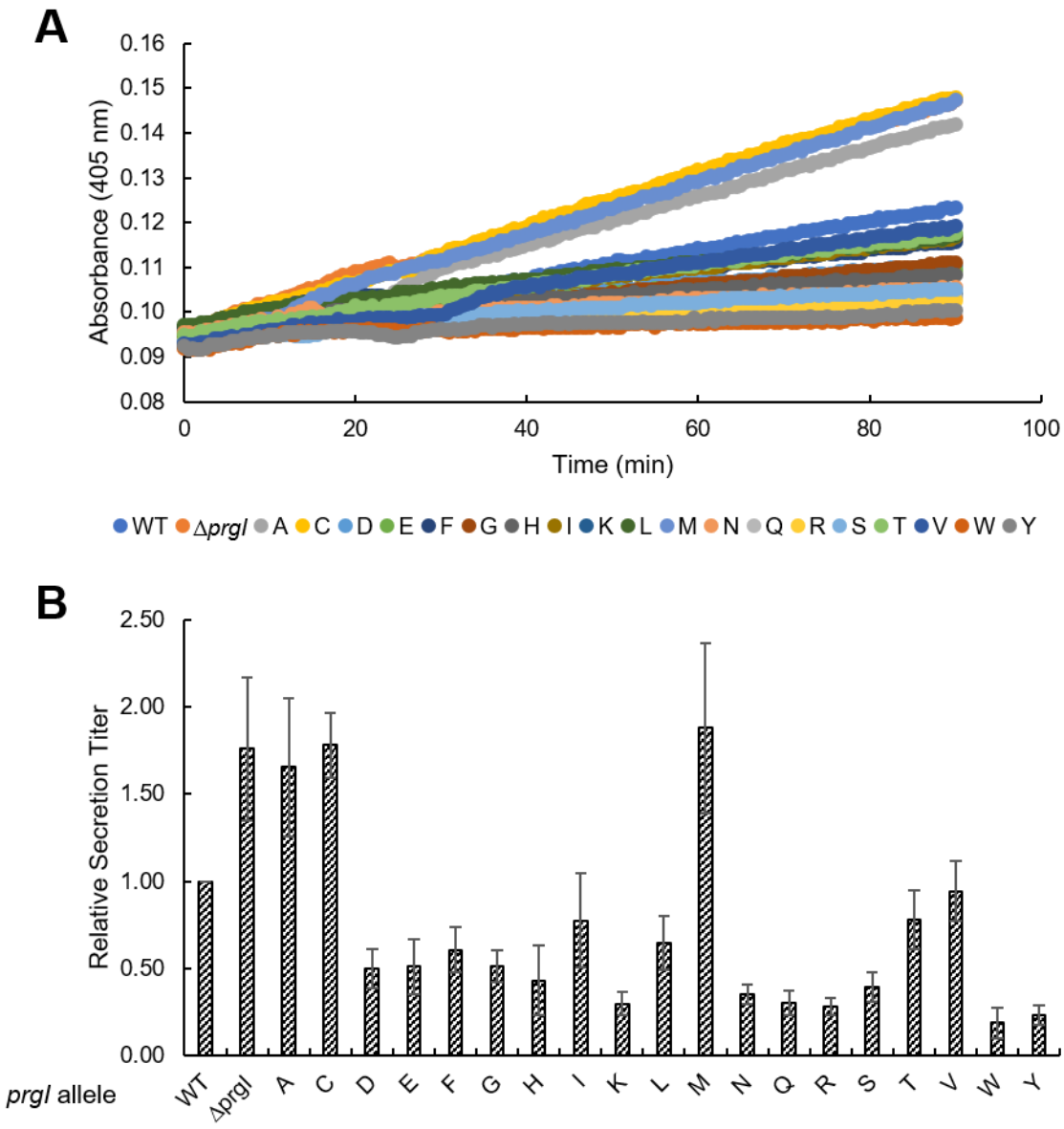


Figure 5.4 Product formation is slow with alkaline phosphatase secreted from a genomic copy. Absorbance over 90 minutes (**A**) and relative secretion titer (**B**) for genomically integrated alkaline phosphatase. SptP-AP-2xFLAG-6xHis was secreted from ASTE13 *sptP::sptP*$^{(1\text{-}159)}$*-phoA-2xFLAG-6xHis* and the specified PrgI mutants in LB-L. "Δ*prgI*" is ASTE13 *prgI*::*catsacB*. Relative secretion titer was measured by averaging the slope of A_{405nm} versus time for four biological replicates and normalizing to PrgI^{WT}. Absorbance was recorded every 30 seconds for 90 minutes on a BioTek Synergy HTX plate reader at 37°C. Error bars represent standard error.

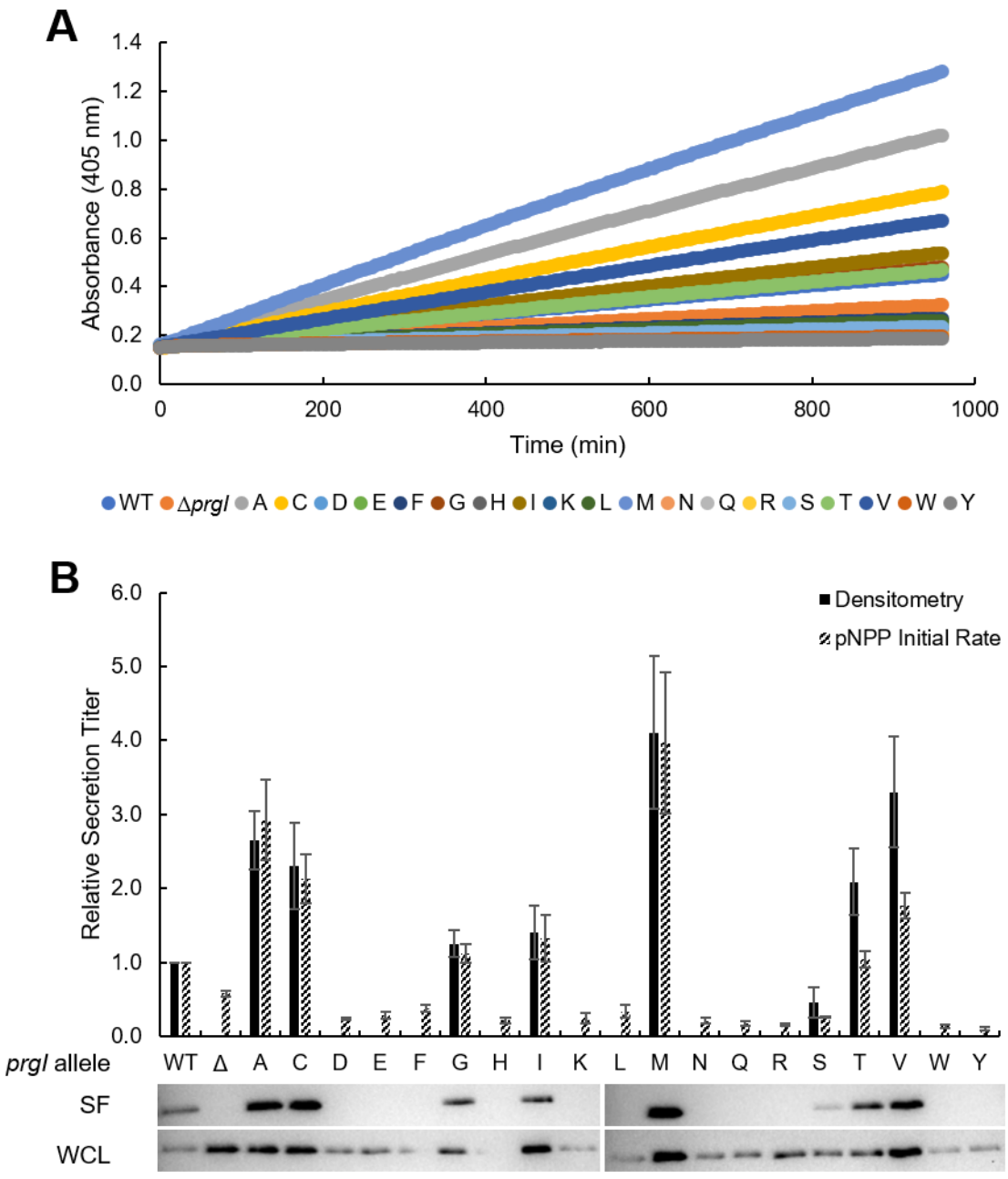


Figure 5.5 Product formation is slow with alkaline phosphatase secreted from a genomic copy. Absorbance over 16 hours (**A**) and relative secretion titer (**B**) for genomically integrated alkaline phosphatase. SptP-AP-2xFLAG-6xHis was secreted from ASTE13 *sptP::sptP*$^{(1\text{-}159)}$*-phoA-2xFLAG-6xHis* and the specified PrgI mutants in LB-L. "Δ*prgI*" is ASTE13 *prgI*::*catsacB*. Relative secretion titer was measured by averaging the slope of A_{405nm} versus time for four biological replicates and normalizing to PrgI^{WT}. Absorbance at 405 nm was recorded every 5 minutes for 16 hours on a BioTek Synergy HTX plate reader at 37°C. Error bars represent standard error.

Up to this point, all secretion experiments in this work were performed in 5 mL cultures in Axygen 24-well deepwell plates. While the 24-well deepwell plates improved sample handling considerably over test tubes and flasks, screening thousands of clones from a library, each in 5 mL of media, would generate a medium supply problem. Comprehensive codon mutagenesis of the entire PrgI monomer would create 1600 unique library members, and we would need to screen a minimum of 4800 members for >95% coverage [188]. Ideally, we would use the same batch of media for the entire project, so 5 mL cultures would require a minimum 48 L of media for overnight cultures and secretion experiments. Our lab is not accustomed to handling such volumes of media, so I attempted to scale down secretion experiments to 2 mL square 96-well deepwell plates. Comparing secretion of SptP-DH-2xFLAG-6xHis in various culture formats and media showed that protein secretion was detectable relative to *hilA* overexpression in 2 mL 96-well deepwell plates with LB-IM (10 g/L tryptone, 5 g/L yeast extract, 17 g/L NaCl) media (Figure 5.6).

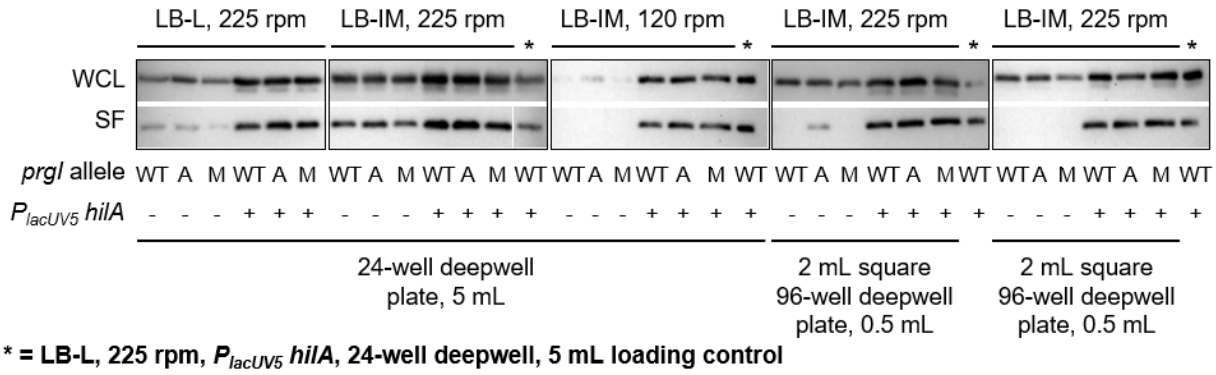


Figure 5.6 SptP-DH-2xFLAG-6xHis secretion in various media and culture formats.
SptP-DH-2xFLAG-6xHis was secreted for 8 hours in the indicated media, strains, and culture vessels. Western blots were performed according to Methods, and each blot included an ASTE13 WT, LB-L 225 rpm, 24-well deepwell plate, 5 mL loading control.

I returned to the genomically integrated SptP-AP-2xFLAG-6xHis to determine if there was detectable enzyme activity in the secreted fraction of a culture grown in 0.5 mL LB-IM in a 2 mL square 96-well deepwell plate at 225 rpm. I also evaluated secretion in 5 mL LB-IM in a 24-well deepwell plate. I shortened the time for recording product formation via absorbance at 405 nm to two hours because 16 hours of kinetic measurements is unreasonable for screening thousands of clones with a single plate reader. Enzyme activity was detectable in the expected pattern in the secreted fraction from the 24-well deepwell plate, but the secreted fraction from the 96-well deepwell plate did not produce detectable enzyme activity (Figure 5.7).

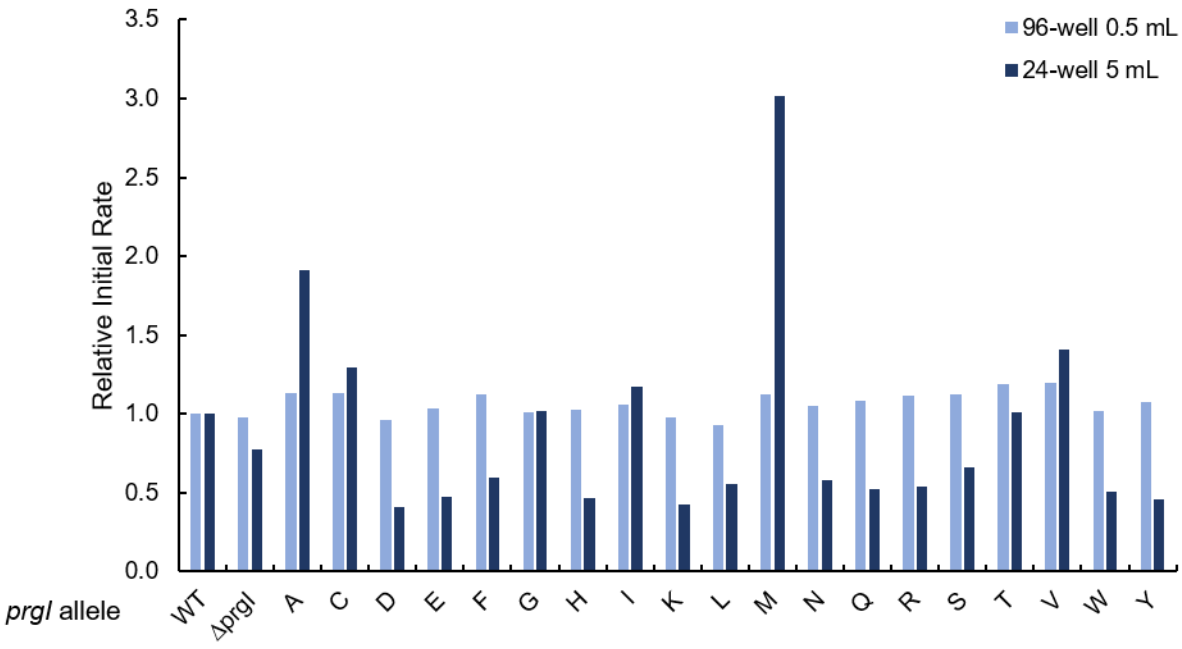


Figure 5.7 Enzyme activity is not detectable in the secreted fraction of a 2 mL square 96-well deepwell plate culture with LB-IM.
Genomically integrated SptP-AP-2xFLAG-6xHis was secreted in 0.5 mL in a 2 mL square 96-well deepwell plate and 5 mL in a 24-well deepwell plate in LB-IM at 37°C and 225 rpm. ASTE13 WT and the specified PrgI mutants contained *sptP::sptP*$^{(1\text{-}159)}$*-phoA-2xFLAG-6xHis*. "Δ*prgI*" is ASTE13 *prgI::catsacB*. Relative secretion titer is represented by relative initial rate, which was the slope of A_{405nm} versus time for a single replicate normalized to PrgI^{WT} from the same growth condition. Absorbance at 405 nm was recorded every two minutes for two hours on a BioTek Synergy HTX plate reader at 37°C.

I next attempted to minimize medium consumption by performing secretion in 2 mL LB-IM in a 24-well deepwell plate. I tried several conditions—2 mL overnight (ON) cultures with 2 mL secretion cultures, 2 mL ON cultures with 5 mL secretion cultures (Figure 5.8A), 2 mL ON cultures containing 0.4% w/v glucose to repress secretion system activity with 5 mL secretion cultures, and 5 mL ON cultures containing 0.4% w/v glucose with 5 mL secretion cultures (Figure 5.8B). All conditions were compared to the standard 5 mL of ON culture with 5 mL of secretion culture. All ON cultures were in LB-L. Unfortunately, no condition including a 2 mL culture matched the standard 5 mL cultures, and glucose in the ON culture had no effect. The Δ*prgI* negative control also had WT-level or higher signal in all conditions (also true in Figure 5.7), suggesting that either LB-IM was causing leakage of SptP-AP-2xFLAG-6xHis or that secretion titer was still too low to develop an appropriate signal-to-noise ratio over two hours of product formation.

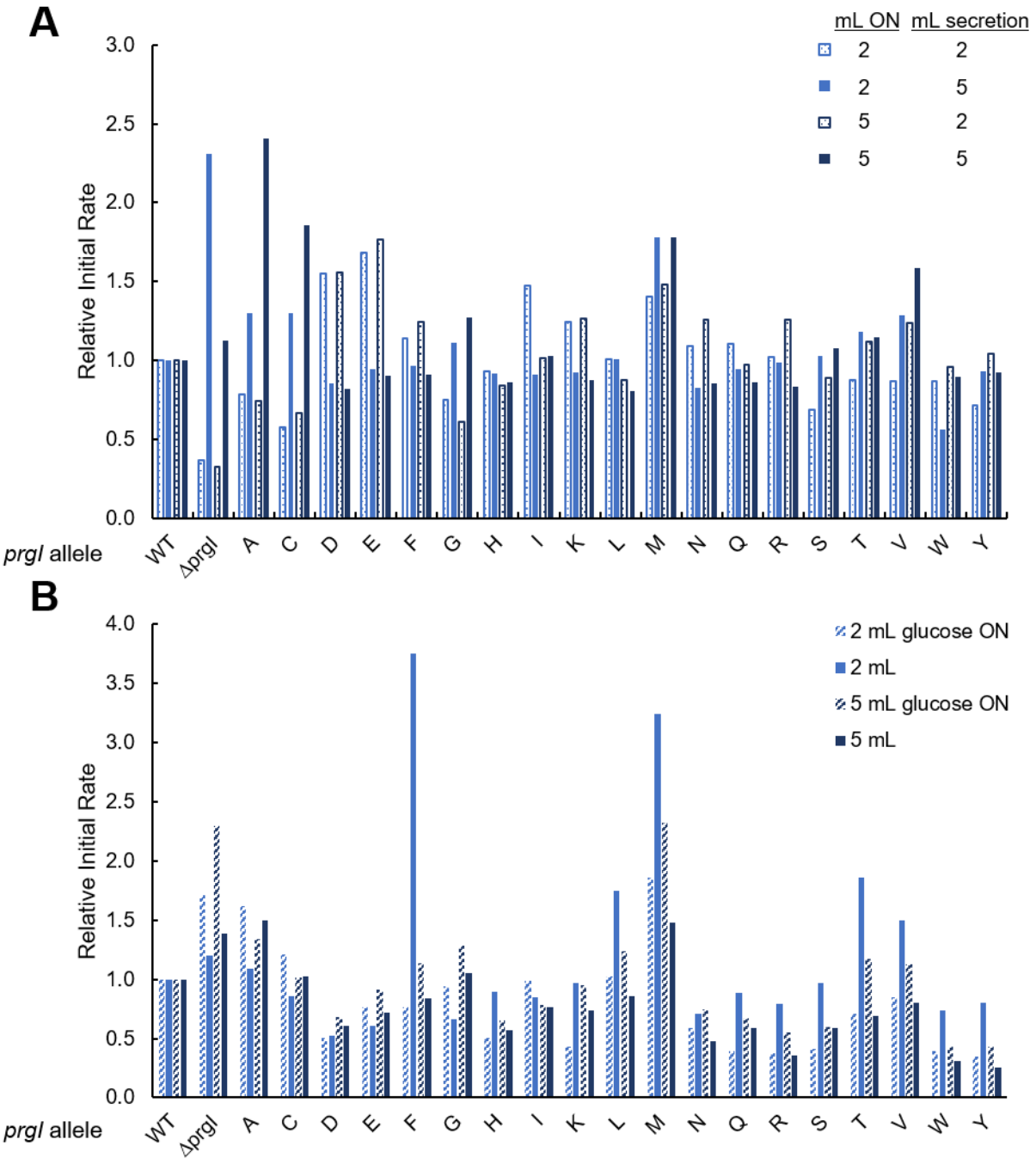


Figure 5.8 Secretion with a genomic copy of SptP-AP is incompatible with 2 mL cultures in 24-well deepwell plates.
Genomically integrated SptP-AP-2xFLAG-6xHis was secreted in LB-IM at the specified conditions. In (**A**), ON cultures were LB-L and secretion cultures were LB-IM in the listed volumes. In (**B**), volumes listed refer to ON cultures of LB-L. Glucose was 0.4% w/v. Secretion cultures were 5 mL LB-IM. ASTE13 WT and the specified PrgI mutants contained *sptP::sptP*$^{(1\text{-}159)}$*-phoA-2xFLAG-6xHis*. "Δ*prgI*" is ASTE13 *prgI::catsacB*. Relative secretion titer is represented by relative initial rate, which was the slope of A_{405nm} versus time for a single replicate normalized to PrgI^{WT} from the same growth condition. Absorbance at 405 nm was recorded every two minutes for two hours on a BioTek Synergy HTX plate reader at 37°C.

As a final effort to scale down cultures for library screening, I secreted genomically integrated SptP-AP-2xFLAG-6xHis from the PrgI position 41 mutants in 0.5 mL terrific broth (TB, without glycerol) in 2 mL square 96-well deepwell plates at 225 and 350 rpm. Overnight cultures were 0.5 mL LB-L in 2 mL square 96-well deepwell plates at the same shaking speed used for secretion. I also performed secretion in TB with the usual 5 mL, 24-well deepwell plate, 225 rpm conditions. The *prgI::catsacB* "Δ*prgI*" control often produced high signal, so I added a Δ*invA* strain with genomically integrated SptP-AP-2xFLAG-6xHis to provide a cleaner negative control.

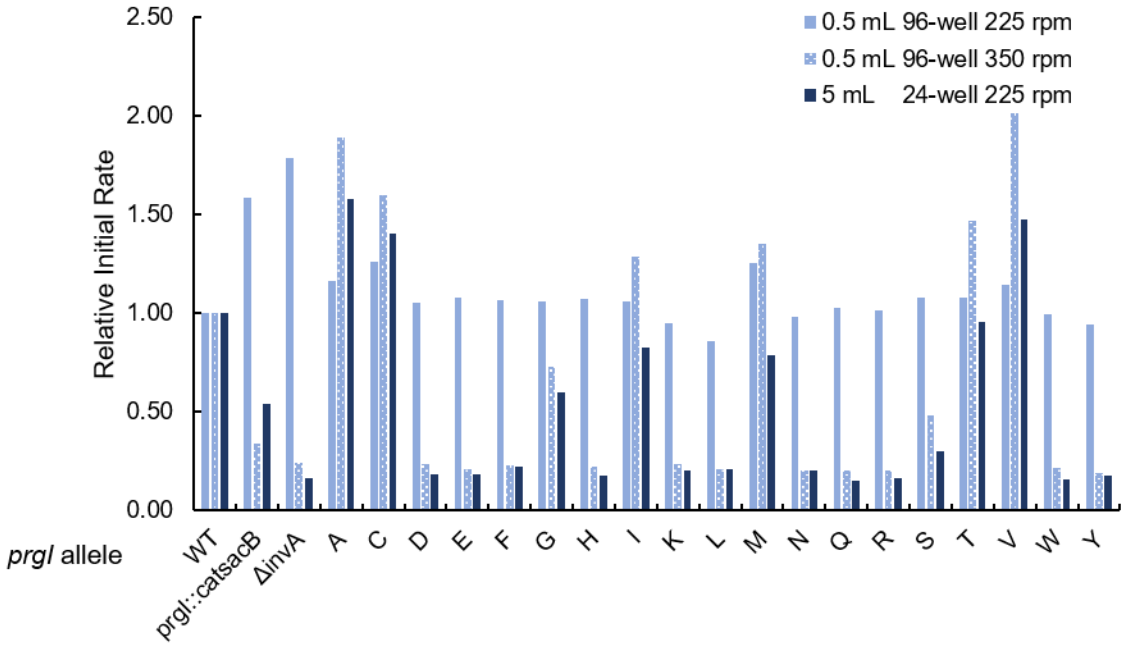


Figure 5.9 TB media and increased shaking speed produce measurable secreted enzyme activity from 2 mL square 96-well deepwell plate cultures.
Genomically integrated SptP-AP-2xFLAG-6xHis was secreted in TB at the specified conditions. ON cultures were LB-L at the same conditions as used for secretion. ASTE13 WT and the specified PrgI mutants contained *sptP::sptP*$^{(1\text{-}159)}$*-phoA-2xFLAG-6xHis*. "Δ*prgI*" is ASTE13 *prgI::catsacB* and Δ*invA* is a clean knockout. Relative secretion titer is represented by relative initial rate, which was the slope of A_{405nm} versus time for a single replicate normalized to PrgI^{WT} from the same growth condition. Absorbance at 405 nm was recorded every two minutes for two hours on a BioTek Synergy HTX plate reader at 37°C.

