The Pennsylvania State University

The Graduate School

Department of Biochemistry and Molecular Biology

THE ROLE AND REGULATION OF THE EXTRACYTOPLASMIC FUNCTION SIGMA FACTOR SIGE IN *BORDETELLA BRONCHISEPTICA*

A Dissertation in

Biochemistry, Microbiology and Molecular Biology

by

Sarah E. Barchinger

©2012 Sarah E. Barchinger

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2012

The dissertation of Sarah E. Barchinger was reviewed and approved* by the following:

Sarah E. Ades
Associate Professor of Biochemistry and Molecular Biology
Dissertation Advisor
Chair of Committee

Eric T. Harvill
Associate Professor of Microbiology and Infectious Disease

Paul Babitzke
Professor of Biochemistry and Molecular Biology

B. Tracy Nixon
Professor of Biochemistry and Molecular Biology

Edward G. Dudley
Casida Development Professor of Food Science
Assistant Professor of Food Science

Craig E. Cameron
Paul Berg Professor of Biochemistry and Molecular Biology
Associate Department Head for Research and Graduate Education

*Signatures are on file in the Graduate School

Abstract

During infection, pathogenic bacteria have to contend not only with the external environment during transmission, but also with the harsh conditions of the host environment, where nutrient availability can be limited and the immune system is poised to attack them. The bacterial cell envelope lies at the interface between a bacterial cell and its environment. To relay information about changes sensed at this interface to the transcriptional machinery and elicit a response, bacteria have developed systems such as extracytoplasmic function sigma factors that allow the cell to rapidly redirect transcription to a particular subset of genes in order to appropriately respond to the stress condition.

I have identified an important extracytoplasmic function sigma factor, SigE, in *Bordetella bronchiseptica,* a respiratory pathogen that causes disease ranging from asymptomatic infection to fatal pneumonia in most mammals. SigE is a member of the RpoE-like group of sigma factors, which includes σ^E from *Escherichia coli* and *Salmonella enterica*, as well AlgU from *Pseudomonas aeruginosa* and *Burkholderia pseudomallei*. In this dissertation, I demonstrate that *B. bronchiseptica sigE* encodes an RpoE-like sigma factor, and that the downstream genes *rseA* and *rseB* encode negative regulators of its activity. SigE is required for response to heat shock, ethanol stress, and specific cell envelope stress, and both no *sigE* and high SigE activity are detrimental to virulence. Cells with high SigE activity, in particular, are defective in early colonization of the lower respiratory tract of mice, and cause less disease in immunocompromised mice. By using multiple methods, I have identified members of the SigE regulon, including some candidate SigE-regulated small regulatory RNAs, and show that the SigE system plays a

role in regulation of cell envelope components, consistent with its role in responding to cell envelope stress. Through this work, I have demonstrated a role for SigE in regulation of both cell envelope stress response and virulence, independently of the well-characterized BvgAS two component system in the bordetellae. *B. bronchiseptica* is closely related to both *B. pertussis* and *B. parapertussis*, the etiological agents of whooping cough in humans, and the *sigE* locus is nearly 100% conserved in all three species, suggesting an important and conserved role for this system in *Bordetella* species.

TABLE OF CONTENTS

LIST OF FIGURES ..... **x**

LIST OF TABLES ..... **xii**

ACKNOWLEDGEMENTS ..... **xiii**

Chapter 1: Introduction **1**
- **Sensing stress at the Gram-negative cell envelope** **1**
- **Extracytoplasmic function** **4**
- **σ^E** **6**
  - Evolutionary conservation of the σ^E system 6
  - Basic components of the σ^E system 8
  - Conserved and varied roles for σ^E systems 9
- **σ^E systems in diverse bacterial species** **10**
  - Escherichia coli 10
  - *Salmonella enterica* serovar Typhimurium 14
  - *Vibrio* spp. 15
  - Pseudomonas spp. 16
    - Pseudomonas aeruginosa 16
    - Pseudomonas syringae and Pseudomonas fluorescens 19
  - Burkholderia spp 20
  - Other bacterial species 21
- **Bordetella** **23**

Chapter 2: *sigE* facilitates the adaptation of *Bordetella bronchiseptica* to stress conditions and lethal systemic infection in mice **26**
- **Abstract** **26**
- **Introduction** **28**
- **Results** **31**
  - *sigE* encodes an active sigma factor 31
  - SigE contributes to the *B. bronchiseptica* stress response 33
  - Growth in the murine respiratory tract is not affected by the lack of *sigE* 38
  - SigE contributes to systemic spread and lethal *B. bronchiseptica* infection in mice lacking B cells and T cells, but not in mice lacking TLR4 or TNF-α. 39
  - SigE contributes to cytotoxicity to macrophages 44
  - RB50Δ*sigE* is more efficiently phagocytosed and killed by PMNs. 46
- **Discussion** **48**
- **Materials and Methods** **52**
  - Strains and Media 52
  - Plasmid construction. 52

Construction of RB50Δ*sigE*. 54
β-galactosidase assays 56
Complementation of E. coli ΔrpoE by B. bronchiseptica sigE 56
Protein purification 57
In vitro transcription 57
Disk diffusion assays 58
Temperature and ethanol stress 58
Complement killing assay. 58
Cytotoxicity assay 59
Phagocytosis and killing by polymorphonuclear leukocytes 59
Animal experiments 60
Statistical analysis 61

Chapter 3: RseA and RseB negatively regulate *Bordetella bronchiseptica* SigE and contribute to colonization and virulence in mice 62
Abstract 62
Introduction 64
Results 67
B. bronchiseptica RseA and RseB share similarity with other negative regulators of RpoE-like sigma factors 67
RseA and RseB negatively regulate SigE 70
RB50Δ*rseAB* is more resistant to some cell envelope stress 73
Constitutive SigE activity leads to decreased colonization in the lower respiratory tract 76
Early clearance from the lower respiratory tract is not due to altered interactions with adaptive immunity 78
RB50Δ*rseAB* does not cause systemic, lethal infection in TLR4^{def} mice and has altered interactions with components of innate immunity. 81
Discussion 85
Materials and Methods 89
Strains and Media 89
Plasmid constructions 90
Construction of RB50Δ*rseAB* 91
β-galactosidase assays 93
Protein purification 93
In vitro transcription 94
Temperature stress 95
Disk diffusion assays 95
Cytotoxicity assay 96
Complement killing assay 96
Animal experiments 96
Lung pathology 97
Statistical analysis 98
Acknowledegements 98

Chapter 4: SigE transcribes genes important at the cell envelope 99
Abstract 99
Introduction 100
Results and Discussion 102
Identifying candidate SigE regulon members 102
Gene expression analysis 102
Run-off transcription coupled to microarray analysis 108
Prediction of putative SigE-regulated promoters in *B. bronchiseptica* 111
Combining methods to predict the SigE regulon 123
Confirming SigE-dependent transcription of candidate regulon members 127
Role of the SigE regulon at the cell envelope and in pathogenesis 130
Cells with constitutively active SigE have an altered cell envelope protein composition 130
Different antibodies are produced during RB50Δ*rseAB* infection compared to RB50 infection 132
Cells with high SigE activity have minor differences in lipopolysaccharides 134
Possible roles of SigE regulon members in pathogenesis 136
Cytotoxicity 136
Interactions with BvgAS 138
Conservation of SigE-regulated promoters in the classical bordetellae 140
Materials and Methods 143
Bacterial strains and growth 143
RNA isolation, preparation of labeled cDNA, and microarray analysis 144
qRT-PCR 144
Promoter prediction 145
Protein purification 146
In vitro transcription and ROMA 146
Separation of cell envelope proteins and western blotting 150
Preparation of lipopolysaccharides 150
Cytotoxicity assays 151
Acknowledegements 151

Chapter 5: Conclusions and Future Directions 152
Components and regulation of the SigE system 153
SigE, RseA, and RseB 153
Regulation of the proteolysis pathway 155
ppGpp 157
Transcriptional regulation of the *sigE* locus 158
The role of the SigE system in cell envelope stress response 160
The role of the SigE system in virulence 162
Colonization and persistence in the lower respiratory tract 162
Interactions with the host immune system 164
The SigE regulon 167
SigE-regulated sRNAs 169
Other ECF sigma factors in *B. bronchiseptica* 170
Final Comments 174

Appendix A: Small RNAs in *Bordetella bronchiseptica* 176
Abstract 176
Introduction 177
Results and Discussion 181
Identifying sRNAs in *Bordetella bronchiseptica* 181
Prediction of sRNAs bioinformatically 181
Sequencing sRNAs from the bordetellae 203
Comparing sequenced and predicted sRNAs 212
Analysis of identified sRNAs 213
Distance to nearest ORF 213
Confirmation of sRNAs by northern blot 214
Predicting targets for the candidate sRNAs 215
Putative SigE-regulated sRNAs 217
Comparing identified sRNAs in the classical bordetellae 222
Materials and Methods 226
Predicting candidate sRNAs with SIPHT 226
Bacterial strains 226
RNA isolation and library preparation 226
SOLiD sequencing and analysis 227
Genome context and sRNA conservation analysis 227
Northern blotting 228
TargetRNA analysis 228

Appendix B: ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor σ^E in *Escherichia coli* by both direct and indirect mechanisms 229
Abstract 229
Introduction 230
Results 235
ppGpp regulates σ^E activity 235
ppGpp affects the activity, not production, of σ^E 241
Regulation of σ^E-dependent transcription by ppGpp 242
Discussion 250
ppGpp and the regulation of σ^E expression 250
Direct regulation of σ^E-dependent transcription by ppGpp 251
Indirect regulation of alternative sigma factor transcription by ppGpp 252
Materials and Methods 255
Strains and growth conditions 255
Plasmid constructions 255
Western blotting 256
β-galactosidase assays 257
Protein purification 258
RNAP-promoter complex decay assays 259

Multi-round transcription assays 259
Acknowledegements 262

Appendix C: Regulated Proteolysis: Control of the *Escherichia coli* σ^E-dependent Cell Envelope Stress Response 263
Abstract 263
Introduction 266
Cell Envelope Stress and σ^E 270
Regulation of σ^E activity 272
The proteolytic cascade 274
Activating σ^E via regulated proteolysis 276
Step 1: Signal recognition 276
Step 2: Activate DegS 281
Role of the PDZ domains 282
Mechanism of activation 285
Positive cooperativity 287
Allosteric regulation- The MWC model 287
Step 3: Activate RseP 291
Substrate selectivity 293
Regulation of RseP 293
Role of the PDZ domains 294
Step 4: Releasing σ^E from RseA-cyto 296
Step 5: Activation of the σ^E regulon and return to a resting state 297
A common system for regulation of membrane-localized sigma/anti-sigma modules 297
Regulation of *Pseudomonas aeruginosa* AlgU 298
Regulation of *Bacillus subtilis* σ^W 300
Regulation of virulence in *Mycobacterium tuberculosis* 301
Perspectives 301

References 304

LIST OF FIGURES

Figure 1-1: The Gram-negative bacterial cell envelope....3
Figure 1-2: Gene context of RpoE-like sigma factors in multiple species....8
Figure 1-3: The basic components of most known σ^E systems....9
Figure 1-4: σ^E regulation in *E. coli*....12

Figure 2-1: *B. bronchiseptica* SigE is a functional sigma factor....32
Figure 2-2: Role of SigE in response to environmental stresses....37
Figure 2-3: Colonization of the respiratory tract of C57BL/6 mice by RB50 and RB50Δ*sigE*....39
Figure 2-4: SigE is not required for survival or systemic colonization in TNFα$^{-/-}$ or TLR4^{def} mice....41
Figure 2-5: Survival and systemic colonization of Rag1-/- mice following infection with RB50 and RB50Δ*sigE*....43
Figure 2-6: RB50Δ*sigE* is not more susceptible to complement-mediated killing....44
Figure 2-7: RB50Δ*sigE* is less cytotoxic to macrophages....45
Figure 2-8: RB50Δ*sigE* is more efficiently phagocytosed and killed by PMNs....46

Figure 3-1: Sequence and structural similarity in the negative regulators of the SigE system....69
Figure 3-2: *B. bronchiseptica* RseA and RseB are negative regulators of SigE activity....71
Figure 3-3: *B. bronchiseptica* RseA inhibits $SigE_{Bb}$, but not σ^E_{Ec}; *E. coli* RseA inhibits both $SigE_{Bb}$ and σ^E_{Ec}....73
Figure 3-4: RB50Δ*rseAB* exhibits reduced thermotolerance, but increased resistance to some cell envelope stress....76
Figure 3-5: RB50Δ*rseAB* exhibits decreased colonization in the lower respiratory tract....78
Figure 3-6: Defects in colonization and early clearance cannot be explained by differences in interactions with antibodies....80
Figure 3-7: RB50Δ*rseAB* does not cause lethal, systemic infection in TLR4^{def} mice....82
Figure 3-8: Cells lacking *rseA* and *rseB* are less cytotoxic to macrophages....84

Figure 4-1: Preliminary qRT-PCR results for *fam* and T3SS locus genes....108
Figure 4-1: Alignment of σ^E-dependent promoters in *E. coli*....112
Figure 4-2: Combining all three methods to determine the SigE regulon....124
Figure 4-3: SigE-dependent promoters in *B. bronchiseptica*....128
Figure 4-4: RB50Δ*rseAB* has an altered cell envelope composition....134
Figure 4-5: RB50Δ*rseAB* does not increase cytotoxicity in response to serum to the same extent as RB50....136
Figure 4-6: Comparison of candidate SigE-regulated promoters in the bordetellae....141

Figure 5-1: The *E. coli* σE system and *B. bronchiseptica* SigE system 157
Figure 5-2: The *B. bronchiseptica sigE* operon has different architecture from the *E. coli rpoE* operon 159
Figure 5-3: Other predicted ECF sigma factors in *B. bronchiseptica* 171

Figure A-1: Distance of candidate sRNAs from the nearest ORF 214
Figure A-2: Confirming sRNA transcription with northern blotting 215
Figure A-3: SigE transcribes the promoter of either *bb3108* or *bbsRNA058*, as well as a promoter upstream of *hfq* 222
Figure A-4: Distance of candidate sRNAs from the nearest ORF 224

Figure B-1: σ^E activity increases under conditions in which ppGpp levels are known to increase 238
Figure B-2: Growth phase-dependent regulation of σ^E activity is shown for strains lacking proteins required for RNAP to respond to ppGpp 240
Figure B-3: ppGpp does not affect the steady state level of σ^E 242
Figure B-4: Suppression of the requirement for ppGpp for the increase in σ^E activity during entry into stationary phase 245
Figure B-5: Effects of ppGpp and DksA on Eσ^E-dependent transcription 249

Figure C-1: Overview of transmembrane signal transduction by regulated intramembrane proteolysis 266
Figure C-2: Regulation of the σ^E-dependent envelope response 269
Figure C-3: σ^E bound to the cytoplasmic domain of RseA 273
Figure C-4: RseB bound to the periplasmic domain of RseA 280
Figure C-5: DegS trimer 281
Figure C-6: Free and peptide-bound DegS monomer 284
Figure C-7: MWC model for allosteric regulation of DegS by activating peptides 288
Figure C-8: Sequential proteolysis of transmembrane anti-sigma factors as a regulatory paradigm 298

LIST OF TABLES

Table 2-1: Strains and plasmids 53
Table 2-2: Primer sequences 55

Table 3-1: Co-transduction frequency of *nadB*::Tn10 and *rpoE*::Kan^R in *E. coli* cells expressing components of the SigE system 72
Table 3-2: Growth rate of strains under non-stress conditions 74
Table 3-3: Strains 89
Table 3-4: Plasmids 91
Table 3-5: Primer sequences 92

Table 4-1: Genes with increased expression in RB50Δ*rseAB* compared to RB50 103
Table 4-2: Genes with decreased expression in RB50Δ*rseAB* compared to RB50.....105
Table 4-3: Genes identified as transcribed by SigE in vitro by ROMA.....109
Table 4-4: Predicted SigE-regulated promoters in *B. bronchiseptica* 113
Table 4-5: Putative SigE regulon 125
Table 4-6: Promoters tested for in vitro transcription 129
Table 4-6: Genes differentially expressed in serum compared to RB50Δ*rseAB* 138
Table 4-7: Genes differentially expressed in Bvg$^+$ compared to RB50Δ*rseAB* 139
Table 4-8: Strains and plasmids 143
Table 4-9: Primer sequences 148

Table 5-1: C-terminal three amino acids from annotated porins in *B. bronchiseptica* 156

Table A-1: Candidate sRNAs predicted by SIPHT 181
Table A-2: Candidate sRNAs in *B. bronchiseptica* identified by SOLiD sequencing.....204
Table A-3: sRNAs in common to both SIPHT predictions and SOLiD sequencing 212
Table A-4: Predicted targets for selected candidate sRNAs 216
Table A-5: Putative SigE-regulated sRNAs 219

Table B-1: Strains and plasmids.....261

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Sarah Ades, for all her guidance and encouragement through even the toughest of troubleshooting. She has helped me see new ways of approaching problems and finding answers. I would also like to thank the other members of the Ades lab, past and present, for all their helpful scientific discussion, extra hands for complicated experiments, and general camaraderie during my time at PSU: Jenny Hayden, Saumya Golpalkrishnan, Herve Nicoloff, Mary Laubacher, and Alessandra Costanzo. I would especially like to thank the undergraduate researchers that have helped teach me how to mentor, and contributed data to this dissertation: Catherine Shea, Airina Lee, Matt D'Olonzo, Nur Assyakirin Mohamed Paid, Jarinil Perez, and Brad Bauer.

Many collaborators made significant contributions to this research. The majority of this dissertation results from studies done in collaboration with Eric Harvill's lab in the Department of Biomedical and Veterinary Sciences. I would like to thank Eric for his resources and helpful discussion, and would especially like to thank Xuqing Zhang and Sara Hester for their significant contributions to this research. Additionally, other members of the Harvill lab have been important collaborators for various portions of this work: Laura Weyrich, Sarah Muse, Jihye Park, Ying Zhang, Liron Bendor, David Place, and Laura Goodfield.

I would also like to thank some additional collaborators: Tracy Nicholson at the USDA-ARS in Ames, IA, for all her help with the microarray analysis; Maria Rodriguez at La Plata University in La Plata, Argentina, for the phagocytosis experiments, and Amy Banta and Rick Gourse at the University of Wisconsin-Madison for both strains and for

their work looking at the interaction of ppGpp and DksA with *E. coli* σ^E in vitro. Other gifts of strains and resources that have been important to this research have come from: Ken Keiler, Paul Babitzke, Andrew Preston, Sandra Armstrong, Katsu Murakami, Carol Gross, and Scott Stibitz.

A special word of thanks to B. Tracy Nixon for his help designing and coding the program for scanning the *B. bronchiseptica* genome, and to the other members of my committee, past and present: Paul Babitzke, Biao He, and Ed Dudley, for their helpful discussion and advice.

I especially thank my friends and family for all the support and encouragement they have provided over the years: Sarah Matthews, Beth Davidson, Liz Billgren, Jenny Hayden, and the many other friends I have made in my time here at Penn State for their friendship and uncanny ability to distract me and help me to get away from the lab to re-focus; my parents, Jeff and DeLaine Barchinger, and my sister Kelly, who have provided constant encouragement; Roo whose unfailing joy and energy prevents bad moods, and Jon Schueth, whose hugs can cure most any problem.

Chapter 1

Introduction

In order to survive, all organisms must be able to sense and respond to changes in their environment. Bacterial pathogens have to cope with not only environmental stress such as changes in nutrient availability, temperature, or pH upon infection of a host, but must also survive attacks from the immune system, such as host-generated antimicrobial peptides and the oxidative burst associated with phagocytosis [1-4]. Positioned as the interface between the cytoplasm of a bacterial cell and the external environment, the cell envelope is both the sensor and the first line of defense against changing and stressful conditions [5, 6]. This chapter will describe how different bacterial species have adapted the components of a system that senses and responds to cell envelope stress, the σ^E system, to combat particular stress conditions and promote virulence.

Sensing stress at the Gram-negative cell envelope

Gram-negative bacteria have a cell envelope comprised of an outer membrane and an inner membrane, between which is the periplasmic space (Fig. 1-1). Peptidoglycan, composed of chains of glycans cross-linked by peptides, resides in the periplasm, and provides the cell its shape [7, 8]. The inner membrane is a symmetric bilayer composed primarily of phospholipids and proteins. The outer membrane is asymmetric: the inner leaflet contains mostly phospholipids, but the outer leaflet is composed primarily of lipopolysaccharides (LPS) [9, 10]. LPS consists of a lipid moiety (lipid A) anchored in the

outer leaflet of the outer membrane, linked to a core oligosaccharide, ending with a variable polysaccharide termed O-antigen [11, 12]. The LPS of pathogenic bacteria can instigate a robust host immune response, and can be modified to evade this response, to increase resistance to antimicrobial peptides, and in response to other environmental changes [13, 14].

A number of important proteins and protein complexes are membrane-bound or membrane-associated. Lipoproteins, for instance, are anchored to the membrane via lipid modifications, and have diverse and important functions in nutrient acquisition or even virulence [15]. Because the outer membrane creates a good barrier to prevent compounds from entering the cell, outer membrane porins (OMPs) create channels of specific sizes across the outer membrane to allow nutrients and ions to enter the cell [16]. Maintenance of all these envelope components is vital to cell survival. Many large multi-subunit structures span the envelope to extend into the external environment, including flagella and various secretion systems that export effector molecules into the surrounding environment or directly into a host cell [7, 17-19]. In addition to secretion systems, many other virulence factors expressed by pathogenic bacteria can be found at the envelope, such as adhesins, fimbriae, and autotransporters [20-22]. The cell envelope is not a static structure, and is constantly being remodeled in response to changes in the environment and during the course of infection. Therefore, sensing stress or changes at the envelope is both critical for survival and important for pathogenesis.