I finally found a scaled-down condition that produced measurable secreted enzyme activity over a reasonable time frame in the 0.5 mL TB cultures shaken at 350 rpm (Figure 5.9). Relative secretion titers, represented by relative initial rate, showed the expected pattern, and the negative controls were well below PrgI^{WT}. I repeated the experiment in biological triplicate to assess variance and discovered that monitoring product formation for one hour developed a sufficient signal-to-noise ratio for those secretion conditions (Figure 5.10).

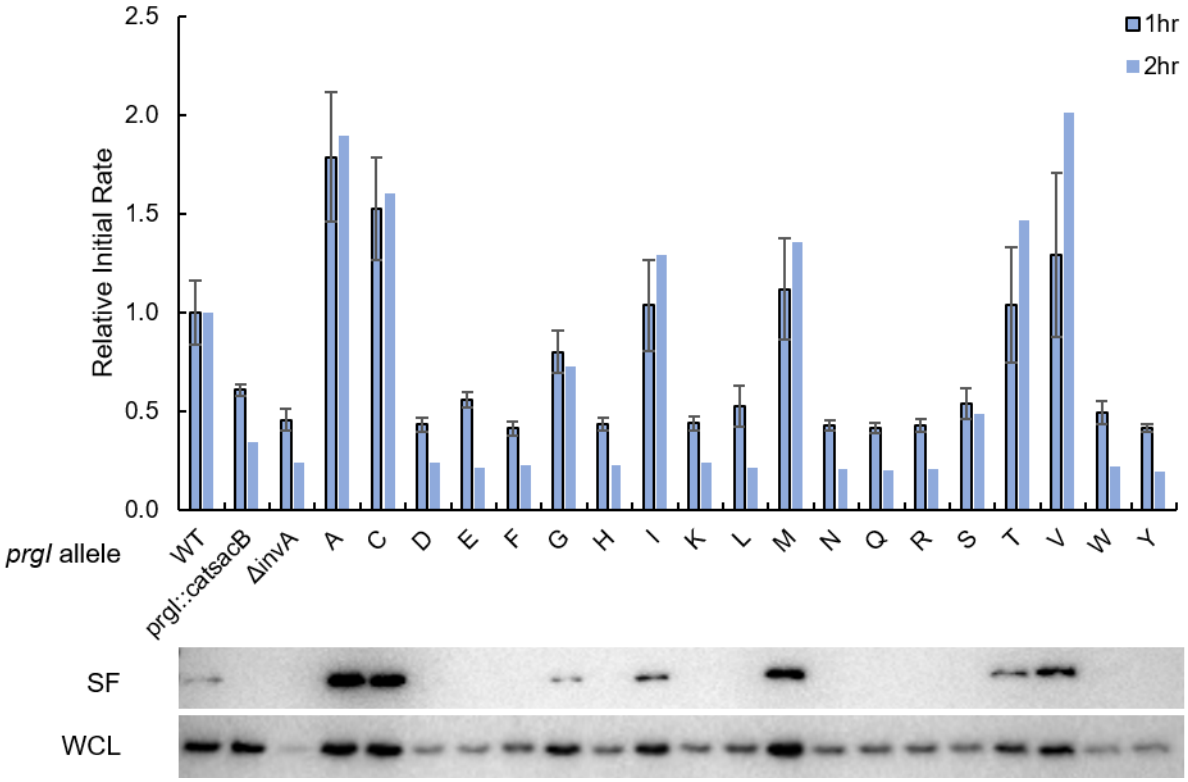


Figure 5.10 One hour is a sufficient measurement range for secreted SptP-AP activity harvested from secretion in 0.5 mL TB, and 350 rpm in 2mL square 96-well deepwell plates.
Genomically integrated SptP-AP-2xFLAG-6xHis was secreted in 0.5 mL TB at 37°C and 350 rpm in 2 mL square 96-well deepwell plates. ON cultures were 0.5 mL LB-L at 37°C and 350 rpm in 2 mL square 96-well deepwell plates. ASTE13 WT and the specified PrgI mutants contained *sptP::sptP*$^{(1\text{-}159)}$*-phoA-2xFLAG-6xHis*. "Δ*prgI*" is ASTE13 *prgI::catsacB* and Δ*invA* is a clean knockout. Relative secretion titer is represented by relative initial rate, which was the slope of A_{405nm} versus time for a single replicate normalized to PrgI^{WT} from the same growth condition. The one-hour data set is the average of three biological replicates; error bars represent standard error. Absorbance at 405 nm was recorded every two minutes for two hours or every two minutes for one hour on a BioTek Synergy HTX plate reader at 37°C.

5.3.3 Generating a genomically-encoded CCM library

The next method requiring validation after the alkaline phosphatase activity assay was construction of the genome-based library via recombineering. Recombineering has been reported as a technique for plasmid-based library construction [189], but to our knowledge no genome-based mutagenesis library exists in the literature. Figure 5.11 provides a schematic for our recombineering process, adapted from Court and colleagues [152].

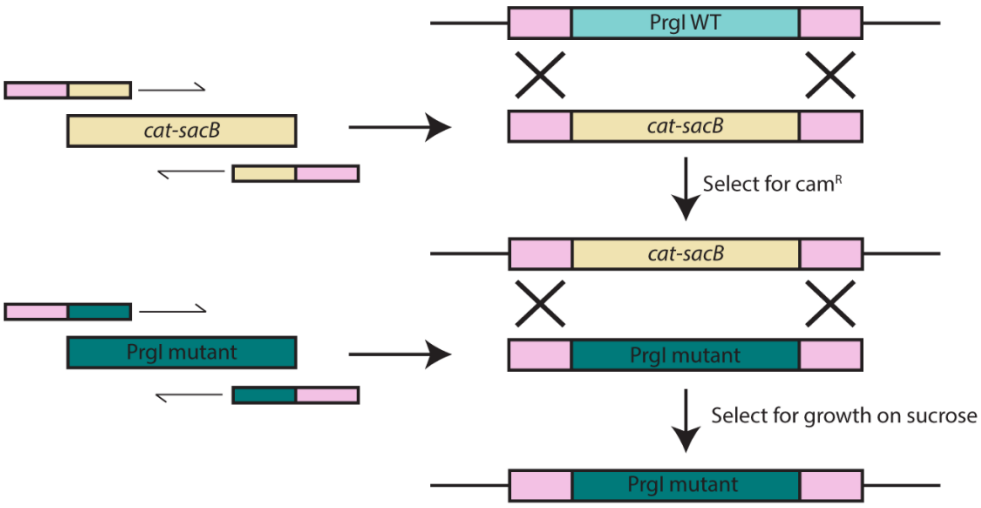


Figure 5.11 Overview of recombineering for genomic modification.
A dual selection *cat-sacB* cassette conferring chloramphenicol resistance and sucrose sensitivity is amplified with 40 bp of homology 5' and 3' to the prgI locus. Enzymes from the λ Red system facilitate homologous recombination. Successful recombinants are selected for chloramphenicol resistance. The *cat-sacB* cassette is replaced in a second step of recombineering, and the λ Red enzymes facilitate homologous recombination of a gene encoding a mutation in prgI with the same 5' and 3' homology as was attached to the *cat-sacB* cassette. Successful recombinants are selected by growth on sucrose.

The first parameter I evaluated was the false positive rate. The *sacB* portion of the *cat-sacB* dual selection cassette confers sucrose sensitivity, but we often observe false positives on sucrose agar plates. I started with an ASTE13 *sptP::sptP$^{(1-159)}$-phoA-2xFLAG-6xHis prgI::catsacB* strain, which was the negative control for the alkaline phosphatase activity assay development. In separate recombineering events, I introduced PrgI^{P41A}, PrgI^{P41C}, PrgI^{P41M}, PrgI^{P41T}, and PrgI^{P41V} to replace the *cat-sacB* cassette. I evaluated recombineering success by phenotype—successful recombinants should lose chloramphenicol resistance with the *cat-sacB* cassette. I patched all colonies for each variant on 10 μg/mL chloramphenicol LB-agar and found that 89% or more of all colonies for each PrgI mutation were successful recombinants (Table 5.7).

Table 5.7 Recombination efficiency at the *prgI* locus.

PrgI Mutation	Number of Colonies Patched	Patched Colonies with cam^R	Successful Recombinants	Total Cells Lacking Sucrose Sensitivity	% False Positive
P41A	17	1	8.0 x 10^5	8.5 x 10^5	6
P41C	42	4	1.9 x 10^6	2.1 x 10^6	9
P41M	47	5	2.1 x 10^6	2.4 x 10^6	11
P41T	35	3	1.6 x 10^6	1.8 x 10^6	9
P41V	49	5	2.2 x 10^6	2.5 x 10^6	10

The next step was to test if a single recombineering experiment could generate a well-distributed, genomically-encoded library. I used the collection of PrgI position 41

variants as a test pool. I performed the second step of recombineering with the ASTE13 *sptP::sptP*$^{(1-159)}$*-phoA-2xFLAG-6xHis prgI::catsacB* strain but instead of the usual single sample of double-stranded PCR product, I electroporated 200 ng of an equimolar pool of double-stranded PCR products for all position 41 mutants. To my surprise, it worked well with no modifications to the recombineering protocol. I performed colony PCR on 68 colonies to provide the greater than threefold sample required for >95% library coverage [188] and Sanger sequenced the clones to analyze how well the mutations were distributed. All mutants were accounted for except P41C, and the distribution of mutations was fairly even (Figure 5.12). Successful recombinants were at least 88% of all screened clones.

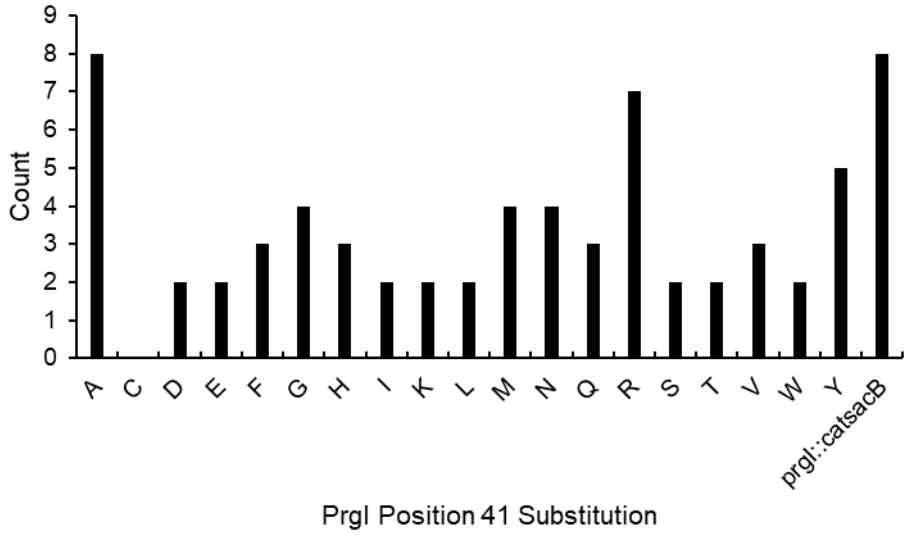


Figure 5.12 Distribution of position 41 mutants from a single recombineering event.
Counts of amino acid substitutions from a single recombineering event with an equimolar pool of PrgI position 41 variants. The "*prgI::cat-sacB*" count represents failed recombinants.

With established library construction and screening methods, I was ready to build and assemble the library. I ordered a library of gene blocks from Twist Biosciences, each with a single amino acid substitution from fully randomized codons. The library contained 1292 unique members, as the library did not contain WT residues or modifications of the first six and last six amino acids because those are essential for proper needle assembly [185,190]. I generated the library in a single recombineering step as described in Methods. I obtained similar recombination efficiency as for the mixture of position 41 mutants—the number of successful recombinants was 5 x 10^5 with 9% false positives.

The screening process was slow, as each clone had to be evaluated individually. The maximum throughput was 372 library members per day, or four 96-well plates each with three control wells for PrgI^{WT}, *prgI::cat-sacB*, and Δ*invA* strains carrying SptP-AP-2xFLAG-6xHis expressed from the genome. I screened 4400 clones total to achieve >95% coverage and compensate for the false positive rate. Secretion titer for each library member was calculated from the initial rate of product formation from alkaline phosphatase activity normalized to that from PrgI^{WT}. I linked genotype and phenotype by

storing glycerol stocks of the overnight cultures in the same layout as I measured the enzyme assay, so each well ID corresponded to a specific relative secretion titer.

A histogram of the screened clones showed multiple populations (Figure 5.13). More than half of all screened clones were functional (54%), and a surprising 2% had secretion titers greater than twofold above PrgI^{WT}. Functional clones were defined as those with secretion titers greater than 0.6 relative to PrgI^{WT}.

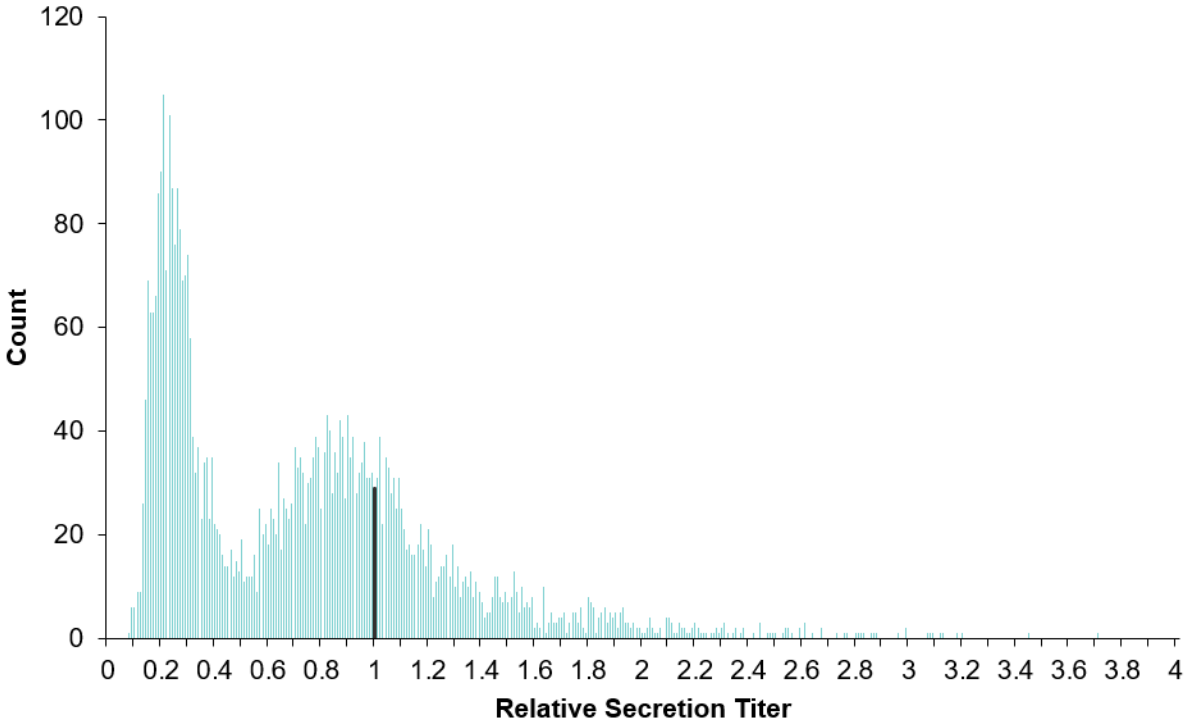


Figure 5.13 Histogram of screened PrgI library members according to relative secretion titer. SptP-AP-2xFLAG-6xHis relative secretion titer was calculated by normalizing the initial rate of product formation for the library member to that of PrgI^{WT} within each assay plate. Secretion titer of PrgI^{WT} is marked by a black line.

5.3.4 A "secretion fitness landscape" predicts secretion titer resulting from a single amino acid change in PrgI

The use of high-throughput sequencing to construct a fitness landscape from comprehensive codon mutagenesis was first developed in our lab to characterize single amino acid mutations that affected assembly of the MS2 viral capsid [187]. The procedure, SyMAPS, was developed for a bulk selection for assembly competence, so we adapted the workflow for the screened PrgI mutants to accommodate the individual clones and known secretion titers. The overall library construction, screening, and analysis workflow is depicted in Figure 5.14.

In SyMAPS, the fitness landscape was assembled by comparing the relative abundance of a mutant after applying the selection pressure to its relative abundance in the naïve library [187]. In the PrgI library, "fitness" is relative secretion titer, so we wanted

to characterize mutations according to those values. Thus, instead of one “selected pool”, we assigned the PrgI library members to ten pools according to their relative secretion titer (Figure 5.14). The PrgI library members were randomly arrayed in glycerol stocks, however, so in order to sequence each pool, we had to physically sort the clones. This took longer than expected because while we successfully used a Tecan Fluent to copy the randomly arrayed glycerol stocks and provide fresh cultures, it failed after reformatting pools F-J. The remaining pools were thus sorted by hand. The reformatting workflow is depicted in boxes with dashed borders in Figure 5.14.

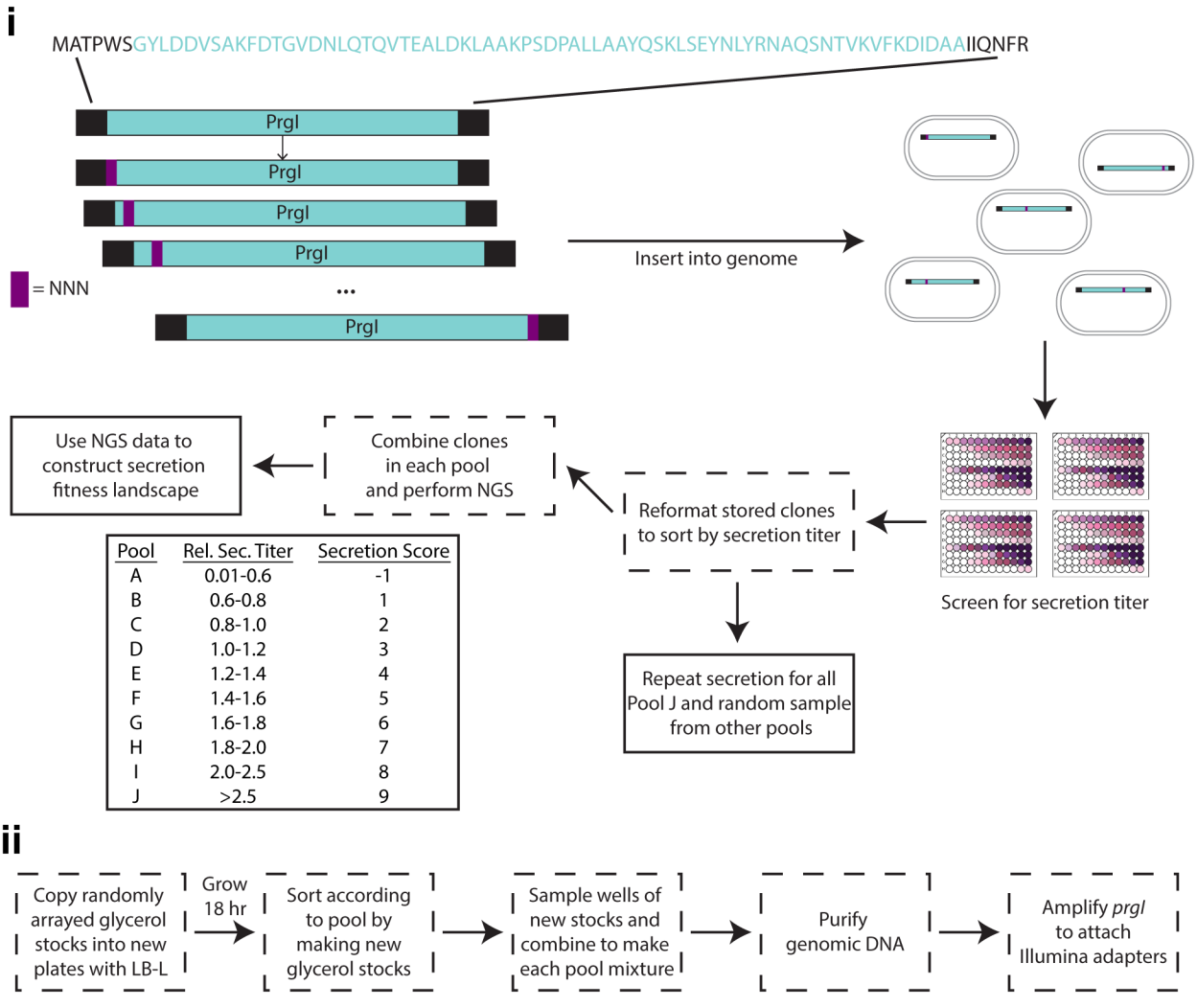


Figure 5.14 Workflow for PrgI library assembly, screening, and analysis.
Gene blocks coding for a single amino acid change in PrgI were introduced to ASTE13 *sptP::sptP*$^{(1\text{-}159)}$*-phoA-2xFLAG-6xHis prgI::catsacB* as a mixture in a single recombineering event to create the library. Individual colonies were inoculated for secretion and screened for alkaline phosphatase activity as described in Methods. The randomly arrayed clones were sorted according to relative secretion titer, combined into their assigned pools, and prepared for sequencing on an Illumina MiSeq (*ii*). Data was analyzed as described to construct the secretion fitness landscape.

Of the 1292 possible variants, 89% were identified in pools A-J, and 100% were identified in the naïve library. To generate the fitness landscape, each pool was assigned a “secretion score” from -1 for dysfunctional secretion (pool A) to 9 for the highest-secreting clones (pool J). Variants appeared in multiple pools, so we calculated a weighted-average secretion score for each library member using its relative abundance in each pool as weights. The laborious nature of the library screening and reformatting process meant that only a single biological and technical replicate was used to construct the “Secretion Fitness Landscape” (SFL), so it is a semi-quantitative assessment (Figure 5.15). The lack of replication made secretion competence uncertain for some amino acid changes—144 variants appeared in functional pools B-J but also in pool A, which contained secretion-incompetent variants. The weighted-average secretion score represents the relative abundance of the amino acid change in each pool, so pink-magenta variants are likely dysfunctional, while teal variants are likely functional.

The SFL revealed a distinct pattern—many substitutions were allowed in the N-terminal helix, but few conferred a significant increase in secretion titer. Conversely, few substitutions were allowed in the C-terminal helix, but the substitutions that were allowed generally conferred a significant increase in secretion titer. Mapping the average secretion score per residue onto the needle structure confirmed this pattern (Figure 5.16).

A recent model of the PrgI needle depicted a right-handed groove with alternating charged and hydrophobic residues forming the edges and lumen of the groove, respectively [175]. The raised, charged groove is highly conserved across all species with a T3SS [175,191], so we were not surprised to discover that modifying those residues inhibited secretion (Figure 5.17*ii*, Figure 5.18). We were surprised to discover that modifying the hydrophobic residues that line the groove and face the interior of the needle increased secretion titer, however. In fact, according to the SFL (Figure 5.15), it was substitution of bulkier and more hydrophobic amino acids at those residues that conferred increased secretion titer.