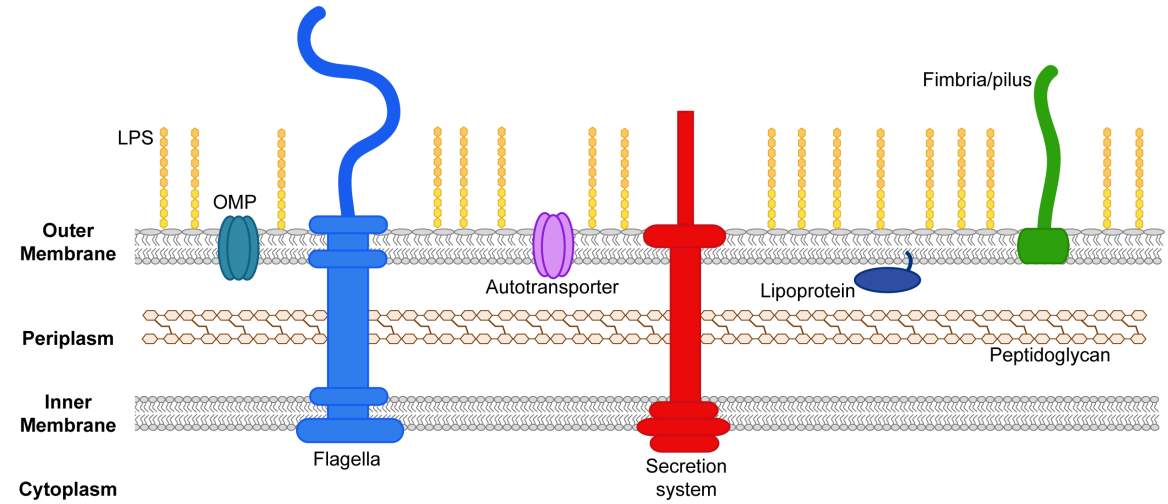


Figure 1-1: The Gram-negative bacterial cell envelope. Many of the structures at the cell envelope are also important for pathogenesis, including lipopolysaccharides (LPS), flagella, outer membrane porins (OMPs), autotransporters, lipoproteins, fimbriae and pili, and secretion systems.

Bacteria have evolved multiple methods of relaying information about changes sensed at the cell envelope to the transcriptional machinery in order to respond to these changes, including activation of an enhancer-binding protein required for transcription, two-component phosphorelay systems, or activation of an extracytoplasmic function sigma factor [23]. In the phage shock protein (Psp) system of *E. coli,* stresses such as overproduction of secretins and disruption of the proton motive force activate the enhancer-binding protein PspF to promote transcription and respond to stress [5, 24, 25]. For two-component systems, a stimulus usually triggers a series of events that results in phosphorylation of a specific residue or residues on the transmembrane sensor kinase component of the system. This phosphate is then transferred ultimately to the response regulator, which is activated to either enhance or block transcription of a particular subset of genes [5, 23]. The BvgAS two-component system of the bordetellae, described in more detail later, senses unknown environmental signals, which triggers up- or down-regulation

of the major virulence factors [26, 27]. The two-component systems CpxAR, BaeRS, and Rcs have been shown to sense and respond to cell envelope stress in *E. coli* [6, 9, 28]. Bacteria can also respond to changes sensed at the envelope by activating an alternative sigma factor, directing RNA polymerase to a specific subset of promoters in order to cope with the change.

Extracytoplasmic function sigma factors

RNA polymerase (RNAP) is a multi-subunit complex that carries out transcription of a DNA template into RNA. The core bacterial RNAP is comprised of two α, one β, one β', and one ω subunit [29]. Sigma factors are the dissociable subunits of bacterial RNAP that confer promoter specificity and associate with core RNAP to form holoenzyme (Eσ) [30, 31]. Two phylogenetically distinct classes of sigma factors can be found in bacteria, σ^{70} and σ^{54} [32, 33]. σ^{54}-family sigma factors recognize sites centered at -26 and -12 nucleotides upstream of a gene's transcriptional start site, and require bacterial enhancer binding proteins and DNA bending proteins for transcription initiation [34].

σ^{70}-family sigma factors recognize sites centered at -35 and -10 nucleotides upstream of a gene's transcriptional start site, and generally do not require accessory proteins to initiate transcription [33, 35]. The structure of σ^{70}-like sigma factors can be divided into four distinct domains. Two of these are conserved in all σ^{70}-family sigma factors, regions 2 and 4, parts of which are required for interactions with promoter DNA. Region 2.4 is involved in recognizing the -10 sequence, region 4.2 in recognizing the -35 sequence, and region 2.3 is involved in promoter melting. Regions 1 and 3 are primarily conserved in the housekeeping and very closely related sigma factors, and are thought to

be involved in antagonizing the DNA-binding activity of sigma and interacting with extended -10 elements, respectively [33, 36-38].

The namesake of this family, σ^{70}, is a housekeeping sigma factor responsible for the bulk of transcription in the cell, including genes involved in ribosome and amino acid biosynthesis [33]. Sequence homology has classified other, alternative, σ^{70}-like sigma factors into three additional groups: group 2, which are most related to σ^{70}, and include the stationary phase sigma factor σ^{S}; group 3, including sporulation sigma factors and the heat shock sigma factor σ^{32}; and group 4, which is the most diverse group of sigma factors, the extracytoplasmic function (ECF) sigma factors [33, 39, 40].

ECF sigma factors have been identified in at least 369 sequenced genomes across nine bacterial phyla [40]. The average bacterial genome contains six ECF sigma factors, but each species can encode from zero (*Rickettsiales, Mycoplasma, Borrelia*) to 40-50 (*Acidobacteria, Actinobacteria, Bacteriodetes*) ECF sigma factors, with the largest number found in *Sorangium cellulosum*, which encodes 83 ECF sigma factors [40]. ECF sigma factors were originally identified for their role in sensing extracytoplasmic stress, but some ECF sigma factors, defined by sequence similarity, actually sense cytoplasmic stress, such as the σ^{E}/ChrR system of *Rhodobacter sphaeroides*, which responds to singlet oxygen [41, 42].

Although this group is large and diverse, there are some characteristics that are common to most ECF sigma factors [39]. These proteins lack regions 1 and 3, described above, but share highly conserved regions 2 and 4. Many, but not all, of the currently described ECF sigma factors are co-transcribed with an anti-sigma factor, and sometimes additional regulatory proteins, which sequester the sigma factor away from RNAP until an

inducing signal is sensed [43]. These sigma factors are usually activated by either conformational changes in the anti-sigma factor or by regulated proteolysis, which degrades the anti-sigma factor and releases the sigma factor, freeing it to bind RNAP and direct transcription of its regulon [44, 45]. The promoter sequences recognized by ECF sigma factors often contain an AAC motif in the -35 region [39, 40]. Starón, et al., have recently classified the ECF sigma factors into groups based on sequence similarity, gene context and type of associated anti-sigma factor [40]. The research presented in this dissertation focuses on a member of ECF02 (RpoE-like sigma factors), which includes *E. coli* σ^E and its orthologs in other species.

σ^E

Evolutionary conservation of the σ^E system

The RpoE-like sigma factors form a group of highly similar sigma factors distributed primarily across the proteobacteria [40]. As mentioned above, these sigma factors are co-transcribed with an anti-sigma factor, and often other genes involved in their regulation. In many cases, the gene encoding σ^E is the first gene in an operon, and has a σ^E-dependent promoter (in addition to at least one σ^{70}-regulated promoter) [46, 47]. However, in a few instances, σ^E does not appear to be the first gene in its operon (Fig. 1-2: *Acidovorax, Polaramonas,* and *Bordetella bronchiseptica*), and it is unknown whether σ^E is still auto-regulated in these systems.

In addition to the sigma and anti-sigma factor, there is some conservation of genomic context for the RpoE-like sigma factors, but not in all bacteria (Fig. 1-2). For instance, in many of the γ-proteobacteria, particularly the enteric pathogens where the σ^E

system has been studied in the greatest molecular detail, the σ^E operon is divergently transcribed from *nadB*, which encodes an NAD biosynthesis enzyme. In some β-proteobacteria, such as *Polaramonas* spp*., Burkholderia* spp., and the bordetellae, the σ^E operon is instead surrounded by genes important in fatty acid biosynthesis. However, not all γ-proteobacteria and β-proteobacteria have these distinct differences in gene context. In the β-proteobacteria *Azoarcus*, the area surrounding the σ^E operon includes both *nadB* and fatty acid biosynthesis genes, and in many other γ-proteobacteria, the σ^E operon is not near *nadB* or *fabF*, but other unique genes, such as *leuA* (leucine biosynthesis) in *Actinobacillus pleuropneumoniae*, or *mscL* (large conductance mechanosensitive channel) in *Haemophilus influenzae* (Fig. 1-2).

As gene context may denote related functions in bacterial genomes, it is possible that the RpoE-like sigma factors, though sharing high sequence conservation, have diverged in function in different bacterial species. As will be described below, different bacterial species, even those where the σ^E locus shares conserved gene context, require the σ^E system for response to a different battery of stress conditions [48-52]. Factors such as environmental niche and presence or absence of other transcriptional regulators that sense similar conditions may have contributed to this divergence of function across species.

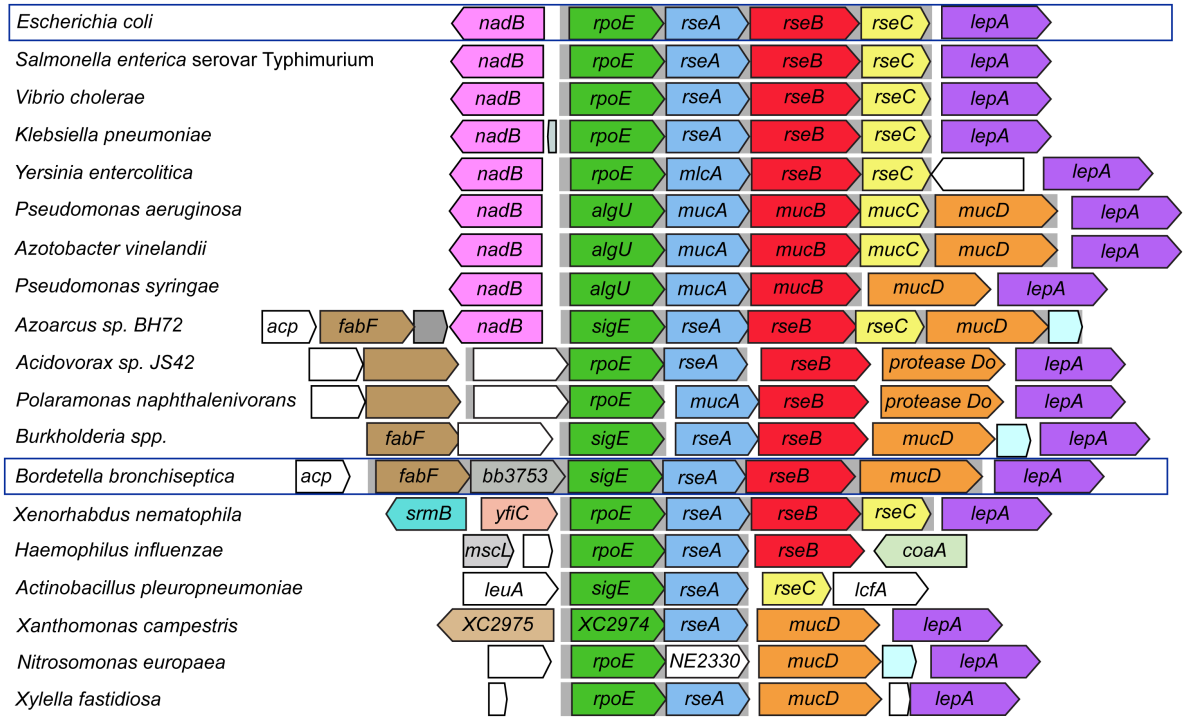


Figure 1-2: Gene context of RpoE-like sigma factors in multiple species. Grey boxes indicate the operon or predicted operon that includes the gene encoding σ^E in each species. The RpoE-like sigma factor operons that are the focus of this work, those of *Bordetella bronchiseptica* and *Escherichia coli,* are indicated with blue boxes.