Functional substitutions at residues that likely affected inter- or intrachain interactions were of more varied character (Figure 5.18). Correct packing of needle monomers is essential for assembly of a secretion-competent needle, and many substitutions only produced functional secretion apparatuses when they replaced a native amino acid of the same character or size. Few residues with restricted substitutions provided increased secretion titers. Curiously, several residues (10, 21, 28, 37, 49, 63, 67) had a high average secretion score and allowed many substitutions despite having close contact with neighboring residues. We hypothesize that the substituted residues altered the conformation of needle monomers to provide a more favorable overall needle structure for increased secretion titer.

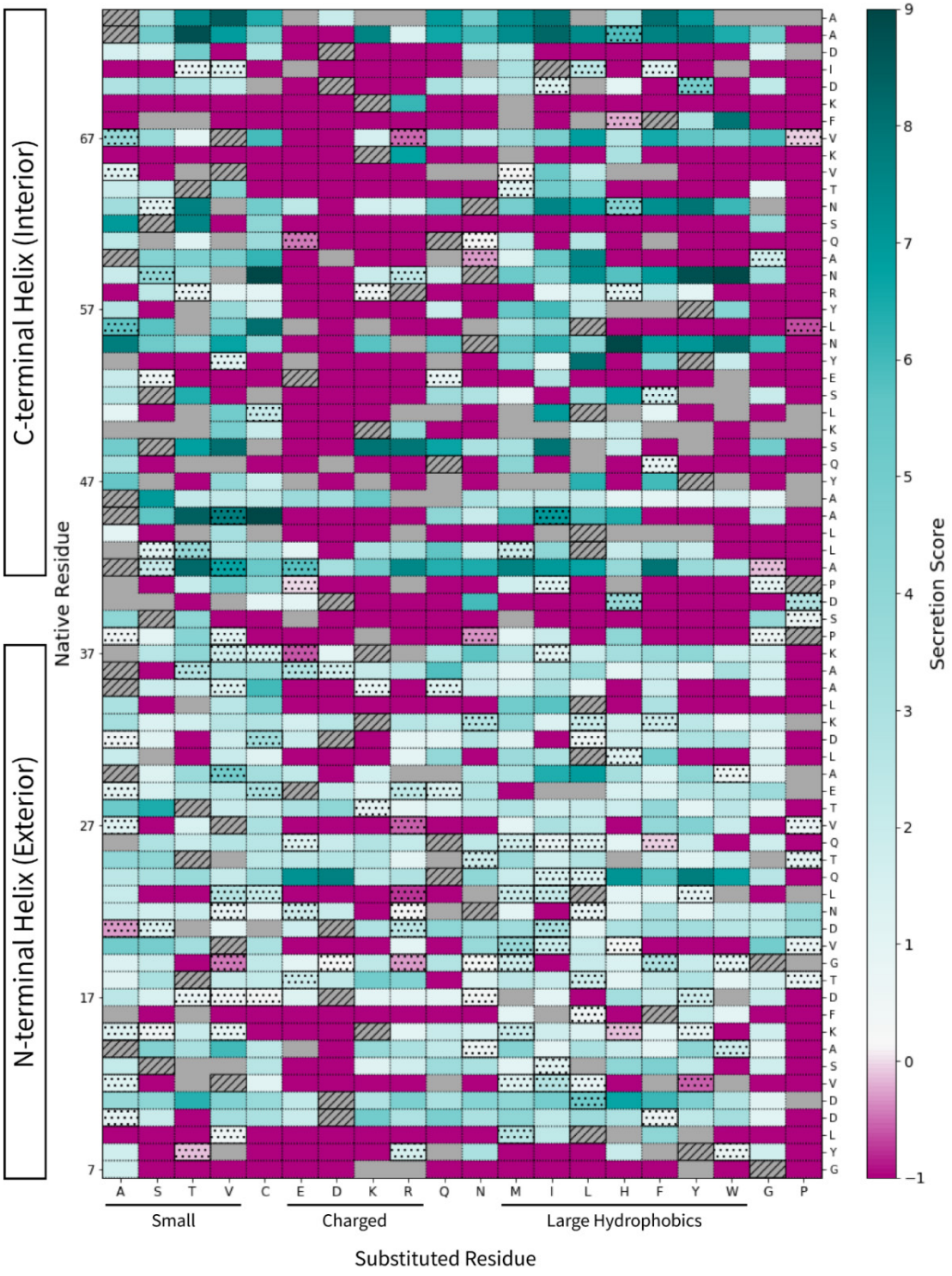


Figure 5.15 Weighted-average secretion scores for all single amino acid variants of the modified residues of PrgI.

Wild-type residues are indicated by hatches. Dots indicate that the variant appeared in pools B-J but also in pool A. Grey boxes denote variants that did not appear in pools A-J; i.e., those variants were not screened. Magenta boxes indicate that the variant was found only in pool A.

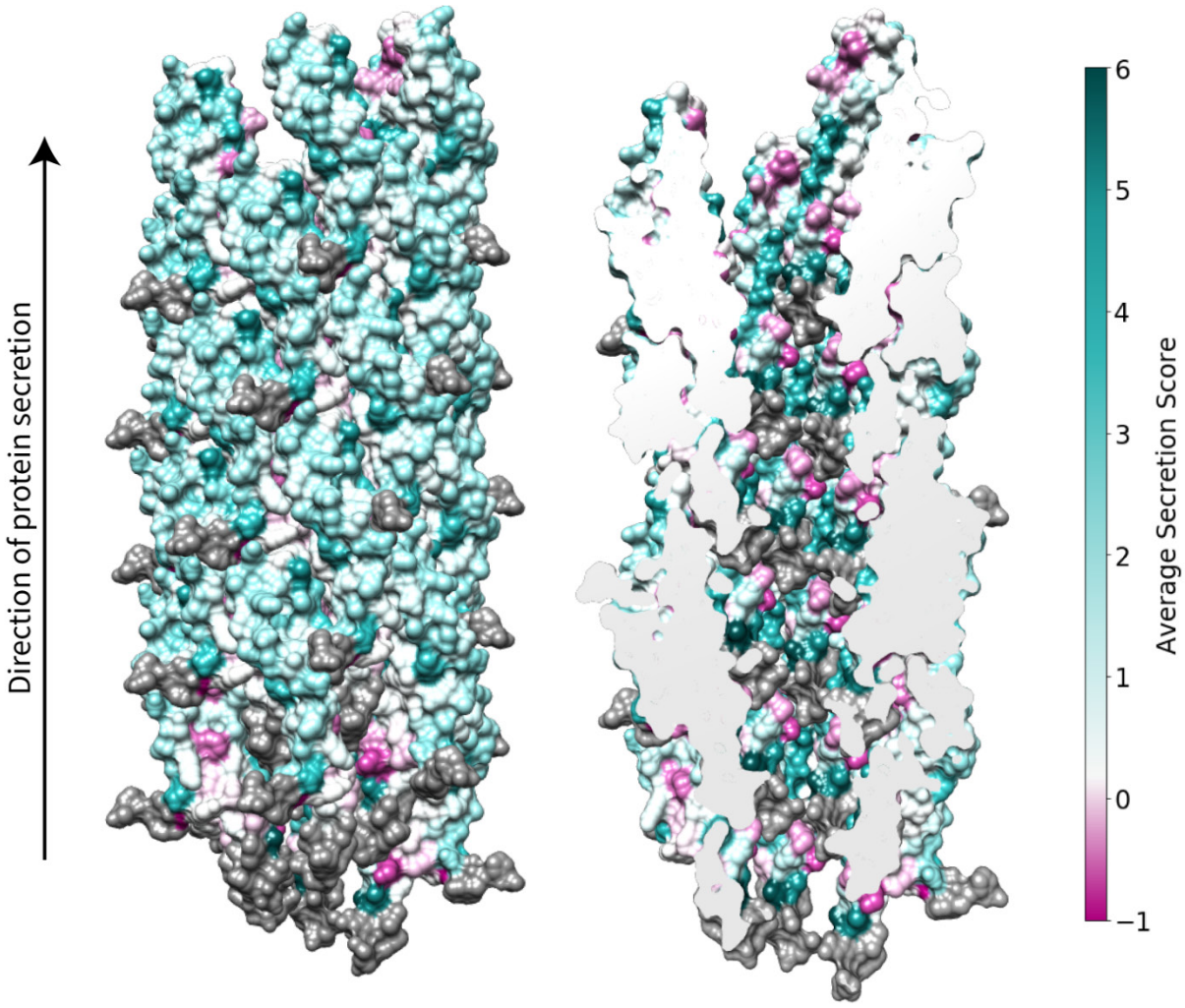


Figure 5.16 Interior and exterior residues show different patterns for secretion fitness.
Secretion scores were averaged across all substitutions at each residue and mapped on to the PrgI needle structure (PDB = 6dwb). Dark grey denotes residues that were not modified in the library design.

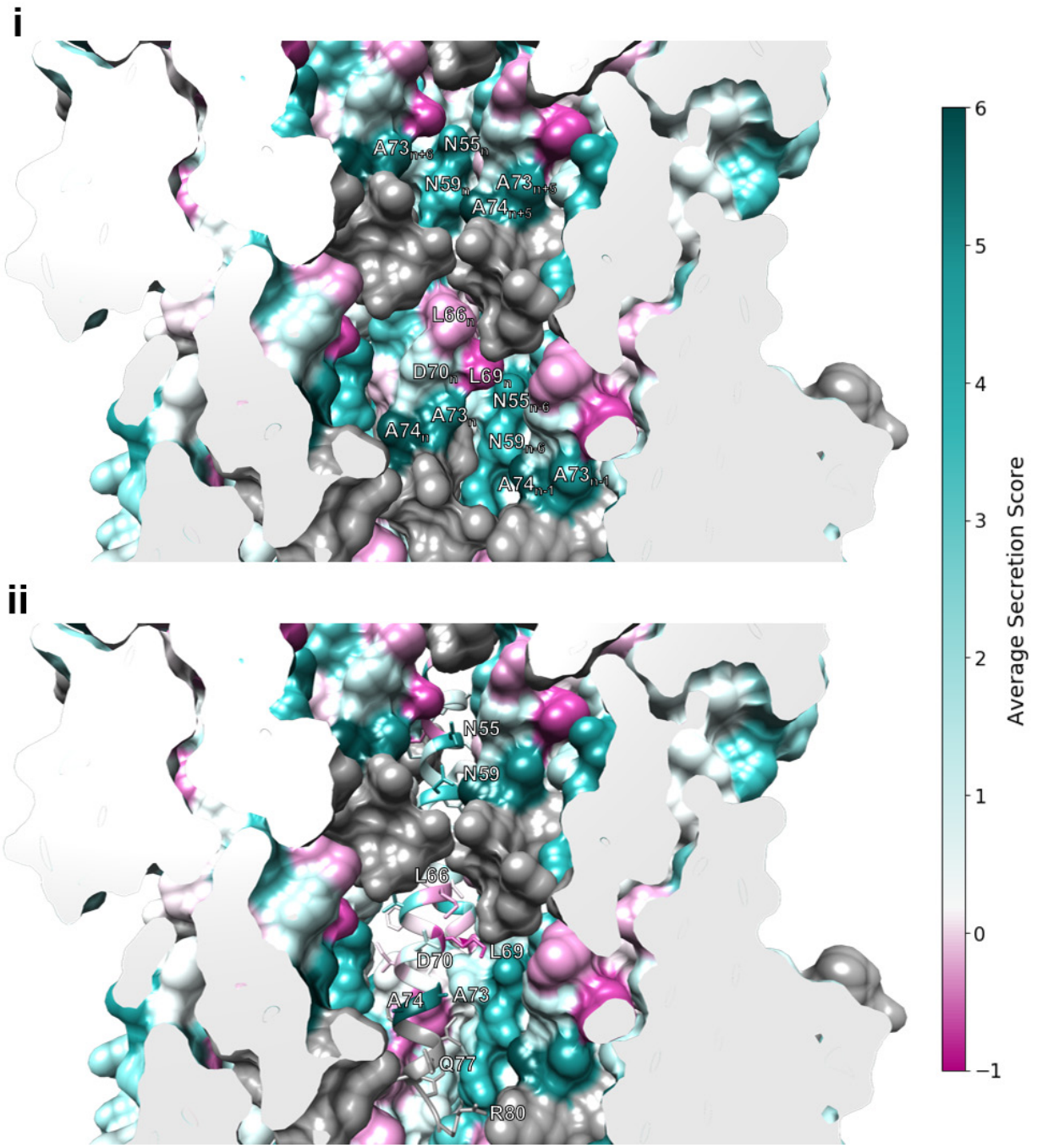


Figure 5.17 Substitution with large hydrophobic amino acids increased secretion titer at residues that line the hydrophobic groove of the needle interior.
The hydrophobic groove in the needle interior is composed of alternating N55, N59, A73, and A74 from the indicated chains (*i*). Residues L66, L69, Q77, and R80 form the charged, raised groove. Residue 70 contributes to the charged raised groove with its native aspartic acid but also tolerated several amino acid substitutions. A ribbon colored by average secretion score shows the predicted orientations of each amino acid (*ii*, PDB = 6dwb).

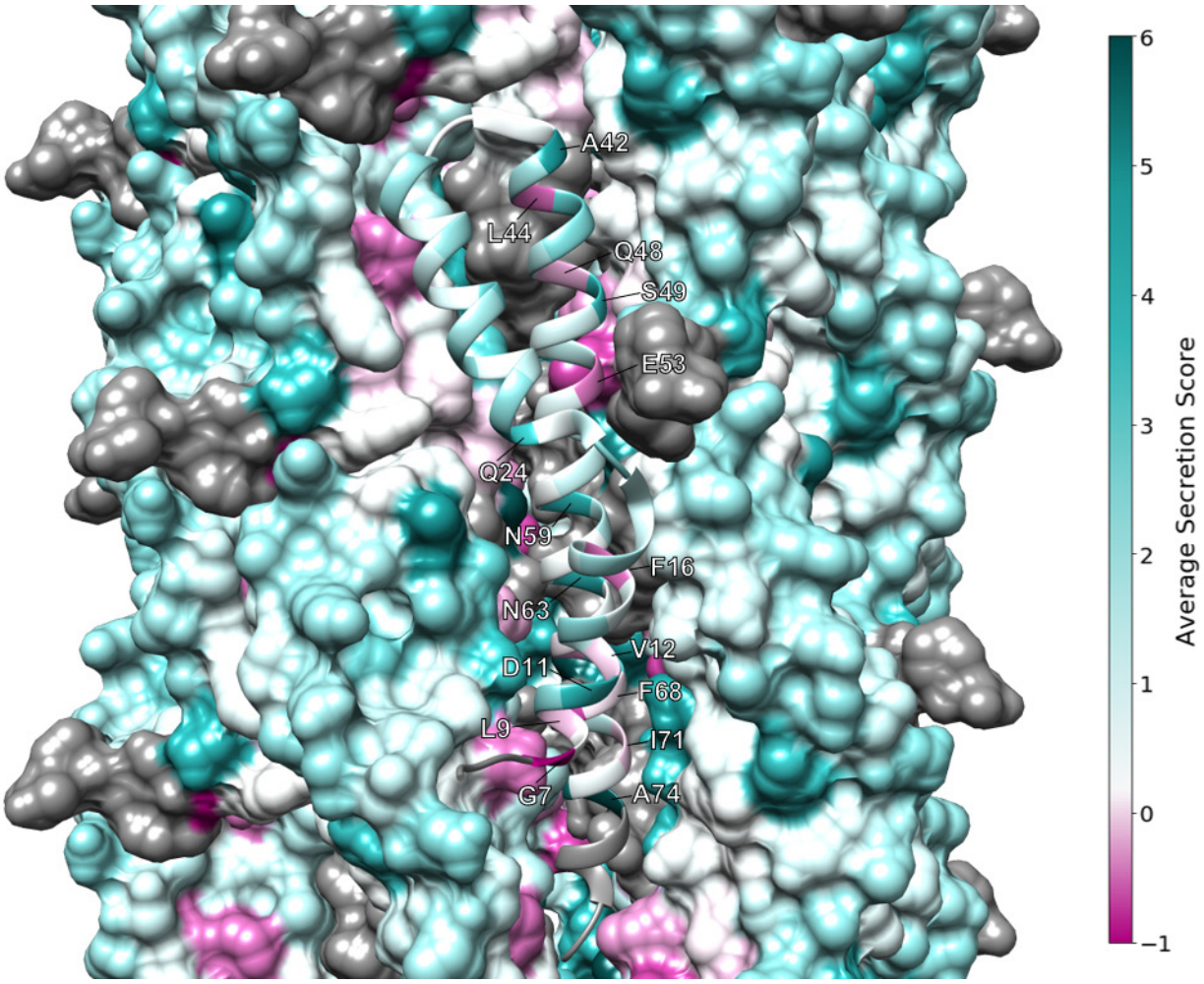


Figure 5.18 Substitutions at residues not facing the interior of the needle likely affect secretion titer by changing inter- and intrachain interactions.
A ribbon view of average secretion score per residue from the exterior of the needle (PDB = 6dwb). Residues with high or low average secretion scores are labeled.

5.3.5 The secretion fitness landscape predicted high-secreting variants

We validated our SFL and library screening assay by repeating secretion titer measurement via alkaline phosphatase activity for a random assortment of variants from pools A-I and all variants from pool J. Relative secretion titer corresponded well (R^2=0.87) with the weighted-average secretion score (Figure 5.19A). Though the re-analyzed pool J variants had generally lower relative secretion titers than the values from the library screen, all were at least 1.5-fold above $PrgI^{WT}$ (Figure 5.19B).

The variants in pool J fell into the two classes observed for high average secretion scores—those that face the interior of the needle apparatus, and those that involve contact with neighboring residues. Multiple substitutions at residues 55, 59, 73, and 74 appeared in pool J, and most were larger and more hydrophobic than the native amino acid. The remaining variants likely changed interactions with neighboring residues, and those substitutions were of varying types. Most amino acid changes were to larger, more

hydrophobic residues, but S49R and A74Q did not follow the pattern of hydrophobic substitutions.

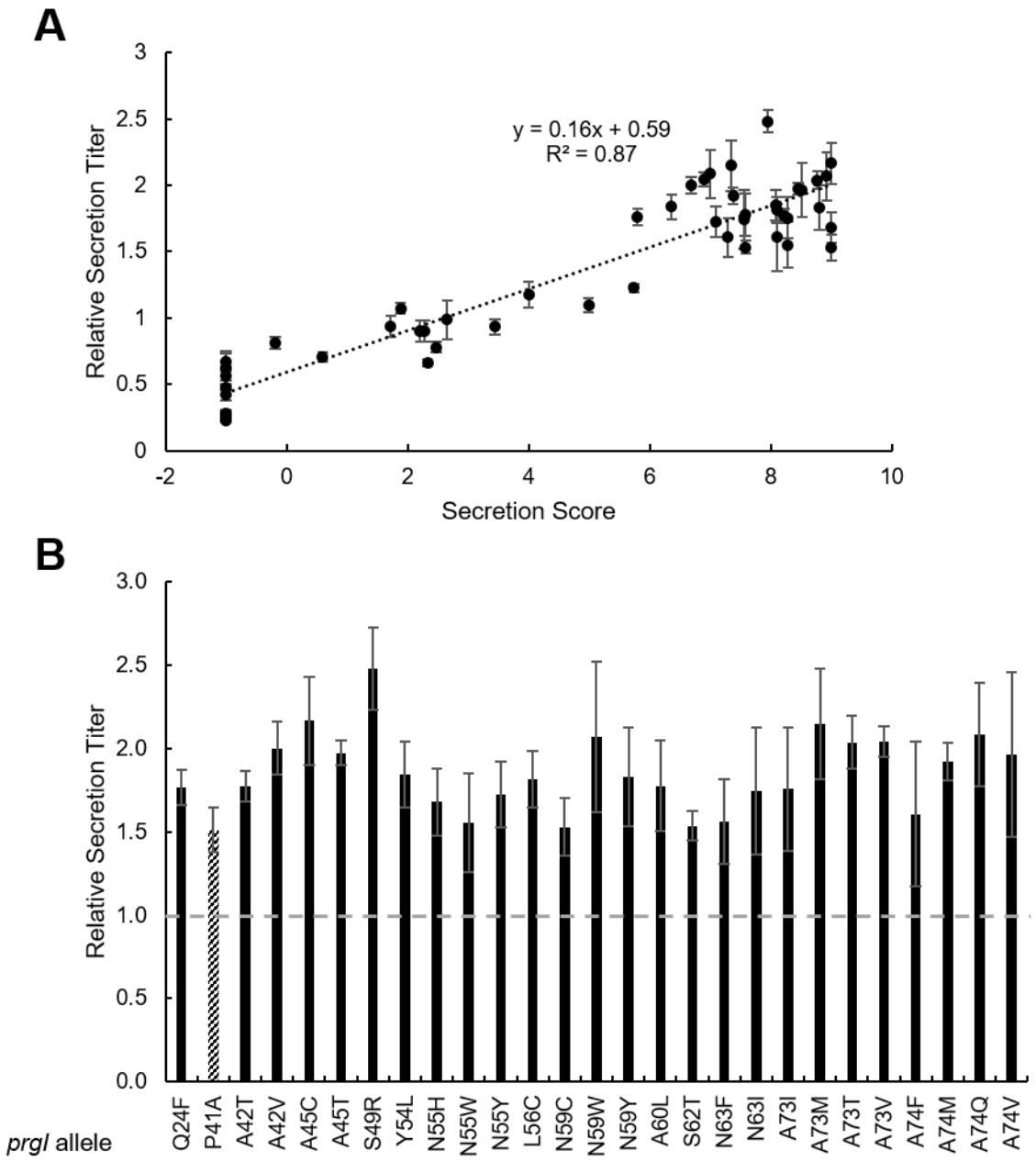


Figure 5.19 Weighted-average secretion scores predict relative secretion titer.
Randomly selected clones from pools A-I and all clones from pool J were patched onto LB-agar plates from the reformatted glycerol stocks. Patched colonies were inoculated for secretion titer measurement, which followed the same workflow as library screening. Clones were Sanger sequenced, and the secretion score was plotted against the newly measured secretion titer (**A**). Error bars represent standard error of three biological replicates. Clones from pool J were plotted separately to highlight differences (**B**). WT-level secretion titer is indicated by a dotted line. PrgI^{P41A} (stripes) was included for comparison as the previous best secreting variant. Error bars represent one standard deviation to allow direct comparison among variants.

5.3.6 Double mutants confer increased secretion titer

The data processing workflow for the high throughput sequencing data deliberately discarded sequences with more than one amino acid change because multiple mutations were assumed to be sequencing or sample preparation errors. Sanger sequencing revealed that pool J contained several double mutants, however (Figure 5.20). We hypothesize that the double mutants arose from assembling the library via recombineering, though it is possible they arose from errors in gene block synthesis. Experimental comparison of the secretion titers from the single and double mutants is necessary to confirm the patterns observed in Figure 5.20, but the preliminary data suggests that an epistatic evaluation of mutant combinations would provide further insight into the design rules of the PrgI needle.

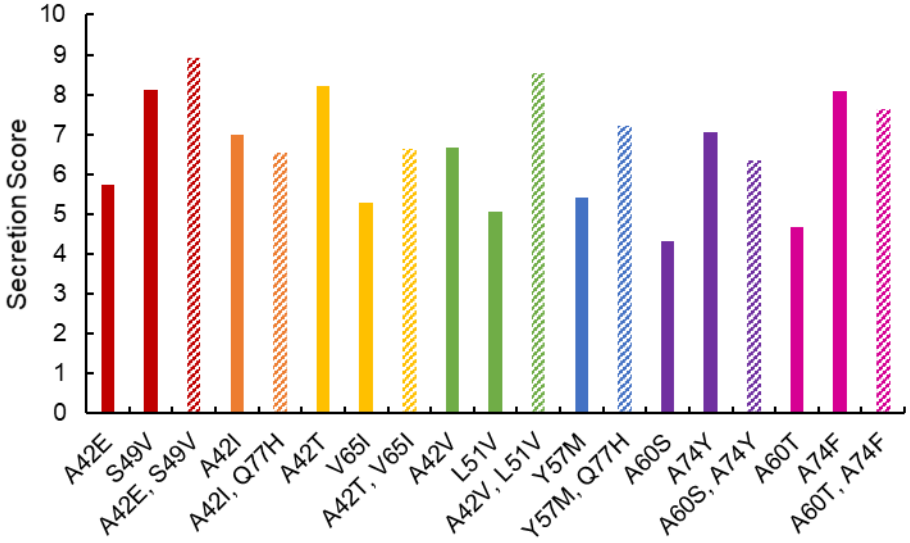


Figure 5.20 Double mutants suggest epistatic effects.
Secretion scores for double mutants were calculated from the correlation between relative secretion titer and secretion score in Figure 5.19A and compared to secretion scores for each of the individual mutations.