Basic components of the σ^E system

In most bacteria, the gene encoding σ^E is co-transcribed with its negative regulators, often annotated *rseA* and *rseB* (for Regulator of SigmaE) [53]. RseA is a transmembrane protein; the cytoplasmic N-terminal domain binds tightly to σ^E, and the periplasmic C-terminal domain binds RseB, which strengthens inhibition of σ^E activity [53-57] (Fig. 1-3). In some bacteria, such as *E. coli,* RseA is more important for negatively regulating σ^E than RseB; deletion of *rseB* results in only a modest two-fold increase in σ^E activity, while deletion of *rseA* results in a 25-fold increase [57]. However, in other

bacteria, such as *Pseudomonas aeruginosa*, RseA and RseB contribute equally to inhibition of σ^E activity [58, 59]. Yet other bacteria, such as *Actinobacillus pleuropneumoniae*, encode the anti-sigma factor RseA, but do not encode a known ortholog of RseB, suggesting that the regulatory pathways to activate σ^E vary among species [60].

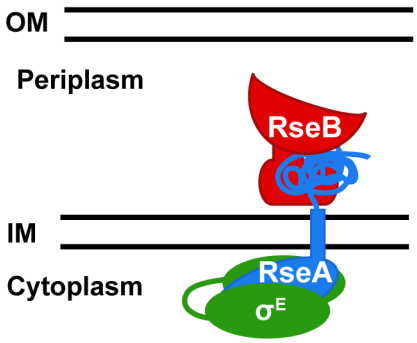


Figure 1-3: The basic components of most known σ^E systems: σ^E, shown in green, is held at the inner membrane by the transmembrane anti-sigma factor RseA (blue), and the periplasmic protein RseB (red).

Conserved and varied roles for σ^E systems

In *E. coli*, *Yersinia* spp., and probably *Vibrio cholerae*, the gene encoding σ^E is essential [61-63]. In other bacteria, the gene encoding σ^E, while dispensable for survival, is required for response to various stress conditions and for virulence [32, 64-66]. As a transcription factor, σ^E directs transcription of many genes that may provide some insight into the role this sigma factor plays in various bacteria.

Rhodius, et al. described a core σ^E regulon predicted to be transcribed by σ^E in multiple species, including the heat shock sigma factor σ^{32}, the σ^E operon, and some lipoproteins [48]. Over half of the predicted core σ^E regulon members are genes important

in LPS biosynthesis and OMP assembly. This observation is consistent with the role σ^E is known to play in response to cell envelope stress for many bacteria. In addition to this small core σ^E regulon, there are many unique genes regulated by σ^E that do not overlap between even very closely related species, such as *Salmonella* spp. and *E. coli*. This is further indication that while the components of the σ^E system (σ^E, RseA, and RseB) are fairly well conserved, this system has been adapted for different functions across bacterial species.

Even in closely related bacteria, σ^E systems sense and respond to different arrays of stress conditions. σ^E systems have been studied in the most detail in the enteric bacteria *E. coli*, *S. enterica* serovar Typhimurium, and *V. cholerae*, but have been studied in many other bacteria, including the opportunistic pathogen *Pseudomonas aeruginosa*, the respiratory pathogens *Burkholderia* spp., and a handful of other species. The following sections briefly describe how the basic, conserved components of the σ^E system have been adapted differently by the different bacterial species in which it has been studied.

σ^E systems in diverse bacterial species

Escherichia coli

σ^E was first discovered for its role in transcribing the heat shock sigma factor, σ^{32}, in *E. coli* [67]. Subsequent work has determined that σ^E is encoded by the essential gene *rpoE* in *E. coli*, where it is required for response to extreme heat shock and ethanol stress [68-70]. σ^E is also required for maintaining the integrity of the cell envelope. Depleting σ^E by overexpressing RseA and RseB causes membrane defects, and multi-copy suppressors of the essentiality of σ^E also suppress these membrane defects, although the mechanism by

which this happens is not yet clear [71]. σ^E has long been known to be activated by misfolded outer membrane proteins (OMPs) [68, 72, 73], and the regulated proteolysis pathway triggered by these misfolded OMPs, which releases σ^E from its anti-sigma factor (described below and extensively in Appendix C) has been studied in the most molecular detail in *E. coli* [45, 74, 75].

Under non-stress conditions, most σ^E is held at the inner membrane by its transmembrane anti-sigma factor, RseA. This inhibition is strengthened by the interaction of the periplasmic domain of RseA with RseB (Fig. 1-4, left panel). When the cell encounters a stress that results in exposing the C-terminus of misfolded OMPs (Fig. 1-4, right panel), the protease DegS is activated, which cleaves RseA in the periplasm. This provides a substrate for a second membrane-associated protease, RseP, which cleaves RseA in the inner membrane. The cytoplasmic domain of RseA is then degraded by the ClpXP degradasome, freeing σ^E to associate with core RNAP and direct transcription of its regulon. For more detail about the proteolytic pathway that releases σ^E from RseA, see Appendix C.

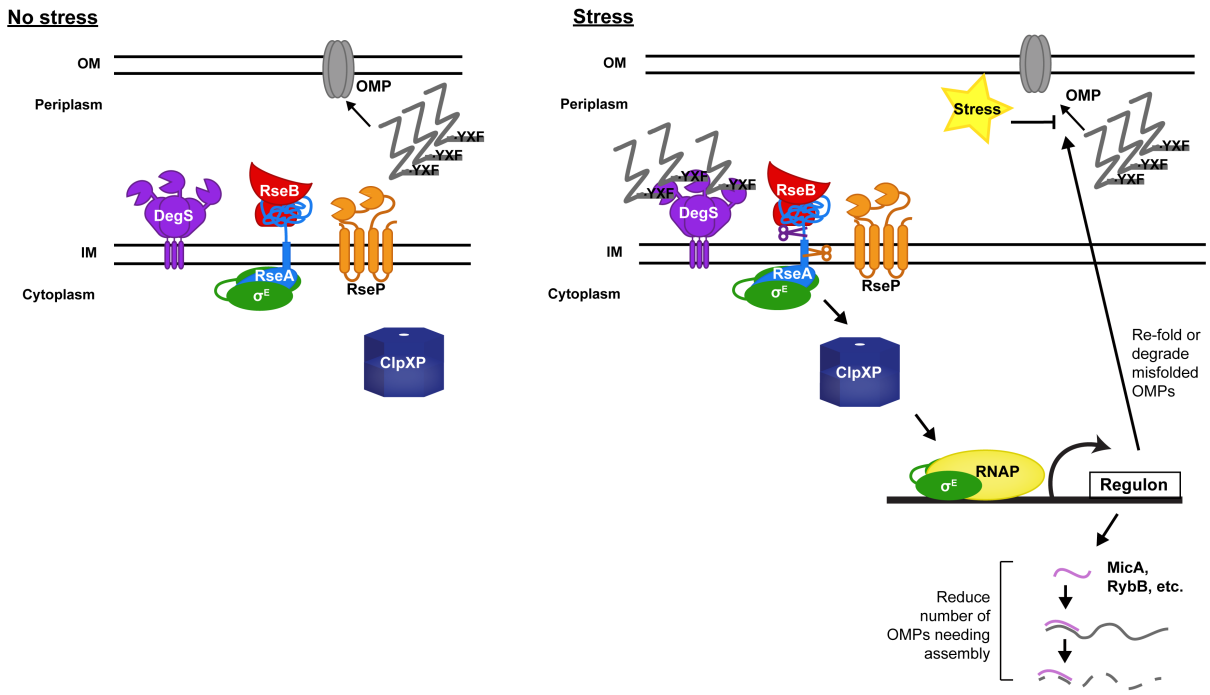


Figure 1-4: σ^E regulation in *E. coli*. On the left, most σ^E is sequestered at the membrane under non-stress conditions. On the right, when the cell encounters a stress, regulated proteolysis frees σ^E to bind core RNAP and direct transcription of its regulon.

E. coli σ^E directs transcription of nearly 100 genes, including many genes important for cell envelope maintenance and stress response. In fact, most of the machinery required to assemble outer membrane proteins and lipopolysaccharides is transcribed by σ^E [48, 76]. This includes biosynthetic enzymes, chaperones important for facilitating folding of mature OMPs or assembling LPS, and proteases to degrade aberrantly folded proteins [7, 48]. *E. coli* σ^E also transcribes many other stress-responsive genes, including the alternative sigma factors σ^{32} and σ^{54} and cytoplasmic proteases, such as *clpX* and *lon*. σ^E also regulates transcription of at least two small non-coding RNAs (sRNAs), RybB and MicA, in *E. coli* [77-79]. These sRNAs, in conjunction with the sRNA-binding protein Hfq, bind to the 5' untranslated region (UTR) of the genes encoding

the outer membrane porins OmpC and OmpA, respectively, targeting these mRNAs for degradation [77, 79]. Thus, the role σ^E plays in relieving the stress of misfolded OMPs is at least two-fold: 1) increasing transcription of genes that aid in the degradation of aberrant proteins and folding and assembly of newly translated OMPs, and 2) decreasing OMP production through targeted degradation of OMP mRNAs, thereby reducing the burden on the system as a whole (Fig. 1-4) [80]. Recent work has demonstrated that these sRNAs regulate expression of multiple targets, suggesting a wider role for σ^E-regulated sRNAs than previously thought [80].

The gene encoding σ^E, *rpoE,* is co-transcribed with the genes encoding its negative regulators, the anti-sigma factor *rseA*, a second negative regulator *rseB*, and a third poorly-characterized gene, *rseC* (Fig. 1-2). The *rpoE* operon is included in the σ^E regulon, so that when σ^E is activated, as described above, transcription of both σ^E and its negative regulators increases [69, 70]. *rseA* also has an additional σ^E-dependent promoter, so that excess σ^E can be quickly inhibited when the activating conditions have been removed [48, 81].

Previous work in our lab uncovered that *E. coli* σ^E is also regulated independently of RseA by the alarmone guanosine 3'-5'-bispyrophosphate (ppGpp) and the protein DksA [82]. During stress conditions such as amino acid and phosphate starvation or entry into stationary phase, ppGpp levels rise [83, 84]. ppGpp and DksA interact with RNAP, decreasing σ^{70}-dependent transcription from ribosomal RNA promoters and increasing transcription from σ^{70}-dependent amino acid biosynthesis promoters [85, 86]. These two factors are also implicated in increasing the activity and/or abundance of the alternative sigma factors σ^{54} and σ^S [87-90].

A couple of models have been proposed as to how ppGpp and DksA-mediated regulation of alternative sigma factor activity occurs. One possibility is that these factors interact directly with the sigma/core RNAP holoenzyme to activate transcription. An alternative explanation is that when ppGpp and DksA decrease transcription from rRNA promoters, which encompasses the majority of σ^{70}-dependent transcription during exponential growth, the pool of core RNAP available to bind alternative sigma factors increases [91]. In the case of *E. coli* σ^E, there is evidence for both direct and indirect activation (see Appendix B for a more detailed discussion) [92].

The *E. coli* σ^E system is generally considered the paradigm for the regulation and role of orthologous σ^E systems in other bacteria. While there are many similarities in how σ^E is regulated and the role it plays in cell envelope stress response, as described below, there are also many differences between species in the specific conditions that activate σ^E and the genes that comprise the σ^E response, even in bacteria closely related to *E. coli*, such as *Salmonella enterica*.

Salmonella enterica serovar Typhimurium

E. coli and *S. enterica* serovar Typhimurium are closely related enteric bacterial pathogens. The regulation of σ^E is conserved, as well as some members of the regulon, including the OMP-targeting sRNAs described above [93, 94]. However, each has adapted the σ^E system to respond to different stimuli. *E. coli* σ^E is essential, and has been shown to be important in response to heat shock and ethanol stress, but not oxidative stress [61, 69, 70]. *S. enterica* serovar Typhimurium σ^E is neither essential nor required for survival

during heat shock. It has, however, been implicated in response to cold shock, oxidative stress, carbon starvation, resistance to antimicrobial peptides, and in virulence [20, 52, 65, 95-97].

S. enterica serovar Typhimurium σ^E is particularly important for response to oxidative stress. A mutant lacking the gene encoding σ^E, *rpoE*, is more sensitive to treatment with hydrogen peroxide and paraquat [52, 65]. An *rpoE* mutant (Δ*rpoE*) is attenuated for virulence in wild-type mice [97]. In a phox$^{-/-}$ mouse, however, which cannot generate superoxide because it lacks the gene encoding a NADPH phagocyte oxidase, Δ*rpoE* is no longer attenuated [52, 65]. Macrophages also generate an oxidative burst as an effective means to combat bacterial pathogens [4]. While Δ*rpoE* still enters macrophages efficiently, it does not survive as well as wild-type *S. enterica* serovar Typhimurium intracellularly [97]. Together, these suggest that σ^E is particularly required for response to oxidative stress in *S. enterica* serovar Typhimurium, demonstrating how σ^E systems can be utilized by different bacterial species, such as *S. enterica* serovar Typhimurium and *E. coli*, to respond to different stresses. Differences in how σ^E systems have been adapted can even be observed between species in the same genus, such as with *Vibrio cholerae/Vibrio vulnificus* and *Pseudomonas aeruginosa/Pseudomonas syrinage/Pseudomonas fluorescens,* described below.

<u>*Vibrio* spp.</u>

In *Vibrio cholerae*, the etiological agent of cholera, σ^E is encoded by *rpoE.* Cells lacking *rpoE* are more sensitive to heat shock, ethanol, and antimicrobial peptides, and have a higher LD_{50} (number of bacteria required to cause death of 50% of the mice) in

mice than wild-type *V. cholerae*, indicating a reduced ability to cause disease [98, 99]. There is some evidence that *rpoE* may, in fact, also be essential in *V. cholerae*; high-throughput sequencing of multiple Δ*rpoE* strains revealed suppressor mutations, most of which reduced expression of a major outer membrane protein [63]. Because misfolded OMPs are the only known signal to initiate release of σ^E from its anti-sigma factor in many bacteria, mutations that lower expression of a major OMP presumably lower the amount of protein that can accumulate, thus lowering the requirement for σ^E [63].

V. cholerae lacking the negative regulator *rseA* does, indeed, show elevated levels of σ^E, but does not have an obvious virulence phenotype . However, this mutation seems unstable, often resulting in suppressor mutations [100]. In the related *V. vulnificus*, an estuarine bacteria that can cause both intestinal and wound infections, an *rseB* deletion has been shown to be attenuated for virulence, where an *rpoE* deletion is not attenuated [101]. As *V. cholerae* lacking *rpoE* does exhibit decreased virulence, this highlights how the role of σ^E can differ between even very closely related bacteria [98]. The colony morphology phenotype associated with a *V. vulnificus rseB* deletion also seems unstable [101], suggesting that regulation of σ^E activity is important for cell viability in *Vibrio* species.

<u>*Pseudomonas* spp.</u>

Pseudomonas aeruginosa

In the opportunistic pathogen *Pseudomonas aeruginosa*, σ^E is encoded by the gene *algU*, named for its role in promoting biosynthesis of the polysaccharide alginate [102-105]. Co-transcribed with *algU* are its negative regulators *mucA* and *mucB, mucC,* and the

HtrA-family protease *mucD*, the names of which are derived from the mucoid phenotype that occurs when alginate is produced [106, 107].

Expression of AlgU in *E. coli* can complement an *rpoE* deletion, indicating that *P. aeruginosa* AlgU is similar enough to *E. coli* σ^E that it can carry out its essential function [108]. The method by which σ^E is activated is conserved, but species-specific [58]. In *E. coli,* when OMPs become unfolded, their C-terminus binds to the PDZ domains of the DegS trimer, activating it to cleave RseA in the periplasm, initiating a proteolytic cascade that results in release of σ^E to direct transcription of its regulon [53, 73, 75]. The *P. aeruginosa* ortholog of DegS is AlgW, which is activated by different sequences. In particular, the C-terminal amino acids of the small *P. aeruginosa* OMP MucE (-WVF), but not of *E. coli* OMPs (-YXF), activate AlgU in an AlgW-dependent manner [109, 110].

Cells lacking *algU* are more sensitive to oxidative stress generated by incubation with paraquat or hypochlorite, heat stress, and are more susceptible to killing by phagocytes [49, 64]. Cells lacking σ^E in many other bacteria (*S. enterica* serovar Typhimurium*, V. cholerae*, e.g.) are attenuated for virulence. However, in two different mouse models of *P. aeruginosa* infection, Δ*algU* was actually more virulent: 1) in a neutropenic mouse model, immunocompromised because of low numbers of neutrophils, the time it took for *P. aeruginosa* lacking *algU* to cause death was shorter than similar numbers of wild-type bacteria [64], and 2) in wild-type C57BL/6 mice, the LD_{50} of Δ*algU* was lower than wild-type, suggesting this strain is actually more virulent than wild-type *P. aeruginosa* in this infection model [64]. In contrast, the LD_{50} of a mutant lacking *mucD*, a homolog of the *E. coli* σ^E-regulated periplasmic serine protease *degP*, was higher,

indicating that this gene is required for full virulence. *P. aeruginosa* lacking *mucD* is also more sensitive to oxidative stress in disk diffusion assays, similar to Δ*algU* [64].

Most isolates of *P. aeruginosa* taken from cystic fibrosis (CF) patients have a mutation in the anti-sigma factor *mucA*, rendering σ^E constitutively active [103, 111, 112]. Because AlgU positively regulates the genes important for alginate biosynthesis, constitutive activation of AlgU generates large amounts of alginate, promoting biofilm formation [113]. Conversion to mucoidy appears to offer some advantage for *P. aeruginosa* colonization and persistence in the lungs of CF patients, but how this is achieved is less clear [112, 114, 115]. Many mucoid isolates are less virulent in animal models of disease, and this is correlated with lower expression of many virulence determinants, such as type three secretion system function, protease production, exotoxin A, and motility [114, 115]. It has been proposed that down-regulation of virulence factors benefits the type of chronic infection seen in CF patients [114], suggesting that maintaining the proper regulation of σ^E may be particularly important during the course of infection.

Some work has been done to identify members of the *P. aeruginosa* AlgU regulon. In addition to the alginate biosynthesis pathway, the genes transcribed by AlgU include *rpoH*, which is also transcribed by σ^E in *E. coli*, genes associated with osmotic stress (*osmC* and *osmE*), and some lipoproteins, which have been shown to induce a cellular immune response in vitro [116]. This is particularly important because although *P. aeruginosa* isolates from cystic fibrosis patients have decreased many of their virulence factors to establish a chronic, persistent infection in the lungs, inflammation ultimately causes enough lung damage to impair respiratory function [114, 116-118].

Pseudomonas syringae and Pseudomonas fluorescens

The operon structure of the *Pseudomonas syringae* AlgT (AlgU) system is different from that of the closely related *P. aeruginosa*. *P. syringae* lacks *mucC*, and transcribes *mucD* independently from *algT, mucA,* and *mucB*, unlike in *P. aeruginosa*, where *algU, mucA, mucB, mucC,* and *mucD* are all co-transcribed (Fig. 1-2) [59, 119, 120]. However, in this phytopathogen, similarly to *P. aeruginosa,* cells lacking *algT* are also more sensitive to osmotic, oxidative, and heat stress [49]. Cells lacking *algT* actually grow better in vitro than wild-type, but do not colonize plant leaves as well [121]. This decrease in virulence is different from *algU* mutants of *P. aeruginosa,* which display enhanced virulence in some infection models [64], again demonstrating how even in closely related species, such as *P. aeruginosa* and *P. syringae*, which share a role for σ^E in response to isolated stress conditions such as temperature and oxidative stress, σ^E has different effects on virulence [49, 64, 104, 120].

Conversely, although *P. syringae* and the plant-beneficial bacteria *P. fluorescens* share a conserved operon structure, *P. fluorescens* lacking *algU* is not more sensitive to oxidative or heat stress, but is more sensitive to osmotic stress and dessication [122]. These examples demonstrate how, while the various components of σ^E systems are largely conserved, these systems can evolve to fit the different needs of each species, and even closely related bacteria can have distinct differences in the role and regulation of this system. Although σ^E is just beginning to be studied in the respiratory pathogens of the *Burkholderia* spp., differences in how each use the σ^E system have already been observed.

Burkholderia spp.