5.4 Discussion

5.4.1 Secretion fitness of substitutions varies by helix

The average secretion score per substituted amino acid confirmed that larger and more hydrophobic amino acids increased secretion titer (Figure 5.21). Evaluating average secretion score per substituted residue in each of the N- and C-terminal helices revealed that the high secretion scores of threonine, valine, isoleucine, and methionine were primarily influenced by substitutions in the C-terminal helix (Figure 5.22).

Proline and charged amino acids were poor substitutions for secretion function in the total structure, though positively charged residues were preferred over negatively charged residues. The low average secretion score for substitutions of negatively charged amino acids was contributed primarily by the C-terminal helix (Figure 5.22). This is likely a result of preserving the specific pattern of alternating charges on the raised edge of the hydrophobic groove in the interior of the needle.

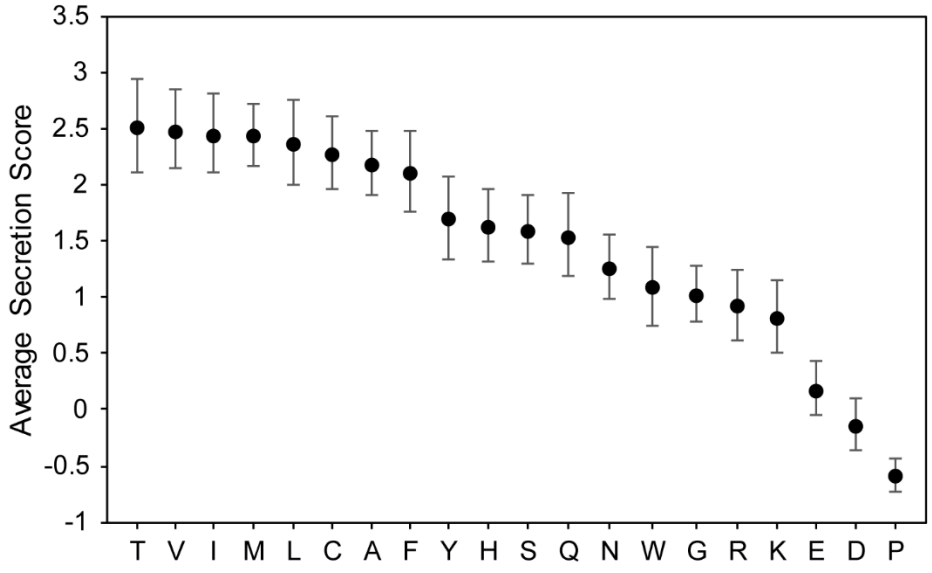


Figure 5.21 Average secretion score by amino acid substitution.
Average secretion score and standard error for each native amino acid mutated to the specified substitution.

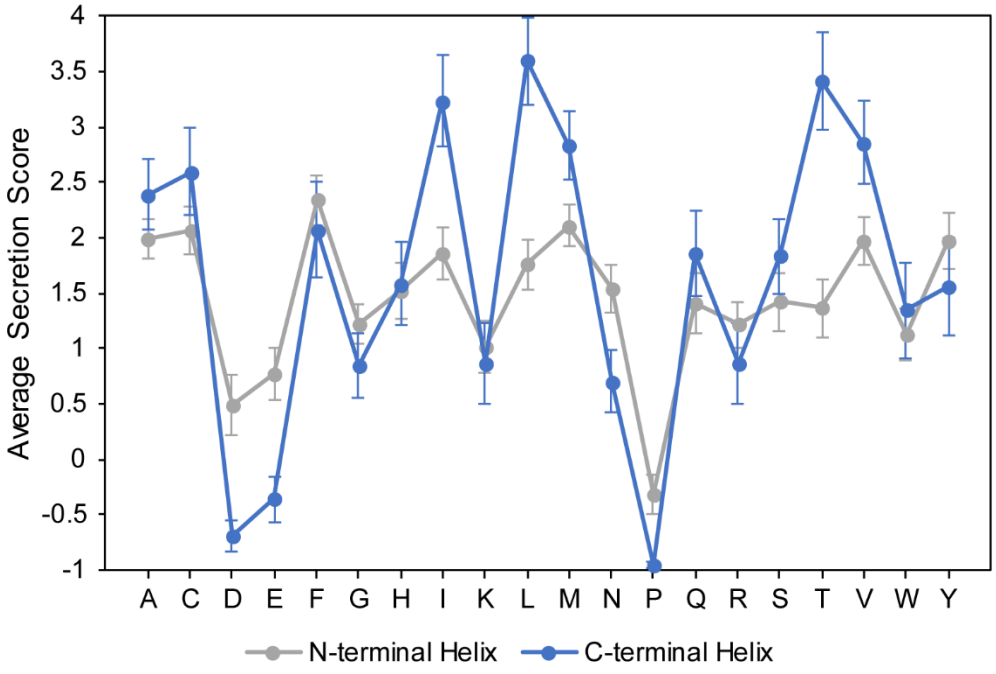


Figure 5.22 Average secretion score of substituted residues differs in the N- and C-terminal helices.
Average secretion score and standard error for each native amino acid mutated to the specified substitution in the N-terminal helix, C-terminal helix, and total monomer.

5.4.2 Outstanding questions

How does secretion fitness relate to native T3SS fitness?

The secretion fitness landscape of PrgI does not provide a roadmap for mutations that enhance native T3SS fitness. Rather, it might provide a roadmap for mutations that decrease native T3SS fitness. The native secretion apparatus did not evolve to secrete a maximum amount of each native substrate—secretion titers of native substrates via the T3SS are tightly controlled in its native context [134,136]. A recent alanine scan of PrgI revealed that mutations that increased secretion of native effectors produced variable invasion levels of *S. enterica* Typhimurium into human intestine epithelial cells, indicating that some mutations interrupted the native program of the T3SS [186]. The high average secretion scores of cysteine and methionine underline this difference, as the native PrgI lacks cysteine, methionine, histidine, and tryptophan.

Are PrgI variants that increase secretion titer general?

The adage “you get what you screen for” is true in another sense—secretion titer was measured only with SptP-tagged alkaline phosphatase. The AP assay development for the library screen showed that the genomic copy of SptP-AP-2xFLAG-6xHis provided similar relative secretion titers as plasmid-based SptP-DH-2xFLAG-6xHis in the PrgI position 41 mutants, but the generality of the secretion increase conferred by the new mutants uncovered in the library screen should be confirmed experimentally with several test proteins. The SptP tag is another important consideration. Several of the PrgI variants with increased secretion titer increased the volume and hydrophobicity of the amino acids that line the groove in the needle interior. Presumably, both the secretion tag and the heterologous protein interact with these residues as they are secreted. Thus, the various secretion tags known to facilitate heterologous protein secretion via the SPI-1 T3SS [8] might respond differently to changes in the needle structure.

Can substitutions at the highly conserved C-terminal residues of PrgI increase secretion titer?

The substitution Q77H was an interesting revelation from the double mutants in pool J because it should not have appeared in the library at all, yet it was part of two double mutations with high secretion titers. Q77 contributes to the charged, raised edge of the groove on the interior of the needle (Figure 5.17). A few of the residues present in the raised edge were included in the library, and they allowed limited substitutions. The positively charged residues in that raised edge, K66 and K69, allowed only leucine and arginine substitutions, and arginine appeared to increase secretion titer. Another residue in that raised edge, D70, permitted several substitutions, but none increased secretion titer. The C-terminal residues should be included in future mutational studies, as this work shows that substitutions in the C-terminal residues can result in functional secretion apparatuses and even increase titers if the overall interior structure is preserved.

What is the mechanistic explanation for the increase in secretion titer?

The recent alanine scan of PrgI lent credence to the hypothesis that amino acid substitutions that alter inter- and intrachain interactions change the conformation and packing of needle monomers. Galán and colleagues reported that PrgI^{S49A} changed the

electrostatic surface in the interior of the needle and proposed that the effect was a result of changes in the orientations side chains that face the lumen [186]. Performing similar structural analyses on the high-secreting variants from this library will reveal the structural variation caused by the amino acid changes.

Secretion titer reports on both needle assembly and average secretion rate, as secretion was performed over the same time frame for all mutants. Thus, mutants with increased secretion titer have a proportional increase in average secretion rate. The maximum increase in secretion titer observed in this work was two- to threefold, which is less than the tenfold difference in secretion rate between activated and passive secretion [134,181]. Single amino acid changes might not be sufficient to achieve a “fully activated” needle conformation, then, or the fully activated state might arise from a confluence of factors. Though we are abstracting the T3SS from its native context in engineering it for heterologous protein secretion, we hope to increase the secretion rate to that achieved in the fully active conformation and therefore increase secretion titer further by combining the results of this library screen with the optimized media from the previous chapters and the strain improvements described in the next chapter.

5.5 Acknowledgments

I extend my thanks to Han Teng Wong, who did the initial quality trimming of the high throughput sequencing reads. Many thanks to Kevin Metcalf and Elias Valdivia for laying the groundwork for this project by developing the recombineering protocol and mapping the MxiH mutations to PrgI. Thank you to Sara Fernandez Dunne and the High Throughput Analysis Core at Northwestern for use of and support for the DigiLab HiGro shakers and the Tecan Fluent used for library reformatting. I would also like to thank Emily Hartman, Nolan Kennedy, Bon Ikwuagwu, Stephanie Robinson, Nolan Kennedy, and Daniel Brauer for helpful discussions on processing and visualizing the fitness landscape data. I would be remiss if I did not also thank all the DTE lab members that helped me decide on the color scheme for the secretion scores.

Chapter 6: Strain Optimization for Use of the SPI-1 T3SS as a Protein Production Platform

Portions of the following were adapted from Tullman Ercek, D., Burdette L., Wong, H.T. RECOMBINANT STRAINS AND MEDIUM FORMULATION FOR ENHANCING SECRETION TITER USING A TYPE III SECRETION SYSTEM. U.S. Patent Application 62/793,226, 2019 with permission.

6.1 Introduction

The use of yeast and bacteria for production of consumer goods can be traced back to early civilizations. Bread, cheese, beer, wine, and yogurt are all produced by specialized strains of yeast or bacteria [192]. Advancements in biotechnology and the genetic tractability of microbial hosts have led to the development of "workhorse" strains, commonly *E. coli* and *Saccharomyces cerevisiae*, for production of chemicals and proteins [193,194]. Similar to traditional synthetic chemical processes, cell-based production of chemicals and proteins requires process optimization to maximize product yield. The cellular reactors are optimized by engineering the specific chemical pathways involved in synthesizing the product, but they have also been "domesticated" over time, purposefully and by natural evolution in a production environment, to remove unnecessary cellular processes and make them more efficient [90,195,196].

The utility of the SPI-1 T3SS as a heterologous protein production platform is clear from the preceding chapters and the wide range of heterologous proteins it can secrete. *S. enterica* Typhimurium is not an optimized protein production strain, however. The complexity of SPI-1 T3SS regulation complicates moving the system to a production strain of *E. coli* [65,92,98]. Thus, engineering *S. enterica* Typhimurium itself to remove pathogenic and non-essential cellular processes will create a safe, efficient strain for heterologous protein secretion in bacteria.

Another barrier to using the SPI-1 T3SS as a protein production platform is the two-plasmid system required to maximize protein secretion titers. One plasmid carries the heterologous protein product with an N-terminal secretion tag under the control of a promoter that participates in the activation signal cascade for the T3SS, and the second is a synthetically inducible plasmid for *hilA* overexpression [41]. Activating the T3SS via plasmid-based synthetic overexpression of *hilA* is undesirable for two reasons—first, multiple plasmids increase the complexity of the system and require cellular resources for maintenance, and second, an expensive small molecule, IPTG, must be added to the culture to induce *hilA* overexpression. Engineering the strain to achieve similar levels of T3SS induction without plasmid-based *hilA* overexpression would create a more efficient production system.

In this chapter, I describe the effect of several non-essential gene knockouts on heterologous protein expression and secretion and identify genomic modifications that mimic the effects of synthetic *hilA* overexpression. Individual gene knockouts of several pathogenic operons and native SPI-1 substrates had minimal effects on protein expression and secretion titer. The SPI-1 regulator HilD acts upstream of *hilA*, so we hypothesized that genomic modifications that increased HilD expression would have the same effect as synthetic *hilA* overexpression. Indeed, deleting the HilD repressor HilE

and inserting genes between the *hilD* coding sequence and its 3' UTR achieved similar levels of heterologous protein expression and secretion as synthetic *hilA* overexpression. In combination with the LB-(KH_2PO_4/K_2HPO_4) medium with glycerol described in Chapter 2, the engineered strain produced higher secretion titers than those achieved with synthetic *hilA* overexpression.

6.2 Methods

6.2.1 Strains and Growth Conditions

Strains and plasmids used in this work are listed in Table 6.1 and Table 6.2. Secretion experiments were started by growing a single colony from a fresh streak of a frozen glycerol stock in the lysogeny broth Lennox formulation (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) with appropriate antibiotics (34 µg/mL chloramphenicol for the plasmid expressing the secreted protein, 50 µg/mL kanamycin for the P_{lacUV5} *hilA* overexpression vector) for 12-16 hours overnight in an orbital shaker at 37°C and 225 rpm. Overnight cultures were diluted 1:100 into the appropriate medium supplemented with appropriate antibiotics and 100 µg/mL isopropyl β-D-1-thiogalactopyranoside (IPTG) if the strain carried P_{lacUV5} *hilA.* All culturing steps were performed in 24-well deepwell plates (Axygen). Secretion was performed for 8 hours at 37°C and 225 rpm in an orbital shaker. Whole culture lysate samples for SDS-PAGE were prepared by adding cell suspension to Laemmli buffer [150] in a 1:2 ratio at the end of secretion. The secretion fraction was harvested by centrifuging cultures at 4000 *x g* for 10 minutes. SDS-PAGE samples for the secretion fraction were prepared by adding supernatant to Laemmli buffer in a 3:1 ratio. All SDS-PAGE samples were boiled at 95°C for 5 minutes immediately after preparation.

Table 6.1 Strains used in Chapter 6.

Strain Name	Comment	Reference
ASTE13	LT2-derived lab strain similar to DW01	This study; DW01 [21]
ASTE13 *hilE*	*hilE* knockout	This study
ASTE13 *hilD::GFPmut2*	*GFPmut2* inserted immediately downstream of *hilD* coding sequence	This study
ASTE 13 *hilD::mCherry*	*mCherry* inserted immediately downstream of *hilD* coding sequence	This study
HASTE15g	*hilE* knockout; *GFPmut2* inserted immediately downstream of *hilD* coding sequence	This study
HASTE15c	*hilE* knockout; *mCherry* inserted immediately downstream of *hilD* coding sequence	This study

Table 6.2 Plasmids used in Chapter 6.

Plasmid Name	ORFs under inducible control		ORI	ab^R	Reference
P_{sic} *DH*	*sicP*	*sptP-DH-2xFLAG-6xHis*	colE1	cam	[22]
P_{sic} *Bla*	*sicP*	*sptP-bla-2xFLAG-6xHis*	colE1	cam	[22]
P_{lacUV5} *hilA*	*hilA*		p15a	kan	[22]

6.2.2 Recombineering

Recombineering was performed in *S. enterica* Typhimurium ASTE13 as described by Thomason *et al.* [152]. Briefly, a *cat-sacB* cassette conferring chloramphenicol resistance and sucrose sensitivity was amplified using primers with 40 bp of homology 5' and 3' to the locus of interest. Genes for GFPmut2 and mCherry were amplified using primers containing the same 40 bp of homology 5' and 3' to the locus of interest as used for the *cat-sacB* cassette. PCR was performed with Phusion DNA polymerase and the primers listed in Table 6.3. *S. enterica* Typhimurium ASTE13 was first transformed with pSIM6. A first round of recombineering was performed to insert the *cat-sacB* cassette at the locus of interest, and a second round of recombineering replaced the *cat-sacB* cassette with *GFPmut2* or *mCherry* for transcriptional fusions or a 60 bp oligo containing the first and last 30 bp of *hilE* for the *hilE* gene knockout. The genomic modifications were confirmed by Sanger sequencing (Quintara), and the strains were cured of pSIM6.

Table 6.3 Primers used in Chapter 6.

Sequence	Used to Construct This Strain
AACTACGCCATCGACATTCATAAAAATGGCGAACCATTAAAT TAAAGAGGAGAAAGGTCATGAG	ASTE13 *hilD::GFPmut2*
TTAATAAAAATCTTTACTTAAGTGACAGATACAAAAAATGTTA TTTGTATAGTTCATCCATGCCATG	ASTE13 *hilD::GFPmut2*
AACTACGCCATCGACATTCATAAAAATGGCGAACCATTAAAT TAAAGAGGAGAAAGGTCATGGTTTCCAAGGGCG	ASTE13 *hilD::mCherry*
TTAATAAAAATCTTTACTTAAGTGACAGATACAAAAAATGTTA TTTGTACAGCTCATCCATGC	ASTE13 *hilD::mCherry*
AACTACGCCATCGACATTCATAAAAATGGCGAACCATTAATG TGACGGAAGATCACTTCG	ASTE13 *hilD::GFPmut2,* ASTE13 *hilD::mCherry*
ATAAAAATCTTTACTTAAGTGACAGATACAAAAAATGATCAAA GGGAAAACTGTCCATAT	ASTE13 *hilD::GFPmut2,* ASTE13 *hilD::mCherry*
ACGAAATGGCTGGAAAATGGAACGTTCTTTCATTGTTGGCTG TGACGGAAGATCACTTCG	ASTE13 *hilE*
GTCCTCATCGCCACAGCGCCTGTCGGTGAAGAGGCCGCCA TCAAAGGGAAAACTGTCCAT	ASTE13 *hilE*
ATGGCTGGAAAATGGAACGTTCTTTCATTGTTGGCGGCGGC CTCTTCACCGACAGGCGCTGTGGCGATGA	ASTE13 *hilE*

6.2.3 Protein Separation and Western Blotting

Samples were separated by SDS-PAGE and transferred to a polyvinylidene fluoride membrane (PVDF, Millipore) for western blotting using the Bio-Rad Criterion blotter. Membranes were probed with mouse anti-FLAG per manufacturer's instructions (Sigma Aldrich). To facilitate chemiluminescent detection, a secondary labeling step was performed with goat anti-mouse IgG (H+L) HRP conjugate according to manufacturer's instructions (Thermo Fisher). Bands were detected with the SuperSignal West Pico Plus substrate (Thermo Fisher) and a ChemiDoc XRS + imaging system (Bio-Rad). All relative protein quantities were calculated by performing densitometry using Image Lab software (Bio-Rad) and normalizing to the specified condition.

6.2.4 Flow Cytometry

Strains were grown and induced as specified in "Strains and Growth Conditions". Samples were prepared by diluting cultures to an optical density at 600 nm of 0.005 to 0.05 in PBS with 1 mg/mL kanamycin in sterile round bottom tubes (Fisher #14-956-3B) and stored at 4°C for analysis. For each sample, an Attune NxT flow cytometer (Life Technologies) was used to collect 10,000 events within a gated population determined to be cells. Data was analyzed using FlowJo 10.5.3 (TreeStar, Inc.).

6.3 Results

6.3.1 SPI-1 effector knockouts do not alter secretion titer

Deleting native substrates of the SPI-1 T3SS (effectors) achieves two goals simultaneously: first, they are incapable of executing their pathogenic function, and second, their absence increases the purity of the secreted protein. Though no literature evidence suggested that knocking out non-structural native SPI-1 T3SS substrates would change secretion behavior, it was important to evaluate the effect of each knockout individually. Fortunately, of the seven effector knockouts tested, none had a significant impact on secretion titer of SptP-DH-2xFLAG-6xHis (Figure 6.1). All of the knockouts increased relative expression by a small margin.

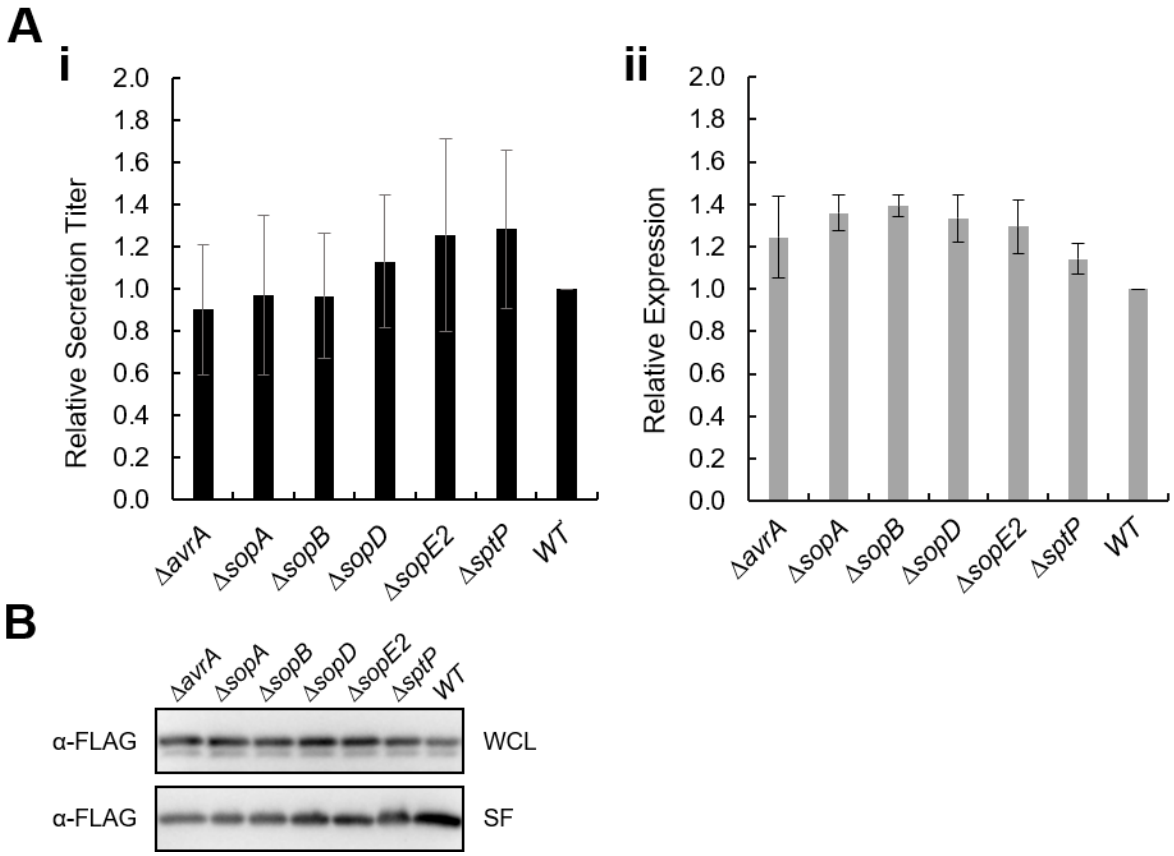


Figure 6.1 Relative expression and secretion titer for individual SPI-1 effector knockouts.
A Relative bulk secretion titer (*i*) and expression (*ii*) of SptP-DH-2xFLAG-6xHis in ASTE13 WT and SPI-1 effector knockouts in LB-L with *hilA* overexpression. Expression and secretion titer were measured via semi-quantitative western blotting and normalized to ASTE13 WT. Error bars represent standard error. **B** Western blots are representative of three biological replicates.

6.3.2 A T3SS tip complex knockout increases secretion titer twofold

Chapter 2 and previous work showed that Δ*sipD* strains increase heterologous secretion titer via the SPI-1 T3SS [52], but the impact of deleting the remaining tip complex members SipB and SipC and the secreted effector SipA was unknown. In combination with *hilA* overexpression, the individual knockout strains Δ*sipB*, Δ*sipC*, and Δ*sipD* had the same effect of increased secretion titer, but they were not additive (Figure 6.2). We were surprised to observe that the individual and combined knockouts of SipB, SipC, and SipD produced similar heterologous protein expression and secretion titer in strains with and without the *hilA* overexpression vector.

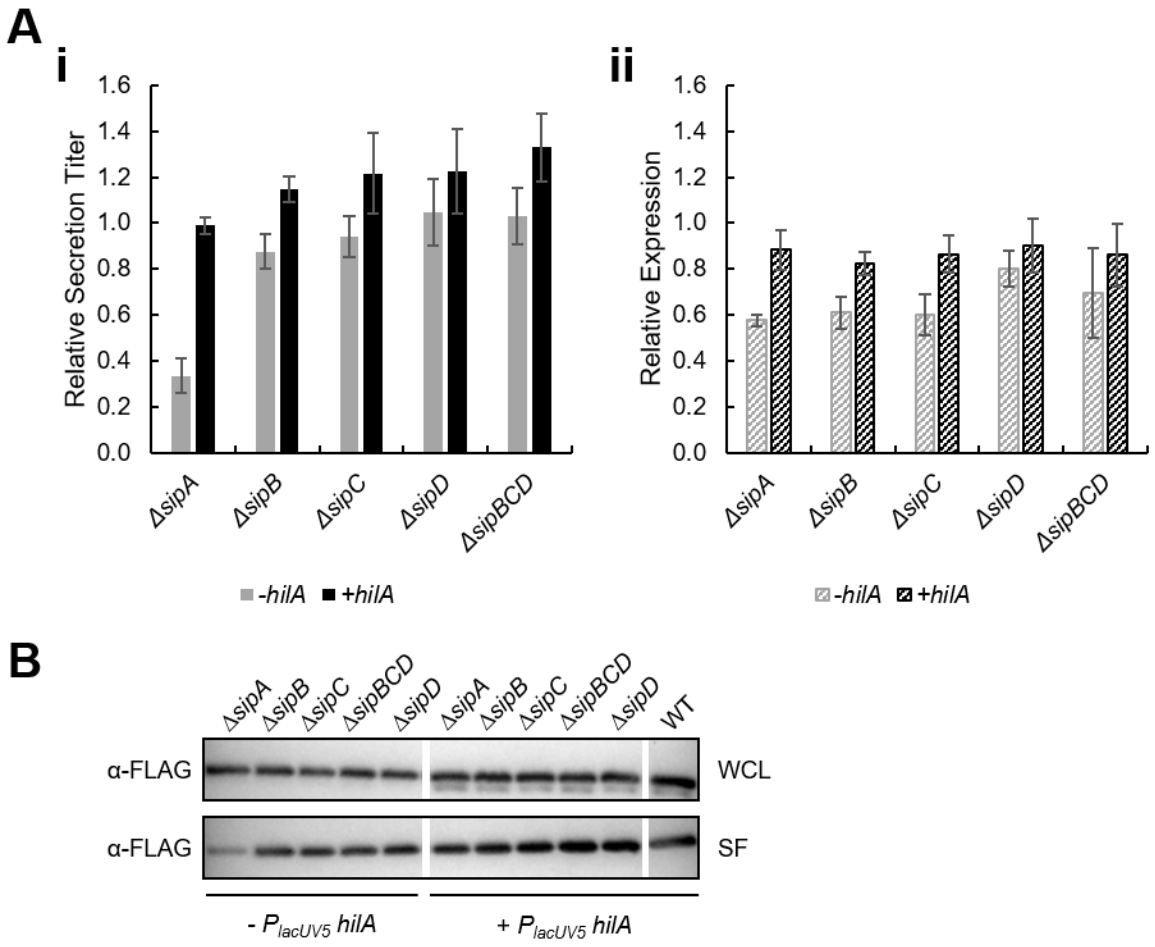


Figure 6.2 Relative expression and secretion titer for *sip* knockouts.
A Relative bulk secretion titer (*i*) and expression (*ii*) of SptP-DH-GFP11-2xFLAG-6xHis in ASTE13 WT and the listed *sip* knockouts in LB-L with *hilA* overexpression. Expression and secretion titer were measured via semi-quantitative western blotting and normalized to ASTE13 WT. Error bars represent standard error. **B** Western blots are representative of three biological replicates.

6.3.3 Non-essential pathogenic gene knockouts do not alter secretion titer

The genome of *Salmonella enterica* Typhimurium contains many operons that facilitate its pathogenic program, and these genes should be removed to create a non-pathogenic organism and improve heterologous protein production by freeing cellular resources (Table 1.7). As with the native substrates of the SPI-1 T3SS, it was important to evaluate the effects of removing those operons individually. Fortunately, the impact of knocking out the flagellar master regulators FlhDC, the type I fimbriae-producing *fim* operon, and *Salmonella* pathogenicity islands 4 and 5 (SPI-4 and SPI-5) was minimal (Figure 6.3). No significant differences in expression were observed among the knockout strains. The Δ*flhDC* and Δ*SPI-5* strains showed a 10-20% decrease in secretion titer, but that will likely be negligible in combination with other strain and process improvements.

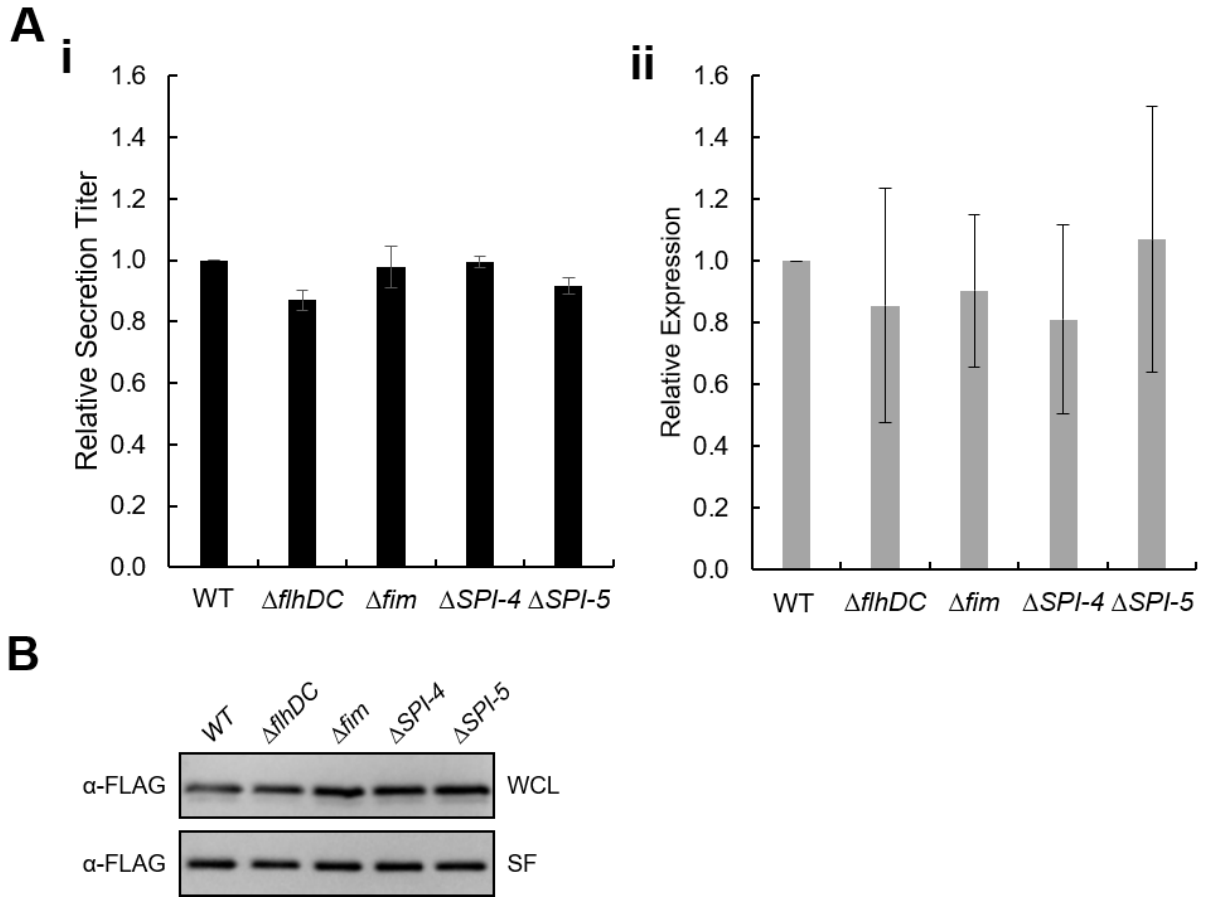


Figure 6.3 Relative expression and secretion titer for non-essential pathogenic gene knockouts. **A** Relative bulk secretion titer (*i*) and expression (*ii*) of SptP-DH-2xFLAG-6xHis in ASTE13 WT and the specified gene knockouts in TB with *hilA* overexpression. Expression and secretion titer were measured via semi-quantitative western blotting and normalized to ASTE13 WT. Error bars represent one standard deviation. **B** Western blots are representative of three biological replicates.

6.3.4 Methods to overexpress *hilD* confer increased secretion

SPI-1 T3SS activation is facilitated directly by HilA and its activator, HilD [97,197,198]. The HilD regulator integrates numerous environmental signals to ensure that the secretion system is activated via HilA only when it senses the appropriate environmental conditions (Figure 1.3). HilA activation in turn triggers a transcriptional activation cascade that leads to a functional T3SS and secretion. While knowledge of the regulatory network is ever-expanding, it is highly complex and integrated with essential cellular functions, which makes it difficult to engineer directly.

We focused on engineering expression of *hilD*, as more points of control are known for *hilD* than *hilA*. First we targeted HilE, which is known to reduce *hilA* activation by binding to HilD. Previous studies showed that *ΔhilE* strains confer increased *hilA* activity [101,199]. López-Garrido ***et al.*** demonstrated that HilD activity and therefore *hilA* activation is also controlled by *hilD* mRNA stability via its 3' UTR [200]. In that study, eliminating a portion of the *hilD* 3' UTR increased transcript stability and secretion system activation. We found, however, that inserting a β-barrel fluorescent protein such as GFP or mCherry at the start of the *hilD* 3'UTR (i.e., immediately after the coding sequence of *hilD*) and leaving the 3' UTR otherwise intact significantly increased secretion titer (Figure 6.4A).

Secretion titer increased ~2.5-fold in the *hilD::GFP* strain and threefold in the *hilE* and *hilD::mCherry* strains compared to the wild type strain without *hilA* overexpression. The combination of *hilE* and *hilD::GFP* provided a threefold increase, while the *hilE* and *hilD::mCherry* strain yielded the highest secretion titer, a ~3.2-fold increase. Flow cytometry performed at the endpoint of the experiment showed that adding the *hilE* knockout to the *hilD::GFP* or *hilD::mCherry* strains increased *hilD* expression (Figure 6.4C, Figure 6.6). The enhanced strains carrying only the plasmid expressing the heterologous protein product achieved secretion titers at 80-90% of the state-of-the-art system two-plasmid system.

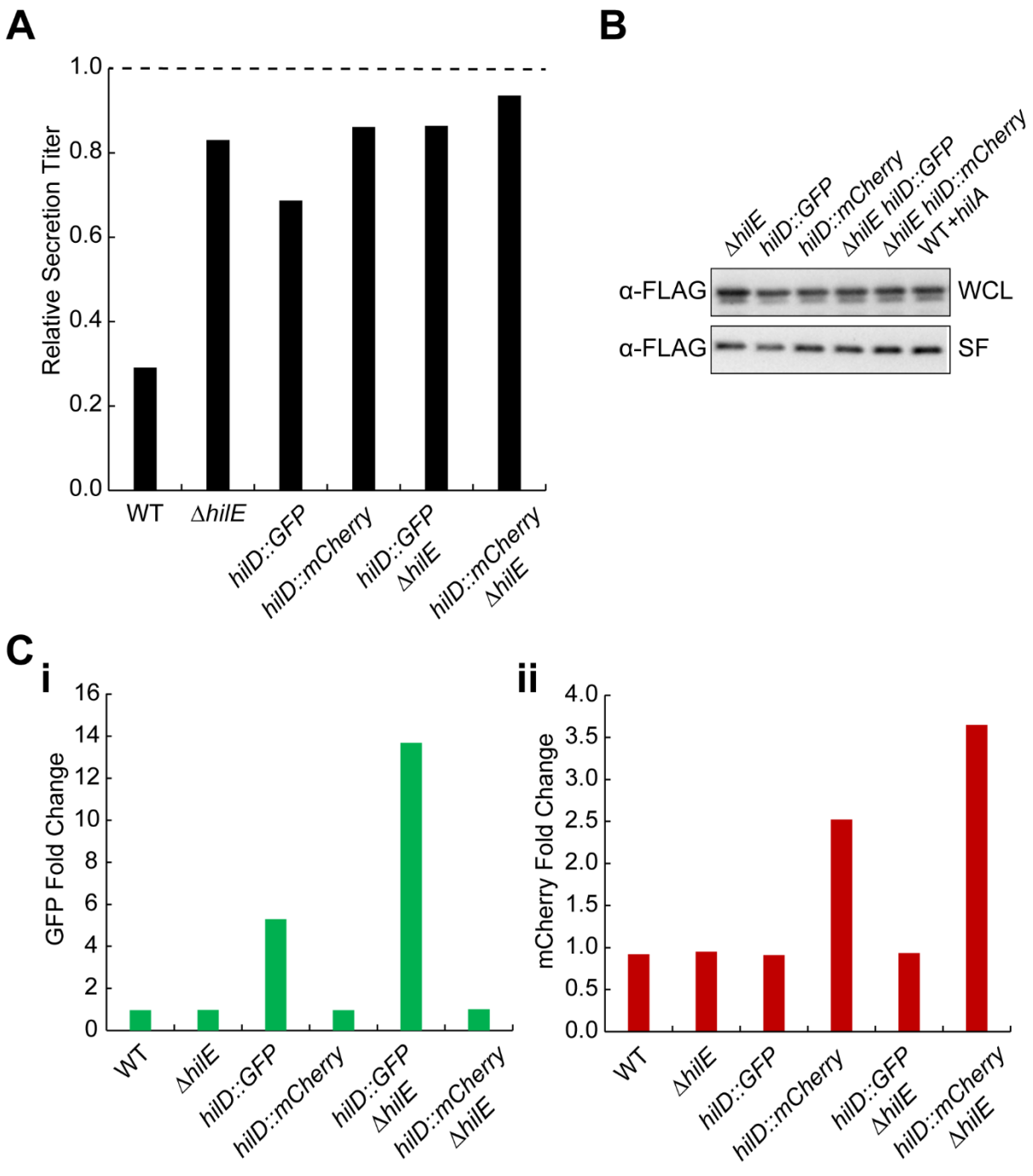


Figure 6.4 Secretion titer and transcriptional activity of genomically modified strains compared to the state-of-the-art system.
Densitometry and fluorescence values were normalized to ASTE13 WT in LB-L with *hilA* overexpression. Results represent a single biological replicate. **A** Relative secretion titer of SptP-DH-2xFLAG-6xHis by densitometry. ASTE13 in LB-L with *hilA* overexpression is represented as a dashed line. **B** Western blots used for densitometry (SF) and that depict total protein expression (WCL). "SF" is secreted fraction and "WCL" is whole culture lysate. **C** Fold change of fluorescence geometric mean for GFP and mCherry according to flow cytometry.

6.3.5 Secretion titer is maximized by overexpressing *hilD* in media optimized for heterologous protein secretion

While the enhanced strains, HASTE15g and HASTE15c (Table 6.1), and an optimized medium, LB with 0.4% w/v glycerol and 90 mM potassium phosphate pH 7.4, separately increased the utility of the T3SS for protein production, they had a more than additive effect when combined (Figure 6.5A). An approximately fourfold increase in secretion titer was observed with the enhanced strain in the optimized medium compared to a threefold increase in the optimized medium observed with the state-of-the-art two-plasmid system. Flow cytometry performed at the endpoint of the experiment showed that the combination of the modified strains and the optimized medium further increased *hilD* expression (Figure 6.5C, Figure 6.6).

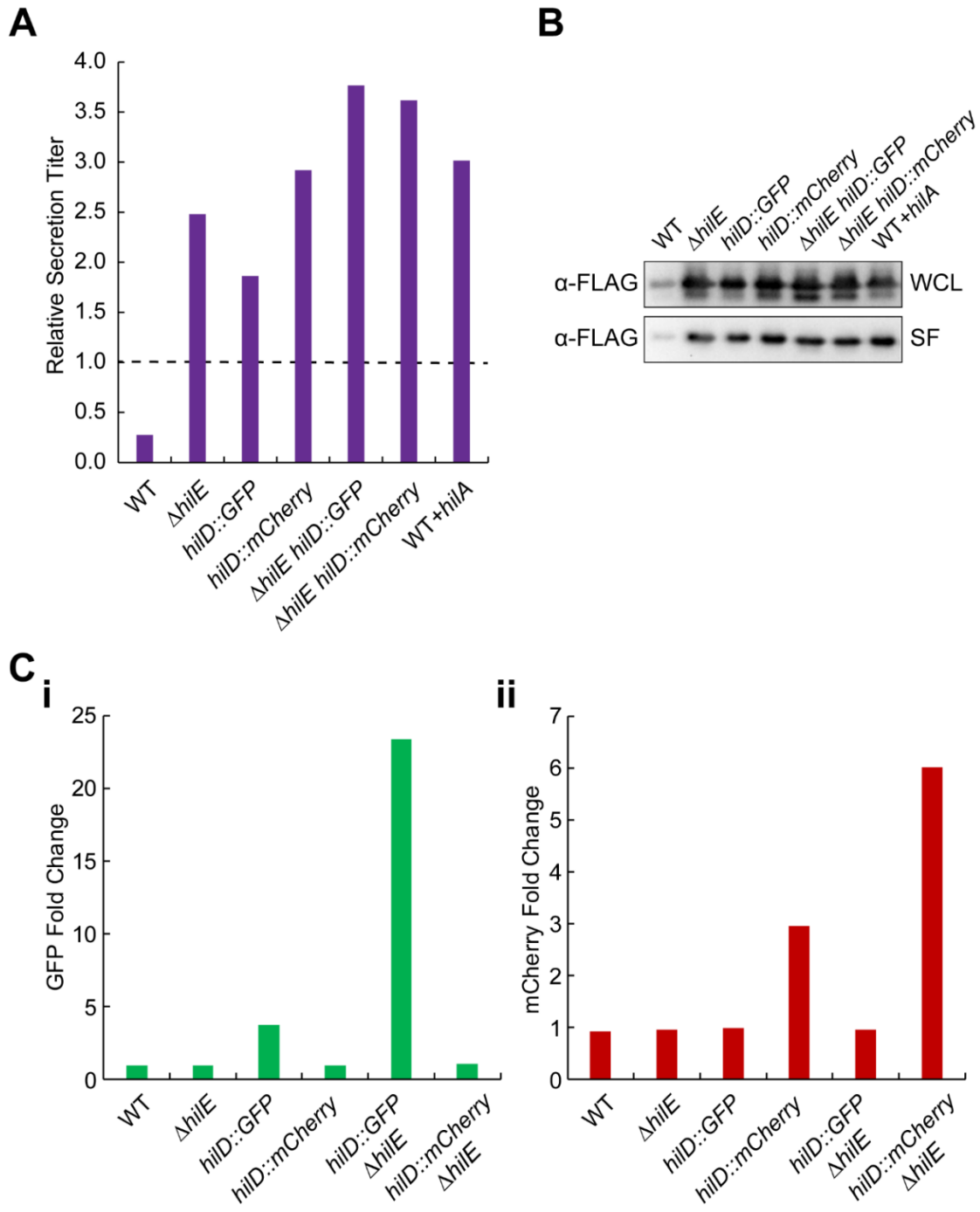


Figure 6.5 Secretion using the enhanced strains in an optimized medium.
A Secretion titer of SptP-DH-2xFLAG-6xHis from the genomically modified strains was normalized to ASTE13 WT in LB-L with *hilA* overexpression by performing densitometry on western blots. “WT” denotes a genetically wild type strain without the *hilA* overexpression plasmid. **B** Western blots used for densitometry (SF) and that depict total protein expression (WCL). “SF” is secreted fraction and “WCL” is whole culture lysate. **C** Fold change of fluorescence geometric mean for GFP and mCherry according to flow cytometry. Results represent a single biological replicate.

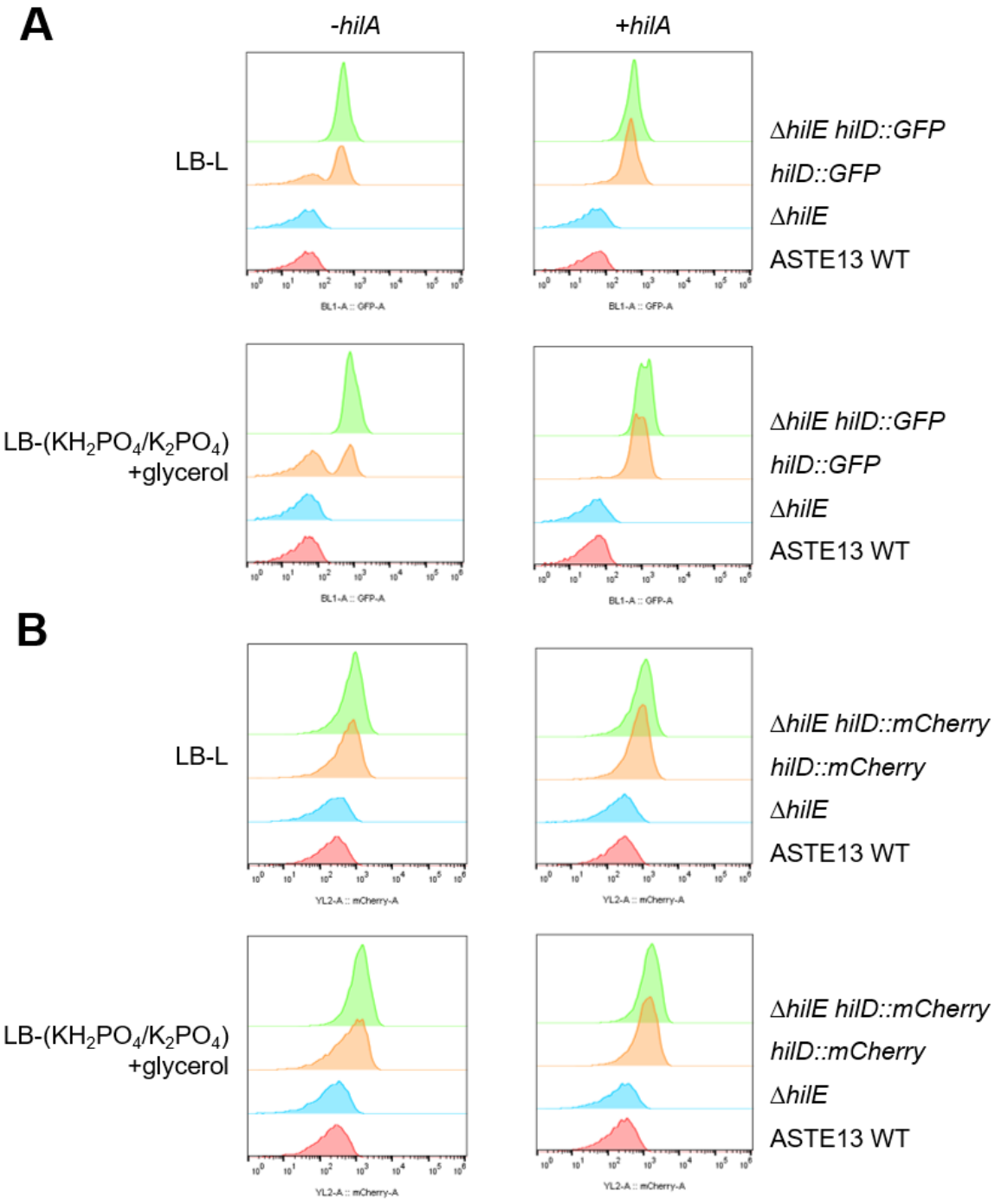


Figure 6.6 Histograms for flow cytometry of strains in Figures 6.4 and 6.5.
A GFP fluorescence and **B** mCherry fluorescence of engineered strains in LB-L and the optimized medium after an eight-hour secretion experiment with and without *hilA* overexpression.

6.4 Discussion

Strain optimization is essential to maximize the yield of any biomolecule, but it is especially important for heterologous protein secretion via the SPI-1 T3SS in order to create an efficient, non-pathogenic production platform. The minimal effect on heterologous protein secretion from many of the individual gene knockouts shown here suggests that combining those knockouts would have a similarly minimal impact. The increase in secretion titer resulting from a knockout of the proteins that compose the tip complex, Δ*sipBCD*, is promising for both increased secretion titer and reduced pathogenicity—*S. enterica* Typhimurium strains lacking those proteins are incompetent for invasion into mammalian cells [201].