Burkholderia pseudomallei is a bacteria typically found in soil that can cause diseases such as melioidosis in humans. It is endemic in Southeast Asia, and often results in fatal septicemia. *B. pseudomallei* is known to be particularly resistant to stress conditions, which is partly why it is considered a potential bioterrorism agent [123, 124]. In *B. pseudomallei*, the σ^E operon consists of *rpoE-bprE(rseA)-rseB-mucD* (Fig. 1). *rpoE* is required for survival during heat shock, osmotic stress, and oxidative stress [123, 124]. Unlike in some bacteria, *rpoE* mutants display altered colony morphology, large intracellular vacuoles, and tend to chain instead of clustering [123]. *rpoE* has also been shown to be involved in biofilm formation, though its exact role is unclear [123]. Cells lacking *rpoE* survive less well in macrophages, similarly to σ^E mutants in *Haemophilus influenzae*, *S. enterica* serovar Typhimurium, and, as this dissertation will demonstrate, *Bordetella bronchiseptica* (Chapter 2) [97, 123, 125].

Burkholderia cepacia and *Burkholderia cenocepacia* are members of the *Burkholderia cepacia* complex (Bcc), which, similarly to *P. aeruginosa*, are opportunistic pathogens often isolated from cystic fibrosis patients [126]. Bcc species are also very common to the rhizosphere. Because of the wide range of niches Bcc species inhabit, this complex is generally considered to be stress tolerant [50]. σ^E in *B. cenocepacia*, encoded by *rpoE*, is required for response to heat stress and osmotic stress, but not oxidative stress, unlike in the closely related *B. pseudomallei*, which requires σ^E for response to oxidative stress [50, 123]. *B. cenocepacia* lacking *rpoE* also has altered interactions with macrophages, and is unable to delay fusion with phagolysosomes, an important step in intracellular survival [50].

Other bacterial species

σ^E has been identified in at least 122 known bacterial genomes [40]. As demonstrated in the previous sections, the role σ^E plays in response to particular stress conditions and the details of its regulation vary from species to species, even between closely related species, such as *P. aeruginosa* and *P. syringae, B. pseudomallei* and *B. cenocepacia*, or even *E. coli* and *S. enterica* serovar Typhimurium. The following section briefly describes the role of σ^E in bacterial species where, to date, this sigma factor has been identified as important in stress response or virulence, but has been studied in less detail than in the species described earlier.

In the phytopathogen *Xanthomonas campestris* pathovar campestris (*X. campestris*), the gene encoding σ^E, *rpoE*, is co-transcribed with the genes encoding an anti-sigma factor *rseA*, and a serine protease, *mucD* [127, 128]. *rpoE* is required for survival during stationary phase, heat shock, ethanol stress, and in the presence of cadmium, but does not seem to play a role in response to oxidative, osmotic, acid, or detergent stress [127]. Orthologs of the proteases DegS and RseP were identified; however, the contribution of each is not completely consistent with what is known for *E. coli.* In *X. campestris,* σ^E activity increases during heat shock even in cells lacking the DegS and RseP orthologs, suggesting that other proteases may be able to activate σ^E [127]. This could be because the *X. campestris* genome does not encode *rseB*, which, in *E. coli,* is important for preventing cleavage of RseA by RseP or other proteases before DegS cleaves RseA in the periplasm [129].

Xylella fastidiosa is a phytopathogen phylogenetically linked to *X. campestris* [130]. Like for *X. campestris*, *X. fastidiosa* σ^E is required for survival during heat shock

and ethanol stress, but is not required for response to oxidative or osmotic stress [130]. However, unlike in *X. campestris* or most other described σ^E systems, in *X. fastidiosa*, σ^E is not auto-regulated. There is no σ^E-dependent promoter upstream of *rpoE*, but there is one upstream of the gene encoding the anti-sigma factor *rseA* [130]. Differences observed in the regulation of *X. campestris* and *X. fastidiosa* σ^E systems demonstrates that although the locus shares significant sequence similarity, different bacterial species have developed even different methods of regulating these conserved components to fit the system for their own requirements.

Xenorhabdus nematophila lives as a symbiont with the nematode *Steinernema carpocapsae*, and this bacteria-nematode combination together are insect pathogens; both are required for insect killing. A mutant lacking *rpoE* in *X. nematophila* displayed decreased motility, lipase, and protease activities, and could no longer efficiently colonize its nematode host [131].

Similarly to *P. aeruginosa*, AlgU of the cyst-forming soil bacteria *Azotobacter vinelandii* plays a role in the regulation of the exopolysaccharide alginate, which is important for forming mature cysts that are dessication resistant [132, 133]. *A. vinelandii* AlgU, like *P. aeruginosa,* can complement a deletion of *rpoE* in *E. coli* [132]. Mutants lacking the entire *mucABCD* locus, or even just *mucA*, are unstable and readily accumulate spontaneous mutations [132]. Despite this, while *A. vinelandii* AlgU is required for encystment, constitutive activation of AlgU does not seem to affect the encystment process [132, 133].

In the porcine respiratory pathogen *Actinobacillus pleuropneumoniae,* constitutive activation of σ^E through deletion of *rseA* or overproduction of plasmid-encoded σ^E

increases biofilm formation through direct regulation of the polysaccharide PGA (poly-β-1,6-N-acetyl-D-glucosamine) locus [60, 134]. Despite increased biofilm production, a mutant lacking the anti-sigma factor *rseA* is attenuated for virulence, while a mutant lacking *rpoE* is not attenuated [60]. Other than the work presented in this dissertation (Chapter 3), this is one of the only examples where constitutive activation of σ^E is more detrimental to pathogenesis than deletion of the gene encoding σ^E.

RpoE-like sigma factors can be found in many and diverse proteobacterial species. Although this system is highly conserved at the sequence level, the examples described above clearly show that this system has been adapted by different species to respond to different conditions and regulate different subsets of genes. The cellular role of σ^E and its regulation vary widely from species to species. I have recently identified an RpoE-like sigma factor in the classical bordetellae, encoded by *sigE*. My dissertation research focuses on the role and regulation of the σ^E (SigE) system in *Bordetella bronchiseptica*, and how it has been adapted by this respiratory pathogen to respond to specific stress conditions and promote virulence.

Bordetella

The bordetellae are obligately aerobic β-proteobacteria. There are nine currently identified members of the Bordetella genus: *B. bronchiseptica, B. pertussis, B. parapertussis, B. avium, B. holmesii, B. hinsii, B. trematum, B. ansorpii,* and *B. petrii.* Of these, *B. pertussis, B. parapertussis,* and *B. bronchiseptica* are considered the "classical bordetellae," and are causative agents of respiratory diseases in humans and other

mammals [135, 136]. *B. pertussis* and *B. parapertussis* are the etiological agents of whooping cough in humans, and *B. bronchiseptica* can cause a wide range of disease in predominantly non-human mammal hosts [136-138]. The genomes of the classical bordetellae have been extensively studied, and it is thought that *B. pertussis* and *B. parapertussis* each independently evolved from a *B. bronchiseptica*-like progenitor strain through extensive gene loss [135, 139, 140].

Most work on transcriptional regulation in the bordetellae has focused on the BvgAS two-component phosphorelay. In this system, BvgS is the sensor kinase, and BvgA is the response regulator. A series of phosphorylations activates BvgA, allowing it to bind upstream of promoters in the BvgAS regulon and activate transcription [27, 141]. The exact environmental signals sensed by this system are unclear. What is known is that at low temperature (25 °C), or in the presence of nicotinic acid or magnesium sulfate, the system is off. The cells become motile, and expression of most virulence factors is downregulated. In the absence of these signals, the cells are considered Bvg^+. The expression of many known virulence factors, such as the type three secretion system (T3SS), adhesins, and toxins such as adenylate cyclase toxin or pertussis toxin, is increased, while expression of flagella and O-antigen is decreased [27, 142, 143]. BvgAS has long been considered the "master regulator" of virulence in the bordetellae. However, the work described in this dissertation will show that BvgAS is not the only important system affecting transcriptional control of virulence factors and envelope proteins in *B. bronchiseptica*.

B. bronchiseptica infects a broad range of hosts, and can cause a variety of disease, from asymptomatic infection to fatal pneumonia [136, 138]. In wild-type C57BL/6 mice,

the prototypic *B. bronchiseptica* strain RB50 asymptomatically colonizes the nasal cavity, trachea, and lungs. Bacterial numbers peak on day three post-inoculation, and bacteria are eventually cleared from the lower respiratory tract, but persist in the nasal cavity [144-147]. In mice lacking key parts of the immune response, such as the innate immune factors TNF-α and TLR4, or in mice lacking B cells and T cells ($Rag1^{-/-}$), RB50 can leave the respiratory tract, colonize systemic organs, and cause lethal disease [148-151]. Both courses of disease are important to study, as other strains of *B. bronchiseptica* cause systemic, lethal infection in mice and other hosts.

In addition to causing devastating disease in non-human mammals, from atrophic rhinitis in swine to kennel cough in dogs, and snuffles in rabbits, *B. bronchiseptica* can also cause severe, systemic infections in immunocompromised human hosts [137, 152, 153]. It is, therefore, imperative to understand the mechanisms by which these bacteria survive both within a host and in the environment between hosts. The SigE system is >99% identical between the classical bordetellae, suggesting that this locus plays an important and conserved role for each species. Understanding how the SigE system has been adapted to respond to specific stresses encountered by the bordetellae during infection will help provide insight into the regulation of virulence and persistence in these respiratory pathogens.

Chapter 2

sigE facilitates the adaptation of *Bordetella bronchiseptica* to stress conditions and lethal infection in mice

Abstract

The cell envelope of a bacterial pathogen can be damaged by harsh conditions in the environment outside a host and by immune factors during infection. Cell envelope stress responses preserve the integrity of this essential compartment and are often required for virulence. Although *Bordetella* species possess a large number of putative transcription factors, cell envelope stress responses have not been described in these important respiratory pathogens. Here we report that SigE of *Bordetella bronchiseptica* is a functional sigma factor that mediates a cell envelope stress response. Mutants of *B. bronchiseptica* strain RB50 lacking *sigE* are more sensitive to high temperature, ethanol, and perturbation of the envelope by SDS-EDTA and certain β-lactam antibiotics. Using a series of immunocompromised mice deficient in different components of the innate and adaptive immune responses, we show that SigE plays an important role in evading the innate immune response during systemic lethal infection. SigE is not required, however, for colonization of the respiratory tract of immunocompetent mice. The *sigE* mutant is more efficiently phagocytosed and killed by peripheral blood polymorphonuclear leukocytes (PMNs) than RB50, and exhibits decreased cytotoxicity toward macrophages. These altered interactions with phagocytes could contribute to the defects observed during systemic infection. Much of the work on transcriptional regulation during infection in *B.*

bronchiseptica has focused on the BvgAS two-component system. This study reveals that the SigE regulon is also required during infection, and mediates a discrete subset of functions associated with the response to cell envelope stress and virulence.

* Infection work was done by X Zhang, Cytotoxicity experiments were performed by SE Hester, and phagocytosis experiments were done by ME Rodriguez.