Genomic modifications to effect HilD overexpression in combination with a growth medium optimized for maximal secretion titer increased secretion titer above that produced by the state-of-the-art synthetic *hilA* overexpression system. Flow cytometry performed on the optimized strains at the end of the secretion experiment indicated that the combination of Δ*hilE* and *hilD::GFP* or *hilD::mCherry* maximized secretion titer by increasing the fraction of the population with *hilD* overexpression (Figure 6.6). The results will have to be replicated to confirm the statistical significance of the differences in secretion titer, and the specific effect of the genetic insertion between the *hilD* coding sequence and 3' UTR should be investigated to find the optimal configuration. The optimized strain should also be tested with multiple secreted proteins, as the results of Chapter 2 suggest that optimal T3SS activity is not its maximum for some secreted proteins such as SptP-14B7*.

The results of this chapter advance the development of the SPI-1 T3SS as a heterologous protein production platform, but the work is unfinished. In addition to the experiments outlined for the HASTE15g and HASTE15c strains, the remaining pathogenicity islands and operons from Table 1.7 should be tested individually and in combination to ensure that they have no effect or a positive effect on heterologous protein secretion. The preliminary data shown in this chapter are promising—they show that it is possible to create a minimal, non-pathogenic, high-secreting *S. enterica* Typhimurium strain for heterologous protein production.

6.5 Acknowledgments

I am indebted to several undergraduate students for much of the work presented in this chapter. Jinan Oubaid, a SynBREU student, produced the data for individual SPI-1 effector knockouts. Emily Chen, a Northwestern undergraduate student, created the *sipA*, *sipB*, *sipC*, and *sipBCD* knockouts and performed preliminary secretion experiments. Finally, Aliki Valdes, another SynBREU student, produced the data for the *flhDC*, *fim*, SPI-4, and SPI-5 knockouts. I would also like to thank Han Teng Wong for initially constructing the *hilD::GFP* and *hilD::mCherry* strains and performing preliminary experiments that suggested their positive effect on heterologous protein expression and secretion.

Chapter 7: Conclusion

Bacterial secretion systems have the potential to dramatically decrease the cost of recombinant protein production by decreasing production timelines and simplifying the downstream processing required to collect a pure product. The work in this thesis expands the engineering space for increased protein secretion via the SPI-1 T3SS and lays the foundation for creating a minimal, robust heterologous protein production platform. Much work remains to realize this goal, however.

Aside from high productivity, the most important feature of a cellular factory is safety. Strains employed for protein or chemical production at scales of even tens of liters should have no negative ecological or medical impacts. Ideally, given its pathogenic origin, an engineered SPI-1 T3SS would be moved to a production strain of *E. coli*. If those efforts fail, however, the results of Chapter 6 indicate that it is possible to create a safe, minimal, high-secreting strain based on *S. enterica* Typhimurium.

Most of the SPI-1 T3SS engineering advances are presented in isolation, but an ideal production system would include a combination of those factors. The combinations of optimized media and strain modifications presented in Chapters 2 and 6 suggest that at least some improvements are additive. As additional improvements are discovered, a continuous, systematic evaluation of the collective effect of strain modifications, medium designs, and growth conditions will be necessary to create an optimal heterologous protein secretion system. A high-throughput assay for measuring expression and secretion titer is critical for this endeavor.

The knowledge gained from the work presented here is applicable beyond simply increasing secretion titer. The design rules for growth media, hypersecreting PrgI variants, and strain improvements provide insight into the native functions and properties of the T3SS. New drug targets and alternative uses of the T3SS could be revealed through synthetic manipulation of the system. Though some properties are specific to the SPI-1 T3SS, many are applicable to homologous T3SS systems. Engineering improvements could be applied to other systems that are compatible with a non-pathogenic production host (like mT3 *E. coli* [58]) to provide an alternative to engineering the SPI-1 T3SS in situ.

Even if they are not adopted at a production scale, using bacterial secretion systems to evaluate binding domains, biomaterial properties, or the efficacy of antimicrobial peptides would drastically decrease development timelines. Academic and industrial labs alike would benefit from inexpensive alternative production platforms for difficult-to-express proteins and new biotechnological targets. Shifting beyond the current paradigms of protein production could promote new technological innovation with higher purity, simpler systems, and lower cost.

References

1. Goeddel DV, Kleid DG, Bolivar F, Heyneker HL, Yansura DG, Crea R, *et al.* Expression in *Escherichia coli* of chemically synthesized genes for human insulin. PNAS. 1979;76:106–10.
2. Carlson R. Estimating the biotech sector's contribution to the US economy. Nat Biotechnol. 2016;34:247–55.
3. Sarrouh B, Santos TM, Miyoshi A, Dias R, Azevedo V. Up-To-Date Insight on Industrial Enzymes Applications and Global Market. Journal of Bioprocessing & Biotechniques. 2012;0.
4. Demain AL, Vaishnav P. Production of recombinant proteins by microbes and higher organisms. Biotechnology Advances. 2009;27:297–306.
5. Zorko M, Jerala R. Production of Recombinant Antimicrobial Peptides in Bacteria. Giuliani A, Rinaldi AC, editors. Antimicrobial Peptides: Methods and Protocols. 2010;61–76.
6. Roque ACA, Lowe CR, Taipa MÂ. Antibodies and Genetically Engineered Related Molecules: Production and Purification. Biotechnology Progress. 2004;20:639–54.
7. Lyons RE, Elvin CM, Taylor K, Lekieffre N, Ramshaw JAM. Purification of recombinant protein by cold-coacervation of fusion constructs incorporating resilin-inspired polypeptides. Biotechnol Bioeng. 2012;109:2947–54.
8. Widmaier DM, Tullman-Ercek D, Mirsky EA, Hill R, Govindarajan S, Minshull J, *et al.* Engineering the *Salmonella* type III secretion system to export spider silk monomers. Molecular Systems Biology. 2009;5:309.
9. Park Seujeung, Ramirez W. Fred. Dynamics of foreign protein secretion from *Saccharomyces cerevisiae*. Biotechnology and Bioengineering. 1989;33:272–81.
10. Kim JY, Kim Y-G, Lee GM. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl Microbiol Biotechnol. 2012;93:917–30.
11. Thomas P, Smart TG. HEK293 cell line: a vehicle for the expression of recombinant proteins. J Pharmacol Toxicol Methods. 2005;51:187–200.
12. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the *Pichia pastoris* expression system. Yeast. 2005;22:249–70.
13. Datar RV, Cartwright T, Rosen C-G. Process Economics of Animal Cell and Bacterial Fermentations: A Case Study Analysis of Tissue Plasminogen Activator. Bio/Technology. 1993;11:349.
14. Caunt P, Impoolsup A, Greenfield PF. Stability of recombinant plasmids in yeast. Journal of Biotechnology. 1988;8:173–92.
15. Baneyx F. Protein Expression Technologies: Current Status and Future Trends. Taylor & Francis; 2004.
16. Hatada Y, Hidaka Y, Nogi Y, Uchimura K, Katayama K, Li Z, *et al.* Hyper-production of an isomalto-dextranase of an Arthrobacter sp. by a proteases-deficient *Bacillus subtilis*: sequencing, properties, and crystallization of the recombinant enzyme. Appl Microbiol Biotechnol. 2004;65:583–92.
17. Kang Z, Yang S, Du G, Chen J. Molecular engineering of secretory machinery components for high-level secretion of proteins in *Bacillus* species. J Ind Microbiol Biotechnol. 2014;41:1599–607.

18. ÿztürk S, ÿalık P, ÿzdamar TH. Fed-Batch Biomolecule Production by *Bacillus subtilis* : A State of the Art Review. Trends in Biotechnology. 2016;34:329–45.
19. Park S, Schumann W. Optimization of the secretion pathway for heterologous proteins in *Bacillus subtilis*. Biotechnol Bioproc E. 2015;20:623–33.
20. Westers L, Westers H, Quax WJ. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. Biochim Biophys Acta. 2004;1694:299–310.
21. Anné J, Maldonado B, Van Impe J, Van Mellaert L, Bernaerts K. Recombinant protein production and streptomycetes. Journal of Biotechnology. 2012;158:159–67.
22. Anné J, Vrancken K, Van Mellaert L, Van Impe J, Bernaerts K. Protein secretion biotechnology in Gram-positive bacteria with special emphasis on *Streptomyces lividans*. Biochim Biophys Acta. 2014;1843:1750–61.
23. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol. 2014;5.
24. Reed B, Chen R. Biotechnological applications of bacterial protein secretion: from therapeutics to biofuel production. Research in Microbiology. 2013;164:675–82.
25. Wang A, Lu SD, Mark DF. Site-specific mutagenesis of the human interleukin-2 gene: structure-function analysis of the cysteine residues. Science. 1984;224:1431–3.
26. Lilie H, Schwarz E, Rudolph R. Advances in refolding of proteins produced in *E. coli*. Curr Opin Biotechnol. 1998;9:497–501.
27. Azam A, Li C, Metcalf KJ, Tullman-Ercek D. Type III secretion as a generalizable strategy for the production of full-length biopolymer-forming proteins. Biotechnology and Bioengineering. 2016;113:2313–20.
28. Baneyx F. Recombinant protein expression in *Escherichia coli.* Current Opinion in Biotechnology. 1999;10:411–21.
29. Dalbey RE, Kuhn A. Protein Traffic in Gram-negative bacteria – how exported and secreted proteins find their way. FEMS Microbiol Rev. 2012;36:1023–45.
30. Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, *et al.* Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nature Reviews Microbiology. 2015;13:343–59.
31. Natale P, Brüser T, Driessen AJM. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. Biochim Biophys Acta. 2008;1778:1735–56.
32. Prehna G, Zhang G, Gong X, Duszyk M, Okon M, McIntosh LP, *et al.* A Protein Export Pathway Involving *Escherichia coli* Porins. Structure. 2012;20:1154–66.
33. Zhang Z, Tang R, Zhu D, Wang W, Yi L, Ma L. Non-peptide guided auto-secretion of recombinant proteins by super-folder green fluorescent protein in *Escherichia coli*. Scientific Reports. 2017;7:6990.
34. Gao D, Wang S, Li H, Yu H, Qi Q. Identification of a heterologous cellulase and its N-terminus that can guide recombinant proteins out of *Escherichia coli*. Microbial Cell Factories. 2015;14:49.
35. Wurm DJ, Slouka C, Bosilj T, Herwig C, Spadiut O. How to trigger periplasmic release in recombinant *Escherichia coli*: A comparative analysis. Engineering in Life Sciences. 2017;17:215–22.

36. Chen Z-Y, Cao J, Xie L, Li X-F, Yu Z-H, Tong W-Y. Construction of leaky strains and extracellular production of exogenous proteins in recombinant *Escherichia coli*. Microbial Biotechnology. 2014;7:360–70.
37. Khosa S, Scholz R, Schwarz C, Trilling M, Hengel H, Jaeger K-E, *et al.* An A/U-Rich Enhancer Region Is Required for High-Level Protein Secretion through the HlyA Type I Secretion System. Appl Environ Microbiol. 2018;84:e01163-17.
38. Metcalf KJ, Bevington JL, Rosales SL, Burdette LA, Valdivia E, Tullman-Ercek D. Proteins adopt functionally active conformations after type III secretion. Microbial Cell Factories. 2016;15:213.
39. Sevastsyanovich YR, Leyton DL, Wells TJ, Wardius CA, Tveen-Jensen K, Morris FC, *et al.* A generalised module for the selective extracellular accumulation of recombinant proteins. Microbial Cell Factories. 2012;11:69.
40. Schwarz CKW, Landsberg CD, Lenders MHH, Smits SHJ, Schmitt L. Using an *E. coli* Type 1 secretion system to secrete the mammalian, intracellular protein IFABP in its active form. Journal of Biotechnology. 2012;159:155–61.
41. Metcalf KJ, Finnerty C, Azam A, Valdivia E, Tullman-Ercek D. Using Transcriptional Control To Increase Titers of Secreted Heterologous Proteins by the Type III Secretion System. Applied and Environmental Microbiology. 2014;80:5927–34.
42. Holland IB, Peherstorfer S, Kanonenberg K, Lenders M, Reimann S, Schmitt L. Type I Protein Secretion—Deceptively Simple yet with a Wide Range of Mechanistic Variability across the Family. EcoSal Plus. 2016;7.
43. Baneyx F, Mujacic M. Recombinant protein folding and misfolding in *Escherichia coli*. Nature Biotechnology. 2004;22:1399–408.
44. Missiakas D, Raina S. Protein folding in the bacterial periplasm. J Bacteriol. 1997;179:2465–71.
45. Goosens VJ, Monteferrante CG, van Dijl JM. The Tat system of Gram-positive bacteria. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2014;1843:1698–706.
46. Bakkes PJ, Jenewein S, Smits SHJ, Holland IB, Schmitt L. The Rate of Folding Dictates Substrate Secretion by the *Escherichia coli* Hemolysin Type 1 Secretion System. J Biol Chem. 2010;285:40573–80.
47. Freudl R. Signal peptides for recombinant protein secretion in bacterial expression systems. Microbial Cell Factories. 2018;17:52.
48. Nielsen H. Predicting Secretory Proteins with SignalP. Methods Mol Biol. 2017;1611:59–73.
49. Bendtsen JD, Nielsen H, Widdick D, Palmer T, Brunak S. Prediction of twin-arginine signal peptides. BMC Bioinformatics. 2005;6:167.
50. Dobó J, Varga J, Sajó R, Végh BM, Gál P, Závodszky P, *et al*. Application of a Short, Disordered N-Terminal Flagellin Segment, a Fully Functional Flagellar Type III Export Signal, to Expression of Secreted Proteins. Appl Environ Microbiol. 2010;76:891–9.
51. Eom GT, Oh JY, Park JH, Lim HJ, Lee SJ, Kim EY, *et al*. Alleviation of temperature-sensitive secretion defect of *Pseudomonas fluorescens* ATP-binding cassette (ABC) transporter, TliDEF, by a change of single amino acid in the ABC protein, TliD. Journal of Bioscience and Bioengineering. 2016;122:283–6.

52. Glasgow AA, Wong HT, Tullman-Ercek D. A Secretion-Amplification Role for *Salmonella enterica* Translocon Protein SipD. ACS Synth Biol. 2017;6:1006–15.
53. Low KO, Mahadi NM, Abdul Rahim R, Rabu A, Abu Bakar FD, Abdul Murad AM, *et al*. Enhanced secretory production of hemolysin-mediated cyclodextrin glucanotransferase in *Escherichia coli* by random mutagenesis of the ABC transporter system. Journal of Biotechnology. 2010;150:453–9.
54. Natarajan A, Haitjema CH, Lee R, Boock JT, DeLisa MP. An Engineered Survival-Selection Assay for Extracellular Protein Expression Uncovers Hypersecretory Phenotypes in *Escherichia coli*. ACS Synth Biol. 2017;6:875–83.
55. Haitjema CH, Boock JT, Natarajan A, Dominguez MA, Gardner JG, Keating DH, *et al.* Universal genetic assay for engineering extracellular protein expression. ACS Synth Biol. 2014;3:74–82.
56. Ernst NH, Reeves AZ, Ramseyer JE, Lesser CF. High-Throughput Screening of Type III Secretion Determinants Reveals a Major Chaperone-Independent Pathway. mBio. 2018;9:e01050-18.
57. Terpe K. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Applied Microbiology and Biotechnology. 2006;72:211–22.
58. Reeves AZ, Spears WE, Du J, Tan KY, Wagers AJ, Lesser CF. Engineering *Escherichia coli* into a Protein Delivery System for Mammalian Cells. ACS Synth Biol. 2015;4:644–54.
59. Heel T, Vogel GF, Lammirato A, Schneider R, Auer B. FlgM as a Secretion Moiety for the Development of an Inducible Type III Secretion System. PLOS ONE. 2013;8:e59034.
60. Browning DF, Richards KL, Peswani AR, Roobol J, Busby SJW, Robinson C. *Escherichia coli* “TatExpress” strains super-secrete human growth hormone into the bacterial periplasm by the Tat pathway. Biotechnol Bioeng. 2017;114:2828–36.
61. Eom GT, Lee SH, Oh YH, Choi JE, Park SJ, Song JK. Efficient extracellular production of type I secretion pathway-dependent *Pseudomonas fluorescens* lipase in recombinant *Escherichia coli* by heterologous ABC protein exporters. Biotechnol Lett. 2014;36:2037–42.
62. Eom GT, Oh JY, Song JK. High-level production of *Serratia proteamaculans* metalloprotease using a recombinant ABC protein exporter-mediated secretion system in *Pseudomonas fluorescens*. Process Biochemistry. 2014;49:1718–22.
63. Akeda Y, Kimura T, Yamasaki A, Kodama T, Iida T, Honda T, *et al*. Functional cloning of Vibrio parahaemolyticus type III secretion system 1 in *Escherichia coli* K-12 strain as a molecular syringe. Biochemical and Biophysical Research Communications. 2012;427:242–7.
64. Ruano-Gallego D, Álvarez B, Fernández LÁ. Engineering the Controlled Assembly of Filamentous Injectisomes in *E. coli* K-12 for Protein Translocation into Mammalian Cells. ACS Synth Biol. 2015;4:1030–41.
65. Song M, Sukovich DJ, Ciccarelli L, Mayr J, Fernandez-Rodriguez J, Mirsky EA, *et al.* Control of type III protein secretion using a minimal genetic system. Nature Communications. 2017;8:14737.

66. Ittig SJ, Schmutz C, Kasper CA, Amstutz M, Schmidt A, Sauteur L, *et al.* A bacterial type III secretion-based protein delivery tool for broad applications in cell biology. J Cell Biol. 2015;211:913–31.
67. Fernández LA, Sola I, Enjuanes L, Lorenzo V de. Specific Secretion of Active Single-Chain Fv Antibodies into the Supernatants of *Escherichia coli* Cultures by Use of the Hemolysin System. Appl Environ Microbiol. 2000;66:5024–9.
68. Low KO, Muhammad Mahadi N, Abdul Rahim R, Rabu A, Abu Bakar FD, Murad AMA, *et al.* An effective extracellular protein secretion by an ABC transporter system in *Escherichia coli*: statistical modeling and optimization of cyclodextrin glucanotransferase secretory production. J Ind Microbiol Biotechnol. 2011;38:1587–97.
69. Su L, Chen S, Yi L, Woodard RW, Chen J, Wu J. Extracellular overexpression of recombinant *Thermobifida fusca* cutinase by alpha-hemolysin secretion system in *E. coli* BL21(DE3). Microbial Cell Factories. 2012;11:8.
70. Eom GT, Oh JY, Park JH, Jegal J, Song JK. Secretory production of enzymatically active endo-β-1,4-mannanase from *Bacillus subtilis* by ABC exporter in *Escherichia coli*. Process Biochemistry. 2016;51:999–1005.
71. Lee YJ, Lee R, Lee SH, Yim SS, Jeong KJ. Enhanced secretion of recombinant proteins via signal recognition particle (SRP)-dependent secretion pathway by deletion of *rrsE* in *Escherichia coli*. Biotechnol Bioeng. 2016;113:2453–61.
72. Voulgaris I, Finka G, Uden M, Hoare M. Enhancing the selective extracellular location of a recombinant *E. coli* domain antibody by management of fermentation conditions. Appl Microbiol Biotechnol. 2015;99:8441–53.
73. Gao W, Yin J, Bao L, Wang Q, Hou S, Yue Y, *et al.* Engineering Extracellular Expression Systems in *Escherichia coli* Based on Transcriptome Analysis and Cell Growth State. ACS Synth Biol. 2018;7:1291–302.
74. Ling HL, Rahmat Z, Murad AMA, Mahadi NM, Illias RMd. Proteome-based identification of signal peptides for improved secretion of recombinant cyclomaltodextrin glucanotransferase in *Escherichia coli*. Process Biochemistry. 2017;61:47–55.
75. Zhang G, Brokx S, Weiner JH. Extracellular accumulation of recombinant proteins fused to the carrier protein YebF in *Escherichia coli*. Nature Biotechnology. 2006;24:100–4.
76. Enam F, Mansell TJ. Linkage-Specific Detection and Metabolism of Human Milk Oligosaccharides in *Escherichia coli*. Cell Chemical Biology. 2018;25:1292-1303.e4.
77. Qian Z-G, Xia X-X, Choi JH, Lee SY. Proteome-based identification of fusion partner for high-level extracellular production of recombinant proteins in *Escherichia coli*. Biotechnology and Bioengineering. 2008;101:587–601.
78. Kotzsch A, Vernet E, Hammarström M, Berthelsen J, Weigelt J, Gräslund S, *et al.* A secretory system for bacterial production of high-profile protein targets. Protein Sci. 2011;20:597–609.
79. Narayanan N, Khan M, Chou CP. Enhancing functional expression of heterologous lipase B in *Escherichia coli* by extracellular secretion. J Ind Microbiol Biotechnol. 2010;37:349–61.
80. Berger E, Crampton MC, Nxumalo NP, Louw ME. Extracellular secretion of a recombinant therapeutic peptide by *Bacillus halodurans* utilizing a modified flagellin type III secretion system. Microbial Cell Factories. 2011;10:62.

81. Guo S, Alshamy I, Hughes KT, Chevance FFV. Analysis of Factors That Affect FlgM-Dependent Type III Secretion for Protein Purification with *Salmonella enterica* Serovar Typhimurium. Journal of Bacteriology. 2014;196:2333–47.
82. Gerven NV, Goyal P, Vandenbussche G, Kerpel MD, Jonckheere W, Greve HD, *et al.* Secretion and functional display of fusion proteins through the curli biogenesis pathway. Molecular Microbiology. 2014;91:1022–35.
83. Wang M, Huang M, Zhang J, Ma Y, Li S, Wang J. A novel secretion and online-cleavage strategy for production of cecropin A in *Escherichia coli*. Sci Rep. 2017;7.
84. Zhong C, Gurry T, Cheng AA, Downey J, Deng Z, Stultz CM, *et al.* Self-Assembling Multi-Component Nanofibers for Strong Bioinspired Underwater Adhesives. Nat Nanotechnol. 2014;9:858–66.
85. Chatterjee S, Chaudhury S, McShan AC, Kaur K, De Guzman RN. Structure and Biophysics of Type III Secretion in Bacteria. Biochemistry. 2013;52:2508–17.
86. Büttner D. Protein Export According to Schedule: Architecture, Assembly, and Regulation of Type III Secretion Systems from Plant- and Animal-Pathogenic Bacteria. Microbiol Mol Biol Rev. 2012;76:262–310.
87. Majander K, Anton L, Antikainen J, Lång H, Brummer M, Korhonen TK, *et al.* Extracellular secretion of polypeptides using a modified *Escherichia coli* flagellar secretion apparatus. Nature Biotechnology. 2005;23:475–81.
88. Singer HM, Erhardt M, Steiner AM, Zhang M-M, Yoshikami D, Bulaj G, *et al.* Selective Purification of Recombinant Neuroactive Peptides Using the Flagellar Type III Secretion System. mBio. 2012;3:e00115-12.
89. Singer HM, Erhardt M, Hughes KT. Comparative analysis of the secretion capability of early and late flagellar type III secretion substrates. Mol Microbiol. 2014;93:505–20.
90. Marisch K, Bayer K, Scharl T, Mairhofer J, Krempl PM, Hummel K, *et al.* A Comparative Analysis of Industrial *Escherichia coli* K–12 and B Strains in High-Glucose Batch Cultivations on Process-, Transcriptome- and Proteome Level. PLOS ONE. 2013;8:e70516.
91. Bartra SS, Jackson MW, Ross JA, Plano GV. Calcium-Regulated Type III Secretion of Yop Proteins by an *Escherichia coli hha* Mutant Carrying a *Yersinia pestis* pCD1 Virulence Plasmid. Infect Immun. 2006;74:1381–6.
92. Cangelosi C, Hannagan S, Santiago CP, Wilson JW. Transfer of the cloned *Salmonella* SPI-1 type III secretion system and characterization of its expression mechanisms in Gram negative bacteria in comparison with cloned SPI-2. Microbiological Research. 2015;180:57–64.
93. Carleton HA, Lara-Tejero M, Liu X, Galán JE. Engineering the type III secretion system in non-replicating bacterial minicells for antigen delivery. Nature Communications. 2013;4:1590.
94. Bahrani FK, Sansonetti PJ, Parsot C. Secretion of Ipa proteins by *Shigella flexneri:* inducer molecules and kinetics of activation. Infect Immun. 1997;65:4005–10.
95. Riley G, Toma S. Detection of pathogenic *Yersinia enterocolitica* by using congo red-magnesium oxalate agar medium. J Clin Microbiol. 1989;27:213–4.
96. Lostroh CP, Lee CA. The *Salmonella* pathogenicity island-1 type III secretion system. Microbes Infect. 2001;3:1281–91.
97. Ellermeier CD, Ellermeier JR, Slauch JM. HilD, HilC and RtsA constitute a feed forward loop that controls expression of the SPI1 type three secretion system regulator

hilA in *Salmonella enterica* serovar Typhimurium. Molecular Microbiology. 2005;57:691–705.
98. Golubeva YA, Sadik AY, Ellermeier JR, Slauch JM. Integrating global regulatory input into the *Salmonella* pathogenicity island 1 type III secretion system. Genetics. 2012;190:79–90.
99. Lucas RL, Lostroh CP, DiRusso CC, Spector MP, Wanner BL, Lee CA. Multiple Factors Independently Regulate *hilA* and Invasion Gene Expression in *Salmonella enterica* Serovar Typhimurium. J Bacteriol. 2000;182:1872–82.
100. Bajaj V, Lucas RL, Hwang C, Lee CA. Co-ordinate regulation of *Salmonella* typhimurium invasion genes by environmental and regulatory factors is mediated by control of *hilA* expression. Molecular Microbiology. 1996;22:703–14.
101. Baxter MA, Jones BD. Two-Component Regulators Control *hilA* Expression by Controlling *fimZ* and *hilE* Expression within *Salmonella enterica* Serovar Typhimurium. Infect Immun. 2015;83:978–85.
102. Teplitski M, Goodier RI, Ahmer BMM. Catabolite repression of the SirA regulatory cascade in *Salmonella enterica*. International Journal of Medical Microbiology. 2006;296:449–66.
103. Martínez LC, Yakhnin H, Camacho MI, Georgellis D, Babitzke P, Puente JL, *et al.* Integration of a complex regulatory cascade involving the SirA/BarA and Csr global regulatory systems that controls expression of the *Salmonella* SPI-1 and SPI-2 virulence regulons through HilD. Molecular Microbiology. 2011;80:1637–56.
104. Lawhon SD, Maurer R, Suyemoto M, Altier C. Intestinal short-chain fatty acids alter *Salmonella* typhimurium invasion gene expression and virulence through BarA/SirA. Molecular Microbiology. 2002;46:1451–64.
105. Kim K, Golubeva YA, Vanderpool CK, Slauch JM. Oxygen-dependent regulation of SPI1 type three secretion system by small RNAs in *Salmonella enterica* serovar Typhimurium. Molecular Microbiology. 2019;111:570–87.
106. Mizusaki H, Takaya A, Yamamoto T, Aizawa S-I. Signal Pathway in Salt-Activated Expression of the *Salmonella* Pathogenicity Island 1 Type III Secretion System in *Salmonella enterica* Serovar Typhimurium. J Bacteriol. 2008;190:4624–31.
107. Ono S, Goldberg MD, Olsson T, Esposito D, Hinton JCD, Ladbury JE. H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. Biochemical Journal. 2005;391:203–13.
108. Duong N, Osborne S, Bustamante VH, Tomljenovic AM, Puente JL, Coombes BK. Thermosensing coordinates a cis-regulatory module for transcriptional activation of the intracellular virulence system in *Salmonella enterica* serovar Typhimurium. J Biol Chem. 2007;282:34077–84.
109. Zilkenat S, Franz-Wachtel M, Stierhof Y-D, Galan JE, Macek B, Wagner S. Determination of the stoichiometry of the complete bacterial type III secretion needle complex using a combined quantitative proteomic approach. Mol Cell Proteomics. 2016;mcp.M115.056598.
110. Sukhan A, Kubori T, Wilson J, Galan JE. Genetic Analysis of Assembly of the *Salmonella enterica* Serovar Typhimurium Type III Secretion-Associated Needle Complex. Journal of Bacteriology. 2001;183:1159–67.

111. Lara-Tejero M, Kato J, Wagner S, Liu X, Galán JE. A Sorting Platform Determines the Order of Protein Secretion in Bacterial Type III Systems. Science. 2011;331:1188–91.
112. Wee DH, Hughes KT. Molecular ruler determines needle length for the *Salmonella* Spi-1 injectisome. PNAS. 2015;112:4098–103.
113. Lefebre MD, Galán JE. The inner rod protein controls substrate switching and needle length in a *Salmonella* type III secretion system. Proc Natl Acad Sci USA. 2014;111:817–22.
114. Kato J, Dey S, Soto JE, Butan C, Wilkinson MC, De Guzman RN, *et al*. A protein secreted by the *Salmonella* type III secretion system controls needle filament assembly. Storz G, editor. eLife. 2018;7:e35886.
115. Kubori T, Galán JE. *Salmonella* Type III Secretion-Associated Protein InvE Controls Translocation of Effector Proteins into Host Cells. J Bacteriol. 2002;184:4699–708.
116. Monjarás Feria JV, Lefebre MD, Stierhof Y-D, Galán JE, Wagner S. Role of Autocleavage in the Function of a Type III Secretion Specificity Switch Protein in *Salmonella enterica* Serovar Typhimurium. mBio. 2015;6.
117. Schraidt O, Lefebre MD, Brunner MJ, Schmied WH, Schmidt A, Radics J, *et al.* Topology and Organization of the *Salmonella* typhimurium Type III Secretion Needle Complex Components. PLoS Pathog. 2010;6:e1000824.
118. Lilic M, Vujanac M, Stebbins CE. A Common Structural Motif in the Binding of Virulence Factors to Bacterial Secretion Chaperones. Molecular Cell. 2006;21:653–64.
119. Lee SH, Galán JE. *Salmonella* type III secretion-associated chaperones confer secretion-pathway specificity. Mol Microbiol. 2004;51:483–95.
120. Hardt W-D, Galán JE. A secreted *Salmonella* protein with homology to an avirulence determinant of plant pathogenic bacteria. PNAS. 1997;94:9887–92.
121. Rüssmann H, Kubori T, Sauer J, Galán JE. Molecular and functional analysis of the type III secretion signal of the *Salmonella enterica* InvJ protein. Molecular Microbiology. 2002;46:769–79.
122. Bronstein PA, Miao EA, Miller SI. InvB Is a Type III Secretion Chaperone Specific for SspA. J Bacteriol. 2000;182:6638–44.
123. Kaniga K, Tucker S, Trollinger D, Galán JE. Homologs of the *Shigella* IpaB and IpaC invasins are required for *Salmonella* typhimurium entry into cultured epithelial cells. J Bacteriol. 1995;177:3965–71.
124. Jang JI, Kim JS, Eom JS, Kim HG, Kim BH, Lim S, *et al.* Expression and delivery of tetanus toxin fragment C fused to the N-terminal domain of SipB enhances specific immune responses in mice. Microbiol Immunol. 2012;56:595–604.
125. Tucker SC, Galán JE. Complex Function for SicA, a *Salmonella enterica* Serovar Typhimurium Type III Secretion-Associated Chaperone. J Bacteriol. 2000;182:2262–8.
126. Kaniga K, Trollinger D, Galán JE. Identification of two targets of the type III protein secretion system encoded by the *inv* and *spa* loci of *Salmonella* typhimurium that have homology to the Shigella IpaD and IpaA proteins. J Bacteriol. 1995;177:7078–85.
127. Higashide W, Zhou D. The First 45 Amino Acids of SopA Are Necessary for InvB Binding and SPI-1 Secretion. J Bacteriol. 2006;188:2411–20.
128. Hong KH, Miller VL. Identification of a Novel *Salmonella* Invasion Locus Homologous to *Shigella ipgDE*. J Bacteriol. 1998;180:1793–802.

129. Knodler LA, Bertero M, Yip C, Strynadka NCJ, Steele-Mortimer O. Structure-based mutagenesis of SigE verifies the importance of hydrophobic and electrostatic residues in type III chaperone function. Mol Microbiol. 2006;62:928–40.
130. Wood MW, Williams C, Upadhyay A, Gill AC, Philippe DL, Galyov EE, *et al.* Structural analysis of *Salmonella enterica* effector protein SopD. Biochim Biophys Acta. 2004;1698:219–26.
131. Ehrbar K, Friebel A, Miller SI, Hardt W-D. Role of the *Salmonella* Pathogenicity Island 1 (SPI-1) Protein InvB in Type III Secretion of SopE and SopE2, Two *Salmonella* Effector Proteins Encoded Outside of SPI-1. J Bacteriol. 2003;185:6950–67.
132. Fu Y, Galán JE. Identification of a specific chaperone for SptP, a substrate of the centisome 63 type III secretion system of *Salmonella* typhimurium. J Bacteriol. 1998;180:3393–9.
133. Blocker AJ, Deane JE, Veenendaal AKJ, Roversi P, Hodgkinson JL, Johnson S, *et al.* What's the point of the type III secretion system needle? Proc Natl Acad Sci USA. 2008;105:6507–13.
134. Schlumberger MC, Müller AJ, Ehrbar K, Winnen B, Duss I, Stecher B, *et al.* Real-time imaging of type III secretion: *Salmonella* SipA injection into host cells. PNAS. 2005;102:12548–53.
135. Lee P-C, Rietsch A. Fueling type III secretion. Trends in Microbiology. 2015;23:296–300.
136. Galán JE, Lara-Tejero M, Marlovits TC, Wagner S. Bacterial Type III Secretion Systems: Specialized Nanomachines for Protein Delivery into Target Cells. Annu Rev Microbiol. 2014;68:415–38.
137. Akeda Y, Galán JE. Chaperone release and unfolding of substrates in type III secretion. Nature. 2005;437:911–5.
138. Erhardt M, Mertens ME, Fabiani FD, Hughes KT. ATPase-Independent Type-III Protein Secretion in *Salmonella enterica*. PLoS Genet. 2014;10:e1004800.
139. dos Santos AMP, Ferrari RG, Conte-Junior CA. Virulence Factors in *Salmonella* Typhimurium: The Sagacity of a Bacterium. Curr Microbiol. 2019;76:762–73.
140. Marcus SL, Brumell JH, Pfeifer CG, Finlay BB. *Salmonella* pathogenicity islands: big virulence in small packages. Microbes and Infection. 2000;2:145–56.
141. Hensel M. *Salmonella* pathogenicity island 2. Mol Microbiol. 2000;36:1015–23.
142. Blanc-Potard A-B, Solomon F, Kayser J, Groisman EA. The SPI-3 Pathogenicity Island of *Salmonella enterica*. Journal of Bacteriology. 1999;181:998–1004.
143. Wong K-K, McClelland M, Stillwell LC, Sisk EC, Thurston SJ, Saffer JD. Identification and Sequence Analysis of a 27-Kilobase Chromosomal Fragment Containing a *Salmonella* Pathogenicity Island Located at 92 Minutes on the Chromosome Map of *Salmonella enterica* Serovar Typhimurium LT2. Infection and Immunity. 1998;66:3365–71.
144. Wood MW, Jones MA, Watson PR, Hedges S, Wallis TS, Galyov EE. Identification of a pathogenicity island required for *Salmonella* enteropathogenicity. Mol Microbiol. 1998;29:883–91.
145. Chilcott GS, Hughes KT. Coupling of Flagellar Gene Expression to Flagellar Assembly in *Salmonella enterica* Serovar Typhimurium and Escherichia coli. Microbiol Mol Biol Rev. 2000;64:694–708.

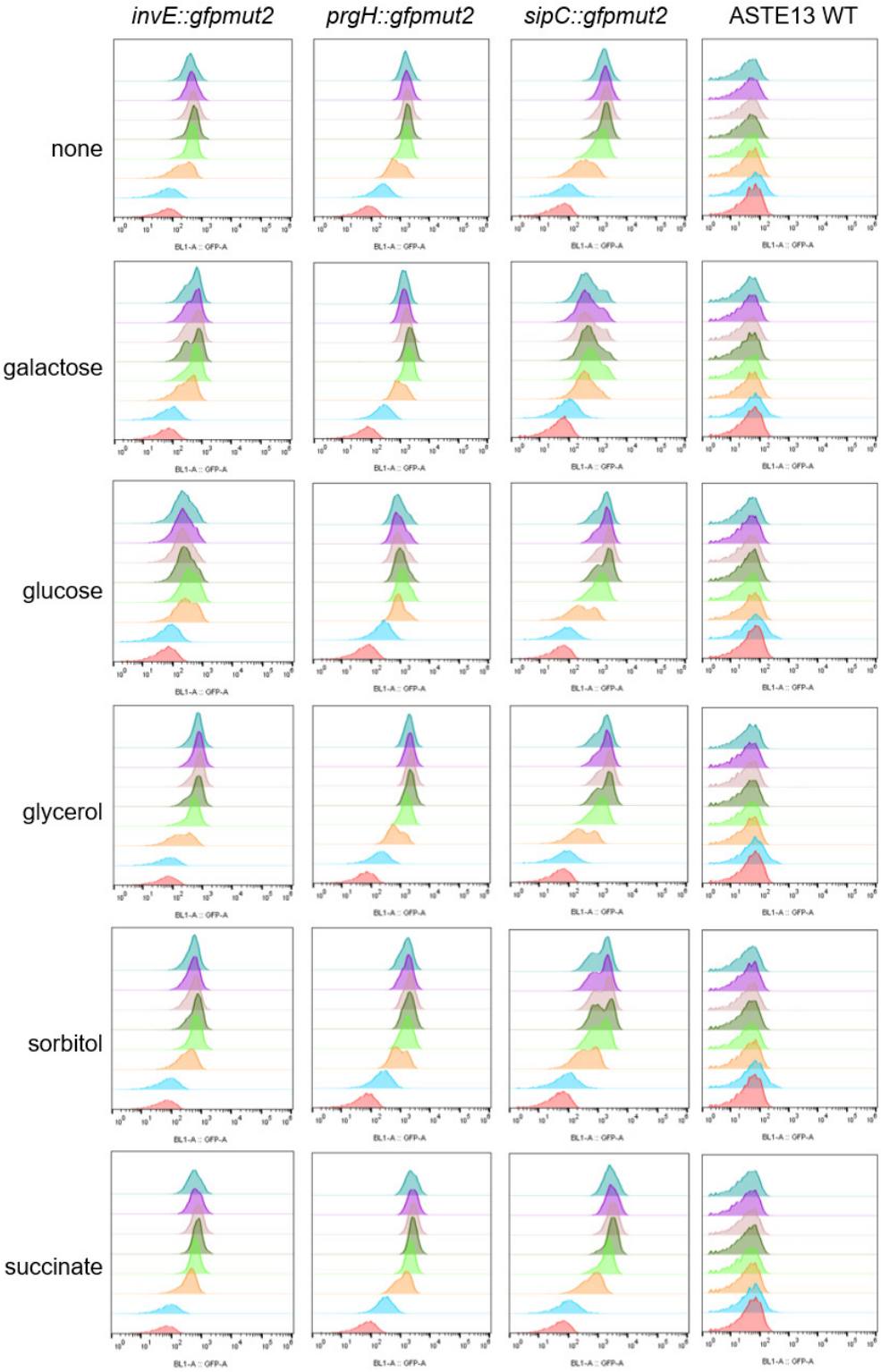
146. van der Velden AWM, Bäumler AJ, Tsolis RM, Heffron F. Multiple Fimbrial Adhesins Are Required for Full Virulence of *Salmonella* typhimurium in Mice. Infect Immun. 1998;66:2803–8.
147. Chu C, Hong S-F, Tsai C, Lin W-S, Liu T-P, Ou JT. Comparative Physical and Genetic Maps of the Virulence Plasmids of *Salmonella enterica* Serovars Typhimurium, Enteritidis, Choleraesuis, and Dublin. Infection and Immunity. 1999;67:2611–4.
148. Gulig PA, Doyle TJ, Hughes JA, Matsui H. Analysis of Host Cells Associated with the Spv-Mediated Increased Intracellular Growth Rate of *Salmonella* typhimurium in Mice. Infection and Immunity. 1998;66:2471–85.
149. Cameron ADS, Dorman CJ. A Fundamental Regulatory Mechanism Operating through OmpR and DNA Topology Controls Expression of *Salmonella* Pathogenicity Islands SPI-1 and SPI-2. PLoS Genet. 2012;8:e1002615.
150. Laemmli UK. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature. 1970;227:680–5.
151. Engler C, Kandzia R, Marillonnet S. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. PLoS ONE. 2008;3:e3647.
152. Thomason LC, Sawitzke JA, Li X, Costantino N, Court DL. Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination. Current Protocols in Molecular Biology. 2014;106:1.16.1-1.16.39.
153. Studier FW. Protein production by auto-induction in high density shaking cultures. Protein Expr Purif. 2005;41:207–34.
154. Tartera C, Metcalf ES. Osmolarity and growth phase overlap in regulation of *Salmonella typhi* adherence to and invasion of human intestinal cells. Infect Immun. 1993;61:3084–9.
155. Ahmad KF, Lim WA. The Minimal Autoinhibited Unit of the Guanine Nucleotide Exchange Factor Intersectin. PLoS ONE. 2010;5:e11291.
156. Kim E, Jakobson CM, Tullman-Ercek D. Engineering transcriptional regulation to control Pdu bacterial microcompartment formation. PLoS ONE. 2014;
157. Neidhardt FC, Bloch PL, Smith DF. Culture Medium for Enterobacteria Culture Medium for Enterobacteria. 1974;119.
158. Good NE, Winget GD, Winter W, Connolly TN, Izawa S, Singh RM. Hydrogen ion buffers for biological research. Biochemistry. 1966;5:467–77.
159. Leysath CE, Monzingo AF, Maynard JA, Barnett J, Georgiou G, Iverson BL, *et al.* Crystal Structure of the Engineered Neutralizing Antibody M18 Complexed to Domain 4 of the Anthrax Protective Antigen. Journal of Molecular Biology. 2009;387:680–93.
160. Maynard JA, Maassen CBM, Leppla SH, Brasky K, Patterson JL, Iverson BL, *et al.* Protection against anthrax toxin by recombinant antibody fragments correlates with antigen affinity. Nature Biotechnology. 2002;20:597–601.
161. Kaleta C, Schäuble S, Rinas U, Schuster S. Metabolic costs of amino acid and protein production in *Escherichia coli*. Biotechnology Journal. 2013;8:1105–14.
162. Carneiro S, Ferreira EC, Rocha I. Metabolic responses to recombinant bioprocesses in *Escherichia coli*. Journal of Biotechnology. 2013;164:396–408.
163. Widmaier DM, Voigt CA. Quantification of the physiochemical constraints on the export of spider silk proteins by *Salmonella* type III secretion. Microbial Cell Factories. 2010;9:78.

164. Berkowitz D, Hushon JM, Whitfield HJ, Roth J, Ames BN. Procedure for identifying nonsense mutations. J Bacteriol. 1968;96:215–20.
165. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003;31:3784–8.
166. Shiloach J, Fass R. Growing E. coli to high cell density—A historical perspective on method development. Biotechnology Advances. 2005;23:345–57.
167. Cabantous S, Waldo GS. In vivo and in vitro protein solubility assays using split GFP. Nat Meth. 2006;3:845–54.
168. Steinberg TH. Protein gel staining methods: an introduction and overview. Meth Enzymol. 2009;463:541–63.
169. Alegria-Schaffer A, Lodge A, Vattem K. Performing and optimizing Western blots with an emphasis on chemiluminescent detection. Meth Enzymol. 2009;463:573–99.
170. McDonough AA, Veiras LC, Minas JN, Ralph DL. Considerations when quantitating protein abundance by immunoblot. Am J Physiol, Cell Physiol. 2015;308:C426-433.
171. Ghosh M, Talukdar D, Ghosh S, Bhattacharyya N, Ray M, Ray S. In vivo assessment of toxicity and pharmacokinetics of methylglyoxal. Augmentation of the curative effect of methylglyoxal on cancer-bearing mice by ascorbic acid and creatine. Toxicology and applied pharmacology. 2006;212:45–58.
172. Adams SR, Campbell RE, Gross LA, Martin BR, Walkup GK, Yao Y, *et al.* New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. J Am Chem Soc. 2002;124:6063–76.
173. Van Engelenburg SB, Nahreini T, Palmer AE. FACS-based selection of tandem tetracysteine peptides with improved ReAsH brightness in live cells. Chembiochem. 2010;11:489–93.
174. Green CA, Kamble NS, Court EK, Bryant OJ, Hicks MG, Lennon C, *et al.* Engineering the flagellar type III secretion system: improving capacity for secretion of recombinant protein. Microb Cell Fact. 2019;18:10.
175. Hu J, Worrall LJ, Hong C, Vuckovic M, Atkinson CE, Caveney N, *et al.* Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin. Nature Communications. 2018;9:1–11.
176. Radics J, Königsmaier L, Marlovits TC. Structure of a pathogenic type 3 secretion system in action. Nat Struct Mol Biol. 2013;21:82–7.
177. Marlovits TC, Kubori T, Sukhan A, Thomas DR, Galán JE, Unger VM. Structural Insights into the Assembly of the Type III Secretion Needle Complex. Science. 2004;306:1040–2.
178. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. J Comput Chem. 2004;25:1605–12.
179. Enninga J, Mounier J, Sansonetti P, Nhieu GTV. Secretion of type III effectors into host cells in real time. Nature Methods. 2005;2:959–65.
180. Widmaier DM, Tullman-Ercek D, Mirsky EA, Hill R, Govindarajan S, Minshull J, et al. Engineering the *Salmonella* type III secretion system to export spider silk monomers. Mol Syst Biol. 2009;5:309.
181. Metcalf KJ. Engineering heterologous protein secretion for improved production. University of California, Berkeley; 2016.