Introduction

The cell envelope of gram-negative bacteria is a dynamic, multifunctional cellular compartment. It provides structure to the cell, acts as a barrier against extracellular hazards, and is involved in important cellular processes, including energy production and nutrient transport. The envelope is also rapidly remodeled in response to changing environmental conditions. Gram-negative bacteria possess an array of stress responses that sense conditions in the cell envelope and alter gene expression to ensure its integrity [5, 23]. In many bacterial pathogens, cell envelope stress responses play a multifaceted role. They provide protection against damage caused by components of the immune system, such as complement and antimicrobial peptides that target the cell envelope [20, 52, 154]. They regulate the expression of chaperones required for proper assembly of cell envelope-associated structures, including outer membrane porins, pili, and fimbrae [52, 97, 155]. In addition, cell envelope stress responses can regulate the expression of virulence factors, ensuring that these factors are expressed at the proper time and location in the host [5, 6]. Despite their importance, no cell envelope stress responses have yet been identified or implicated in pathogenesis in *Bordetella* species.

Bordetella bronchiseptica is a respiratory pathogen that is closely related to *Bordetella pertussis* and *Bordetella parapertussis*, the causative agents of whooping cough in humans [135, 139]. *B. bronchiseptica* causes a range of diseases in various mammals that can be chronic, difficult to completely eradicate, and of variable virulence [136, 138, 156]. It is the etiological agent of atrophic rhinitis in swine, kennel cough in dogs, and snuffles in rabbits [136, 156]. Documented human infections, generally traced to an animal

source, have been observed in immunocompromised individuals, and can be serious, systemic infections [137, 156].

The *B. bronchiseptica*, *B. pertussis* and *B. parapertussis* genomes encode a large number of putative transcription factors relative to their overall genome size [157], suggesting that they have the capacity to extensively regulate gene expression in response to environmental and physiological changes. Despite this finding, only a few *Bordetella* transcription factors have been studied in any detail. Among the predicted transcription factors is an ortholog of the cell envelope stress response sigma factor, σ^E, of *E. coli*. In bacteria, sigma factors are the subunits of bacterial RNA polymerases required for specific promoter recognition and transcription initiation [33]. Alternative sigma factors, like σ^E, are activated in response to specific stresses and rapidly reprogram gene expression by replacing the housekeeping sigma factor and directing RNA polymerase to the genes in their regulons [33, 158].

σ^E belongs to the RpoE-like group of extracytoplasmic function (ECF) sigma factors that have been increasingly implicated as key factors contributing to both bacterial stress responses and virulence [39, 40]. These sigma factors are widely distributed across bacterial phyla. Where studied, they direct a diverse set of stress responses primarily targeted to the cell envelope [5, 6, 40, 47]. In *E. coli* and *Salmonella enterica* serovar Typhimurium, σ^E controls many genes whose products are required for the proper expression of outer membrane porins and lipopolysaccharides (LPS) [48, 53]. During infection, σ^E of *Salmonella enterica* serovar Typhimurium is required for survival and proliferation in epithelial and macrophage cell lines, and in the presence of antimicrobial peptides [65, 95, 97]. In *Pseudomonas aeruginosa*, the σ^E homologue, AlgU, controls the

expression of the exopolysaccharide alginate and conversion to mucoidy. AlgU is constitutively activated in many clinical isolates from cystic fibrosis patients [104, 159]. In addition, σ^E is required for the viability of some bacterial species, but not others. The gene encoding σ^E is essential in *E. coli* and *Yersinia enterocolitica*, but is dispensable in the closely related species *Salmonella enterica* serovar Typhimurium [61, 62, 97]. These observations suggest that the functions of σ^E orthologs have been adapted throughout evolution to combat the challenges each organism faces in its particular environmental niche.

Here we show that the *B. bronchiseptica* σ^E ortholog, encoded by the gene *sigE* (BB3752), is an active sigma factor that mediates a cell envelope stress response. This is the first demonstration of an envelope stress-sensing system in *Bordetella* species. Using a murine infection model, we demonstrate that SigE plays an important role during lethal systemic infection in immunocompromised mice, but not in respiratory tract colonization. This finding has important implications for human disease, because *B. bronchiseptica* has been observed to cause serious systemic infection in immunocompromised humans [137, 156]. This study suggests that SigE is a critical factor during infection, in addition to the BvgAS master virulence regulatory system.

Results

sigE encodes an active sigma factor

SigE of *B. bronchiseptica* shares a high degree of sequence conservation with other RpoE-like sigma factors from proteobacteria, particularly in regions predicted to interact with promoter DNA (Fig. 2-1A) [40]. To determine if SigE is an active sigma factor, we asked whether it could direct transcription from the σ^E-dependent *rpoH*P3 promoter in *E. coli*. This promoter shares a high degree of similarity with a consensus promoter proposed for the RpoE-like sigma factors that was determined from both experimental data and predicted promoter sequences (Fig. 2-1C) [40, 48]. The *sigE* gene from *B. bronchiseptica* strain RB50 was cloned into the pTrc99a expression plasmid and transformed into a derivative of *E. coli* MG1655 that carries an *rpoH*P3::*lacZ* reporter gene fusion integrated at the λ attachment site on the chromosome [68]. When *sigE* expression was induced, LacZ activity increased, indicating that SigE can initiate transcription from this promoter (Fig. 2-1B). Furthermore, we found that the gene encoding σ^E, *rpoE*, which is essential for viability in *E. coli*, could be deleted when *sigE* was overexpressed.

To provide additional evidence that SigE is a functional sigma factor, N-terminally His-tagged SigE was purified and tested for its ability to initiate transcription in vitro from the *E. coli rpoH*P3 promoter. Holoenzyme formed with SigE and *E. coli* core RNA polymerase (ESigE) was able to direct transcription and produced a transcript of equivalent length to that generated by *E. coli* $E\sigma^E$ (Fig. 2-1C). The region immediately upstream of the *B. bronchiseptica rpoH* homologue, encoded by the *fam* gene, contains a sequence that is highly similar to the proposed σ^E-dependent consensus promoter, suggesting that *B.*

bronchiseptica rpoH is regulated by SigE. Indeed, SigE was able to direct transcription from the putative *fam* promoter region in vitro (Fig. 2-1C). Taken together, these results demonstrate that SigE is a functional sigma factor and can initiate transcription from promoter sequences similar to those utilized by other members of the RpoE-like sigma factor family.

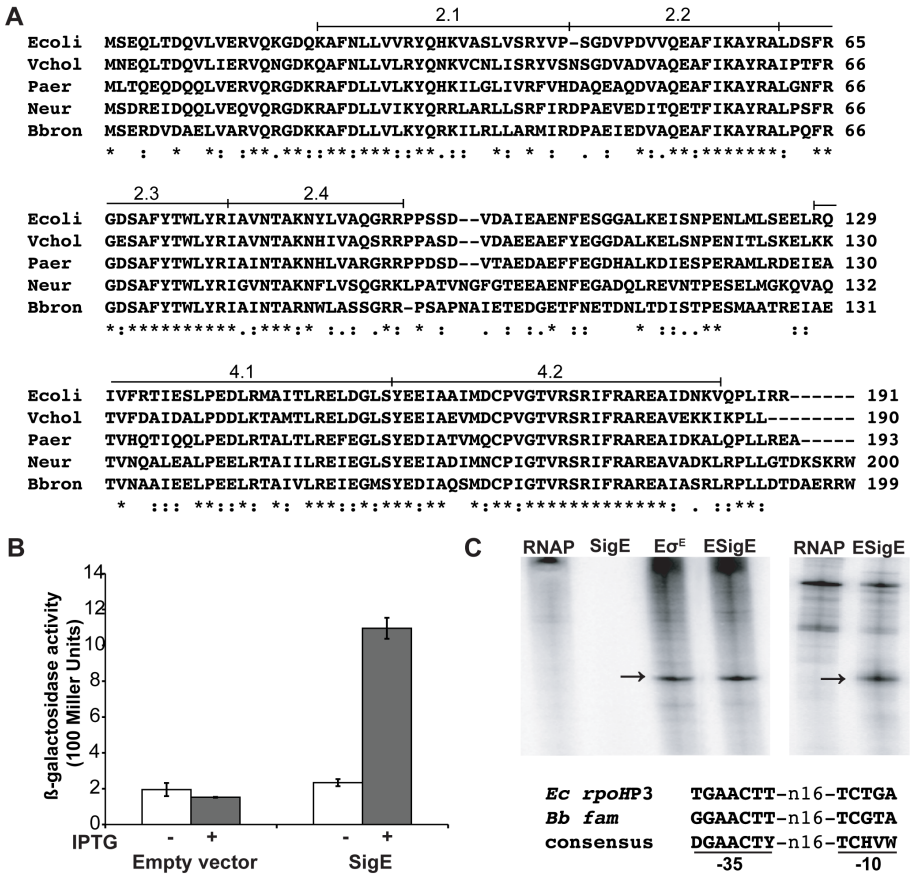


Figure 2-1: *B. bronchiseptica* SigE is a functional sigma factor. (A) Amino acid sequence alignment of RpoE- like sigma factors from *Escherichia coli* (Ecoli)*, Vibrio cholerae* (Vchol), *Pseudomonas aeruginosa* (Paer), *Nitrosomonas europaea* (Neur) and *B. bronchiseptica* (Bbron) using ClustalW2 (EMBL-EBI). Asterisks indicate identity, and one or two dots indicate semi-conservation and conservation, respectively, between amino acid residues. Conserved sigma factor regions 2.1-2.4 and 4.1-4.2 [158] are indicated above the alignment. (B) β-galactosidase activity from the *E. coli rpoH*P3-*lacZ* reporter increases when *B. bronchiseptica sigE* expression is induced from plasmid pSEB006 in strain SEA5005 (grey bars) by the addition of IPTG. No increase is seen upon IPTG addition to

the control strain, SEA008 (white bars), containing the empty vector. The observed difference between the amount of β-galactosidase activity between the two strains in the presence of IPTG is statistically significant (P value <0.001). (C) In vitro transcription from a supercoiled plasmid template containing the *E. coli* σ^E-dependent *rpoH*P3 promoter with *E. coli* core RNA polymerase (RNAP), SigE alone, Eσ^E, and ESigE (left panel). In vitro transcription from a linear template containing the promoter region of *B. bronchiseptica fam*, with *E. coli* core RNAP alone (RNAP), or ESigE (right panel). Arrows indicate transcripts from the *rpoH*P3 and *fam* promoters. Below, an alignment of the *E. coli rpoH*P3 and *B. bronchiseptica fam* promoter sequence with the consensus promoter from Staron, et al. (D=A, G, or T; Y=C or T; H=A, C, or T; V=A, G, or C; W=A or T) [40].

SigE contributes to the *B. bronchiseptica* stress response

To investigate the role of SigE in *B. bronchiseptica*, an in-frame deletion of the *sigE* gene was constructed in RB50 (RB50Δ*sigE*) that removed 176 out of 200 codons of the gene, leaving 22 and 2 codons at the 5′ and 3′ ends of the gene, respectively. The deletion was confirmed by PCR and Southern blotting methods (data not shown). σ^E orthologs are essential in some bacteria, including *E. coli* and *Y. enterocolitica* [61, 62], yet are not required for viability in many other species, such as *S. enterica* serovar Typhimurium, *P. aeruginosa*, and *Burkholderia pseudomallei* [50, 97, 108]. Deletions of *B. bronchiseptica sigE* were readily obtained, suggesting that it falls in the latter class, and is not essential for viability. Furthermore, RB50Δ*sigE* grew at a rate similar to that of RB50 under standard growth conditions (37 °C in Stainer-Scholte broth) (Fig. 2-2A).