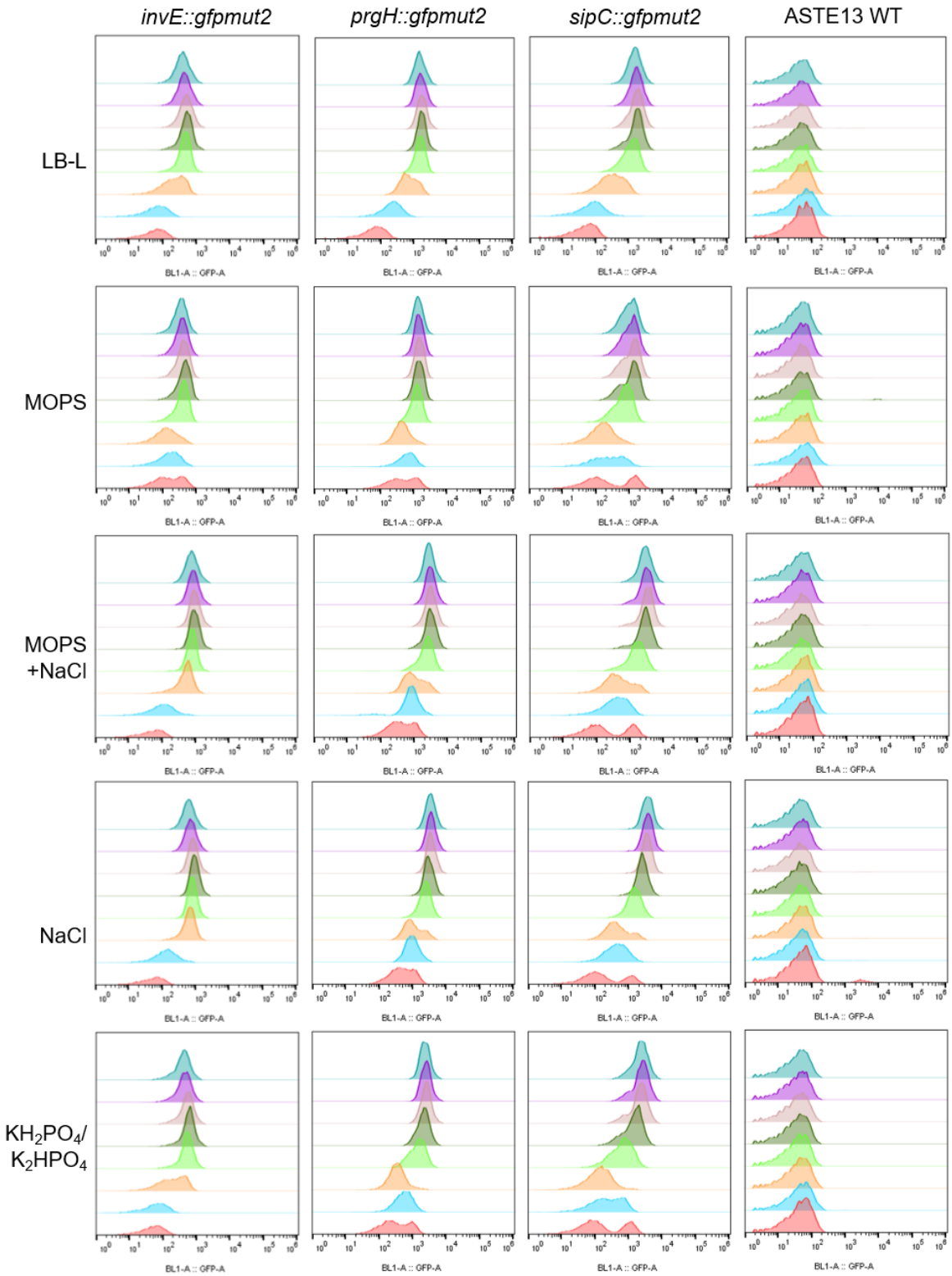
182. Barison N, Gupta R, Kolbe M. A sophisticated multi-step secretion mechanism: how the type 3 secretion system is regulated. Cellular Microbiology. 2013;n/a–n/a.
183. Davis AJ, Mecsas J. Mutations in the *Yersinia pseudotuberculosis* Type III Secretion System Needle Protein, YscF, That Specifically Abrogate Effector Translocation into Host Cells. J Bacteriol. 2007;189:83–97.
184. Torruellas J, Jackson MW, Pennock JW, Plano GV. The *Yersinia pestis* type III secretion needle plays a role in the regulation of Yop secretion. Mol Microbiol. 2005;57:1719–33.
185. Kenjale R, Wilson J, Zenk SF, Saurya S, Picking WL, Picking WD, *et al.* The needle component of the type III secreton of *Shigella* regulates the activity of the secretion apparatus. J Biol Chem. 2005;280:42929–37.
186. Guo EZ, Desrosiers DC, Zalesak J, Tolchard J, Berbon M, Habenstein B, *et al.* A polymorphic helix of a *Salmonella* needle protein relays signals defining distinct steps in type III secretion. PLOS Biology. 2019;17:e3000351.
187. Hartman EC, Jakobson CM, Favor AH, Lobba MJ, Álvarez-Benedicto E, Francis MB, *et al.* Quantitative characterization of all single amino acid variants of a viral capsid-based drug delivery vehicle. Nature Communications. 2018;9:1385.
188. Patrick WM, Firth AE, Blackburn JM. User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries. Protein Eng. 2003;16:451–7.
189. Higgins SA, Ouonkap SVY, Savage DF. Rapid and Programmable Protein Mutagenesis Using Plasmid Recombineering. ACS Synth Biol. 2017;6:1825–33.
190. Darboe N, Kenjale R, Picking WL, Picking WD, Middaugh CR. Physical characterization of MxiH and PrgI, the needle component of the type III secretion apparatus from *Shigella* and *Salmonella*. Protein Science. 2006;15:543–52.
191. Loquet A, Sgourakis NG, Gupta R, Giller K, Riedel D, Goosmann C, *et al.* Atomic model of the type III secretion system needle. Nature. 2012;486:276–9.
192. Buchholz K, Collins J. The roots—a short history of industrial microbiology and biotechnology. Appl Microbiol Biotechnol. 2013;97:3747–62.
193. Ferrer-Miralles N, Domingo-Espín J, Corchero JL, Vázquez E, Villaverde A. Microbial factories for recombinant pharmaceuticals. Microb Cell Fact. 2009;8:17.
194. Nielsen J, Keasling JD. Engineering Cellular Metabolism. Cell. 2016;164:1185–97.
195. Gibbons JG, Rinker DC. The Genomics of Microbial Domestication in the Fermented Food Environment. Curr Opin Genet Dev. 2015;35:1–8.
196. Douglas GL, Klaenhammer TR. Genomic evolution of domesticated microorganisms. Annu Rev Food Sci Technol. 2010;1:397–414.
197. Petrone BL, Stringer AM, Wade JT. Identification of HilD-Regulated Genes in *Salmonella enterica* Serovar Typhimurium. J Bacteriol. 2014;196:1094–101.
198. Sturm A, Heinemann M, Arnoldini M, Benecke A, Ackermann M, Benz M, et al. The Cost of Virulence: Retarded Growth of *Salmonella* Typhimurium Cells Expressing Type III Secretion System 1. PLoS Pathog. 2011;7:e1002143.
199. Baxter MA, Fahlen TF, Wilson RL, Jones BD. HilE Interacts with HilD and Negatively Regulates *hilA* Transcription and Expression of the *Salmonella enterica* Serovar Typhimurium Invasive Phenotype. Infect Immun. 2003;71:1295–305.

200. López-Garrido J, Puerta-Fernández E, Casadesús J. A eukaryotic-like 3′ untranslated region in *Salmonella enterica hilD* mRNA. Nucleic Acids Res. 2014;42:5894–906.
201. Park D, Lara-Tejero M, Waxham MN, Li W, Hu B, Galán JE, et al. Visualization of the type III secretion mediated *Salmonella*–host cell interface using cryo-electron tomography. Musacchio A, Orth K, editors. eLife. 2018;7:e39514.

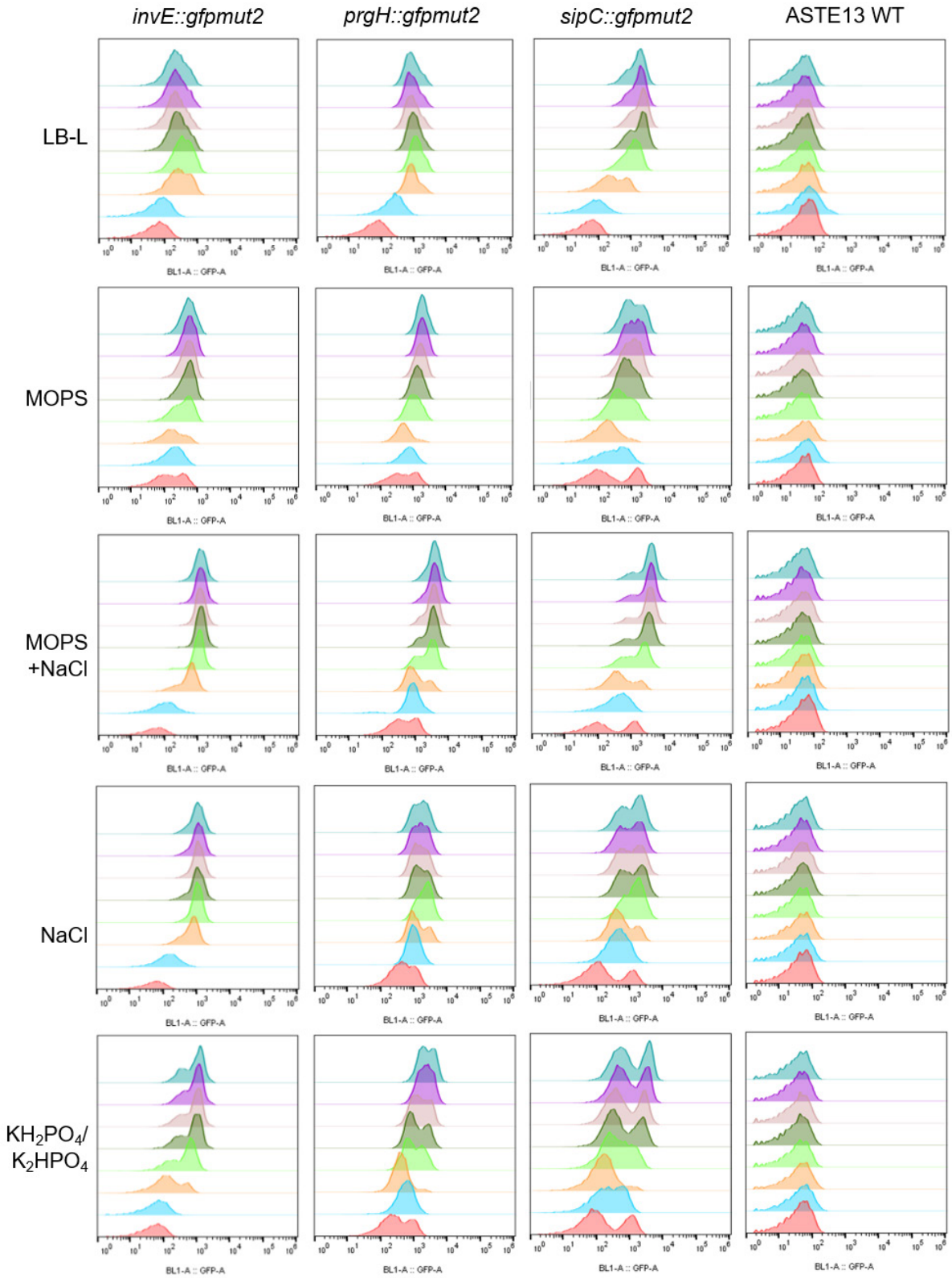
Appendix A: Chapter 2 Supplementary Figures



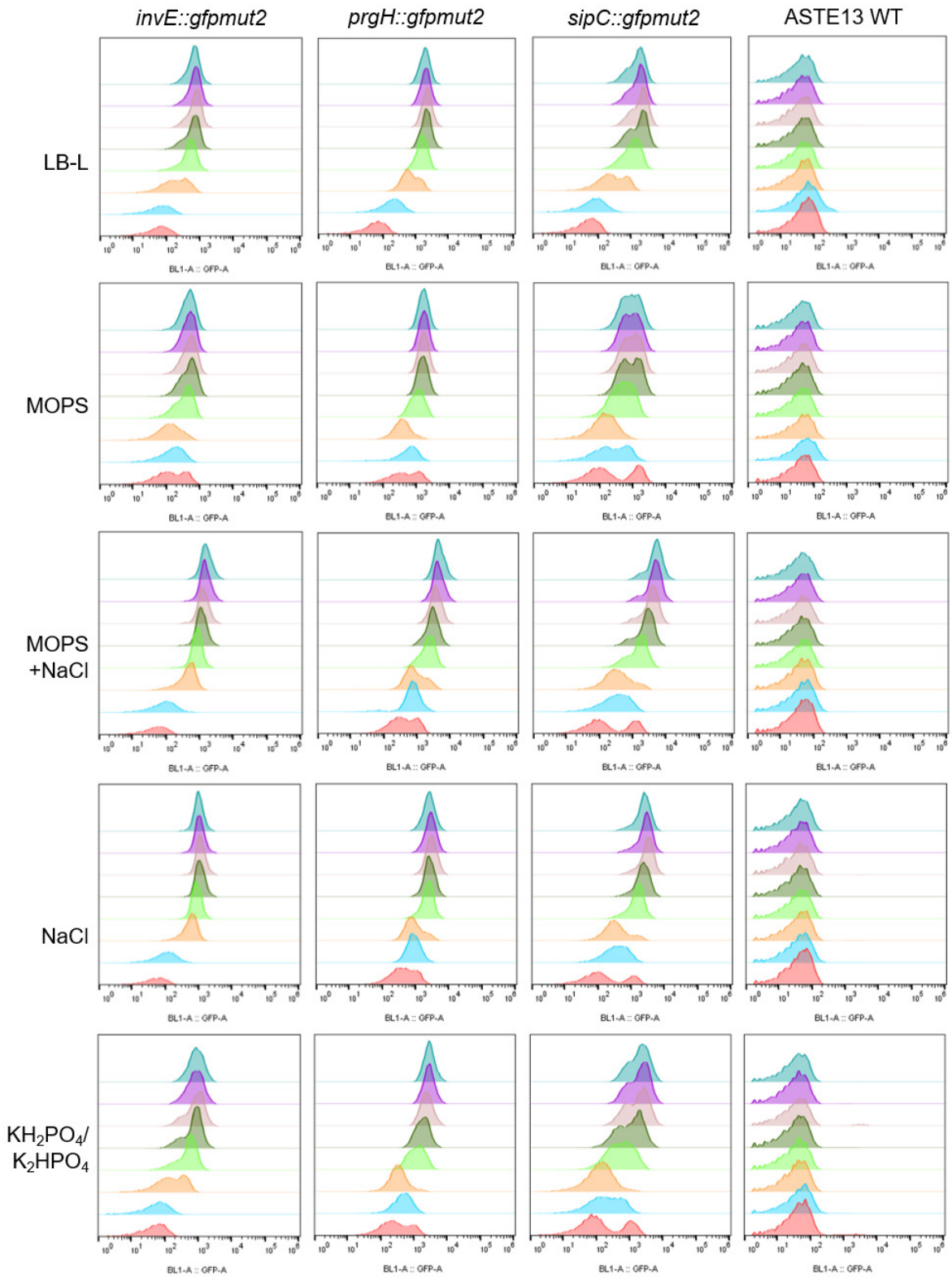
Supplementary Figure 2.1 Histograms of flow cytometry GFP signal for SPI-1 transcriptional activity in LB-L with 0.4% w/v carbon sources (Figure 2.2B). Overlaid populations represent hourly time points from 1 hour (red) to 8 hours (teal). “ASTE13 WT” is a negative control.



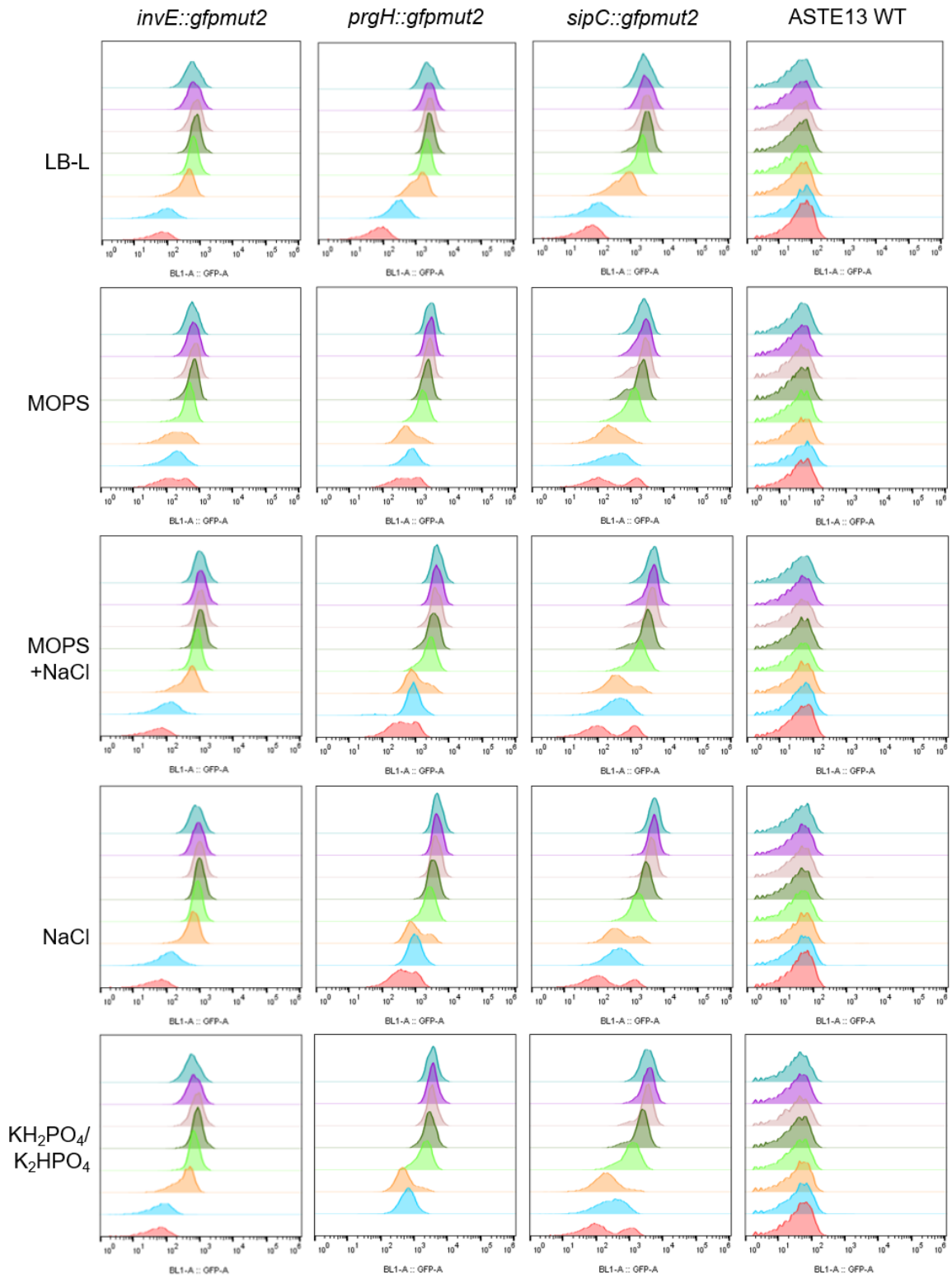
Supplementary Figure 2.2 Histograms of GFP signal for flow cytometry to measure SPI-1 transcriptional activity in LB with various buffers and salts with no added carbon source ("None"; Figures 2.6 and 2.7). Overlaid populations represent hourly time points from 1 hour (red) to 8 hours (teal). "ASTE13 WT" is a negative control.



Supplementary Figure 2.3 Histograms of GFP signal for flow cytometry to measure SPI-1 transcriptional activity in LB with various buffers, salts, and 0.4% w/v glucose (Figures 2.6 and 2.7). Overlaid populations represent hourly time points from 1 hour (red) to 8 hours (teal). "ASTE13 WT" is a negative control.



Supplementary Figure 2.4 Histograms of GFP signal for flow cytometry to measure SPI-1 transcriptional activity in LB with various buffers, salts, and 0.4% w/v glycerol (Figures 2.6 and 2.7). Overlaid populations represent hourly time points from 1 hour (red) to 8 hours (teal). “ASTE13 WT” is a negative control.



Supplementary Figure 2.5 Histograms of GFP signal for flow cytometry to measure SPI-1 transcriptional activity in LB with various buffers, salts, and 0.4% w/v succinate (Figures 2.6 and 2.7). Overlaid populations represent hourly time points from 1 hour (red) to 8 hours (teal). “ASTE13 WT” is a negative control.

Appendix B: Defined Medium Recipes

Supplementary Table 3.1 Defined media from Chapter 3

	Amount (mM)				
	NCE []	**M9**	**M9 +ferric citrate**	**PCN []**	**PCN+P**
NaCl	0	8.6	8.6	50	50
$MgCl_2$	0	0	0	0.52	0.52
$MgSO_4$	1.0	1.0	1.0	0	0
Na_2HPO_4	0	48	48	0	0
NaH_2PO_4	0	0	0	0	0
NH_4Cl	0	19	19	9.5	9.5
K_2HPO_4	34	0	0	1.7	34
KH_2PO_4	29	22	22	0	29
K_2SO_4	0	0	0	0.28	0.28
$FeSO_4$	0	0	0	0.010	0.010
$CaCl_2$	0	0.1	0.1	0.00050	0.00050
$Na(NH_4)HPO_4$	17	0	0	0	0
Ferric citrate	0.050	0.050	0.050	0	0
MOPS	0	0	0	40	40
Tricine	0	0	0	4.0	4.0
Micronutrients	0	0	0	See below	See below

Supplementary Table 3.2 Micronutrients in PCN media

PCN Micronutrients []	**Amount (mM)**
$(NH_4)_4Mo_7O_{24}$	2.9×10^{-7}
H_3BO_3	4.0×10^{-5}
$CoCl_2$	3.0×10^{-6}
$CuSO_4$	9.6×10^{-7}
$MnCl_2$	8.1×10^{-6}
$ZnSO_4$	9.7×10^{-7}

Supplementary Table 3.3 Components of 5X Supplement EZ (Teknova #M2104)

	Amount (mM)
L-Alanine	0.8
L-Arginine HCl	5.2
L-Asparagine	0.4
L-Aspartic Acid, Potassium Salt	0.4
L-Glutamic Acid, Potassium Salt	0.6
L-Glutamine	0.6
L-Glycine	0.8
L-Histidine HCl H_2O	0.2
L-Isoleucine	0.4
L-Proline	0.4
L-Serine	10
L-Threonine	0.4
L-Tryptophan	0.1
L-Valine	0.6
L-Leucine	0.8
L-Lysine HCl	0.4
L-Methionine	0.2
L-Phenylalanine	0.4
L-Cysteine HCl	0.1
L-Tyrosine	0.2
Thiamine HCl	0.01
Calcium Pantothenate	0.01
para-Amino Benzoic Acid	0.01
para-Hydroxy Benzoic Acid	0.01
2,3-diHydroxy Benzoic Acid	0.01

Appendix C: Weighted-Average Secretion Score Calculations

Secretion Titer Score Definitions

m: one of the 20 canonical amino acids.

P_i: pool of screened clones sorted by secretion titer relative to PrgI^{WT}. Nonfunctional clones are in pool A, and pools $B - J$ contain functional clones with relative secretion titer increasing from B to J.

$$P = [A, B, C, D, E, F, G, H, I, J]$$

s_{P_i} : a secretion titer score assigned to each pool P_i.

$$s_P = [-1, 1, 2, 3, 4, 5, 6, 7, 8, 9]$$

$A_{P_i,p,m}$: an abundance score in pool P_i, indexed by position p and mutation m.

$B_{P_i,p,m}$: a binary array recording presence (“1”) or absence (“0”) of a mutation m at position p in pool P_i.

$PA_{P_i,p,m}$: a percent abundance score for each mutation m at position p in pool P_i.

$S_{P_i,p,m}$: a secretion titer score for each mutation m at position p in pool P_i.

$SS_{p,m}$: an array of secretion titer scores for mutation m at position p.

$PAW_{p,m}$: an array of percent abundances for mutation m at position p.

$w_{p,m}$: an array of weighted percent abundances for mutation m at position p.

$WS_{p,m}$: a weighted average secretion titer score for each mutation m at position p in pool P_i.

$\overline{WS_p}$: average secretion titer score per residue

Secretion Titer Score Calculations

Sequences were trimmed and aligned using code written in-house. Following the data processing described in Chapter 5 Methods, a text file was produced containing one sequencing read per line. The text file was read into Python, and only lines starting with “ATG” and ending with “TAA” were kept.

A second quality control step was implemented in Python: sequences with read counts < 20, sequences that matched the wild-type PrgI sequence, and sequences that contained more than one mutation were discarded.

Sequences that survived the more stringent quality control step were translated, and two matrices were populated: an abundance array A with read counts for each mutation at each position, and a binary array B containing a "1" for a mutation present at a position and a "0" for a mutation not present at a position.

Percent abundance for each mutation at each position was generated by dividing by the total number of read counts:

$$PA_{P_i,p,m} = \frac{A_{P_i,p,m}}{\sum A_{P_i,p,m}}$$

A secretion titer score was assigned for $B_{P_i,p,m} = 1$ to generate a new matrix for each pool with the appropriate secretion titer score at mutations that were present in that pool:

$$S_{P_i,p,m} = B_{P_i,p,m} * s_{P_i}$$

A weighted average secretion score was calculated using $PA_{P_i,p,m}$ to compensate for the appearance of mutations in multiple pools. Secretion titer scores and percent abundances from all pools were collected into arrays for each mutation m at position p:

$$SS_{p,m_i} = S_{P_i,p,m}$$
$$PAW_{p,m_i} = PA_{P_i,p,m}$$

An array of weights was generated from the percent abundances and multiplied by the list of secretion scores to calculate a weighted average secretion titer score:

$$w_{p,m_i} = \frac{PAW_{p,m_i}}{\sum PAW_{p,m_i}}$$
$$WS_{p,m} = SS_{p,m} \cdot {w_{p,m}}^T$$

$WS_{p,m}$ is plotted in Figure 5.15. The average secretion score per residue was calculated by averaging $WS_{p,m}$ for each p, ignoring missing values:

$$\overline{WS_p} = \frac{\sum WS_{p,m}}{19 - \sum(\#\ NaN)_p}$$