To investigate whether SigE mediates a cell envelope stress response in *B. bronchiseptica*, we used disk diffusion assays to compare the sensitivity of RB50 and RB50Δ*sigE* to several chemicals that compromise cell envelope integrity and a series of antibiotics that block different steps in peptidoglycan synthesis. The *sigE* mutant was more sensitive than the wild-type strain to the detergent SDS in combination with EDTA (Fig. 2B). The *sigE* mutant was also more sensitive than wild-type RB50 to the antibiotics

mecillinam and ampicillin (Fig. 2-2B), whereas sensitivity to meropenem, aztreonam, and imipenem was not affected (data not shown). Unlike σ^E orthologs in other bacteria, SigE was not required for resistance to the cationic antimicrobial peptide polymyxin B, which targets bacterial membranes, and to osmotic stress (Fig. 2-2B and data not shown) [50, 97, 99, 160]. RB50Δ*sigE* and RB50 were also equally sensitive to antibiotics that inhibit cytoplasmic processes such as translation (chloramphenicol, erythromycin, kanamycin, tetracycline), transcription (rifampicin), and cytoplasmic enzymes DNA gyrase (nalidixic acid), and dihydrofolate reductase (trimethoprim) (data not shown). This lack of sensitivity to multiple antibiotics suggests that the *sigE* mutation does not lead to an overall increase in the permeability of the outer membrane, which would allow more of the antibiotic to enter the cell. These results show that SigE is required for survival in response to particular types of damage to the cell envelope of *B. bronchiseptica*.

We next asked if *sigE* is important for survival following a shift to high temperature, which perturbs both the cell envelope and cytoplasm. RB50 and RB50Δ*sigE* were grown at 37 °C to an OD_{600} of 0.4, then shifted to 50 °C, a lethal temperature for *B. bronchiseptica*. Cell viability, assessed by CFU/ml, was measured after the shift to 50 °C. Survival of the RB50Δ*sigE* strain was lower than that of RB50 (Fig. 2-2C). In attempting to complement this phenotype, we found that plasmid-encoded *sigE* did not restore survival during heat shock (data not shown), although it did complement other phenotypes, as described below. Similar variability in complementation of a σ^E mutant by a plasmid-encoded *rpoE* gene has been seen in other bacteria [49, 50, 65, 123]. Work from *Burkholderia cenocepacia*, a species closely related to the bordetellae, showed that expressing σ^E from a plasmid actually increased sensitivity to heat stress [50]. In *S.*

enterica serovar Typhimurium, an *rpoE* mutant was sensitive to paraquat and did not survive in stationary phase under anaerobic conditions. Expression of *rpoE* from a plasmid partially complemented the former phenotype, but not the latter [65]. Because the anti-sigma factor that regulates σ^E activity was not included in any of these instances, it is likely that proper regulation of SigE activity is required for optimal response to particular stresses, not merely excess SigE activity.

Another aspect of heat shock response is thermotolerance. When bacteria are exposed to an elevated but nonlethal temperature, heat shock responses are induced, resulting in increased production of heat shock proteins, chaperones and proteases that refold or degrade unfolded proteins [161]. Consequently, the cells are preloaded with protective factors and exhibit increased survival following a subsequent shift to a lethal temperature [161]. To investigate the role of SigE in this phenomenon, RB50 and RB50Δ*sigE* were grown to an OD_{600} of 0.1 at 37 °C, shifted to 40 °C for 90 min, then shifted to 50 °C. RB50 cultures incubated at 40 °C before 50 °C survived better at all time points than those directly shifted from 37 °C to 50 °C. For example, 54% of the RB50 cells pre-adapted at 40 °C survived two hours after the shift to 50 °C (Fig. 2-2C), compared to 0.1% survival for those shifted directly from 37 °C to 50 °C (Fig. 2-2C). RB50Δ*sigE* pre-adapted at 40 °C also survived better at 50 °C than when directly shifted from 37 °C to 50 °C. However, only 38% of the RB50Δ*sigE* cells survived after one hour (compared to 76% of the wild-type RB50), and 5% survived after two hours at 50 °C (Fig. 2-2C). These results demonstrate that *B. bronchiseptica* exhibits a typical thermotolerance response and that SigE contributes to this response.

Both ethanol and heat shock lead to protein unfolding and often elicit similar stress responses [98]. To test the role of *sigE* in response to ethanol stress, RB50 and RB50Δ*sigE* were subcultured from mid-exponential-phase cultures into fresh Stainer-Scholte broth with or without 3% ethanol. Both strains grew similarly in medium without ethanol, as noted above. RB50 grew significantly slower in medium containing 3% ethanol than in medium without ethanol (compare the growth curve for RB50 in Fig. 2-2D with that in Fig. 2-2A), but eventually reached a cell density only slightly below that of cultures grown in medium without ethanol. In contrast, the cell density of RB50Δ*sigE* grown in 3% ethanol never surpassed an OD_{600} of around 0.1, even after 24 hours. Expression of plasmid-encoded *sigE* in RB50Δ*sigE* partially complemented this phenotype, restoring growth in 3% ethanol to nearly that of RB50 (Fig. 2-2D), indicating that *sigE* is required for survival during ethanol stress.

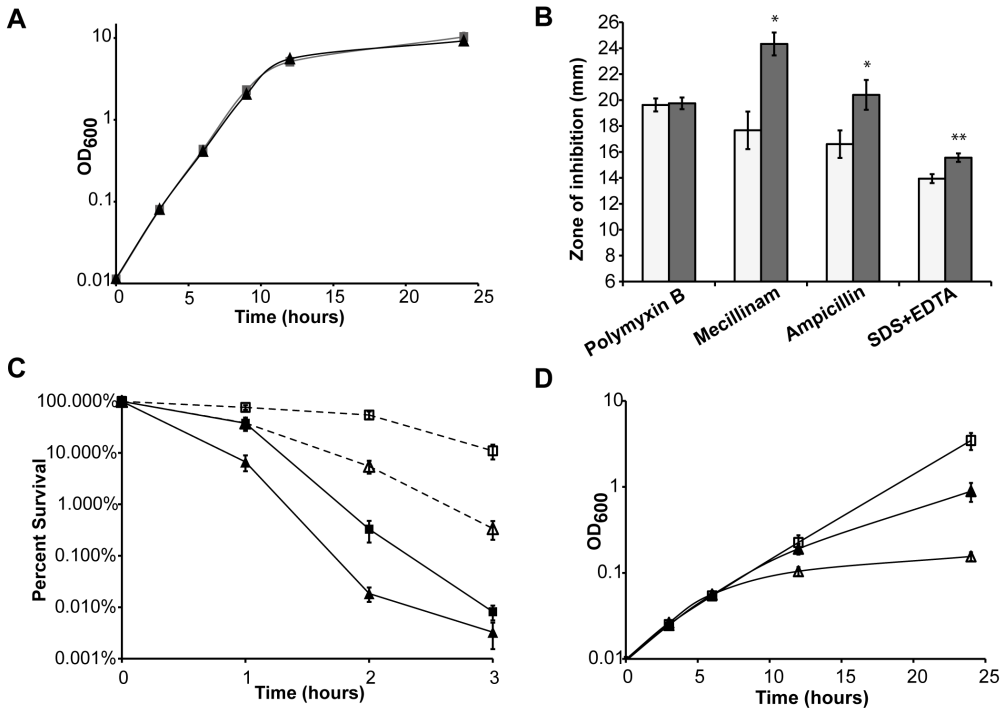


Figure 2-2: Role of SigE in response to environmental stresses. (A) RB50 (squares) and RB50Δ*sigE* (triangles) grow similarly at 37 °C in Stainer-Scholte broth. (B) RB50Δ*sigE* (white bars) is more sensitive than RB50 (grey bars) to treatment with 100 μg mecillinam, 10 μg ampicillin, or 750 μg SDS and 2.9 μg EDTA, but is similarly sensitive to treatment with 300 iu polymyxin B in disk diffusion assays. The average diameters of the zones of inhibition ± SE from at least three independent experiments are shown. The disk diameter was 6 mm. The observed differences between the zones of inhibition for RB50 and the *sigE* mutant are statistically significant for mecillinam, ampicillin, and SDS-EDTA (* indicates a P-value of < 0.05; ** indicates a P-value < 0.01). (C) RB50Δ*sigE* (triangles) is more sensitive than RB50 (squares) to heat shock (solid line, filled symbols) caused by shifting cultures from 37 °C to 50 °C. RB50Δ*sigE* also exhibits reduced thermotolerance (dashed line, open symbols), surviving less well than RB50 when adapted first to 40 ºC before a shift to 50 ºC. The mean percent survival ± SE of fifteen independent experiments for each strain is shown. (D) RB50Δ*sigE* containing the empty cloning vector pEV (open triangles) is more sensitive to treatment with 3% ethanol than RB50 pEV (squares). Expression of plasmid-encoded SigE (RB50Δ*sigE* pSigE) restores growth in 3% ethanol (filled triangles) to near wild-type levels at the 6 and 12 hour time points and partially restores growth at the 24 hour time point. The mean OD_{600} ± SE of at least four independent experiments is shown for each strain.

σ^E homologues have also been found to play a role during oxidative stress in *S. enterica* serovar Typhimurium and *Burkholderia pseudomallei* [65, 123]. However, in disk diffusion assays, SigE was not required for survival in the presence of hydrogen peroxide or paraquat, two inducers of oxidative stress (data not shown). Either SigE is not involved in combating oxidative stress in *B. bronchiseptica,* or other oxidative-stress responsive pathways compensate for SigE when it is deleted.

Growth in the murine respiratory tract is not affected by the lack of *sigE*.

B. bronchiseptica RB50 colonizes the respiratory tract of immunocompetent mice, causing an asymptomatic infection that is eventually cleared by the immune system. To determine whether *B. bronchiseptica* SigE contributes to colonization and persistence in the respiratory tract, groups of C57BL/6 mice were inoculated with RB50 or RB50Δ*sigE*, and colonization was measured in the nasal cavity, trachea, and lung on days 0, 3, 7, 14, 28 and 63 post-inoculation. Both wild-type and *sigE*-deficient RB50 colonized the nasal cavity at comparable levels, peaking on day 3 post-inoculation, and stabilizing at about 10^{4-5} CFU by 2 weeks post-inoculation (Fig. 2-3). Both strains also showed similar colonization kinetics in the lower respiratory tract of C57BL/6 mice, peaking in numbers on days 3 and 7 post-inoculation in the trachea and lungs, respectively, and declining thereafter, with complete clearance in both organs by day 63 post- inoculation (Fig. 2-3). These data indicate that *B. bronchiseptica* SigE is not required for colonization or persistence in the murine respiratory tract.

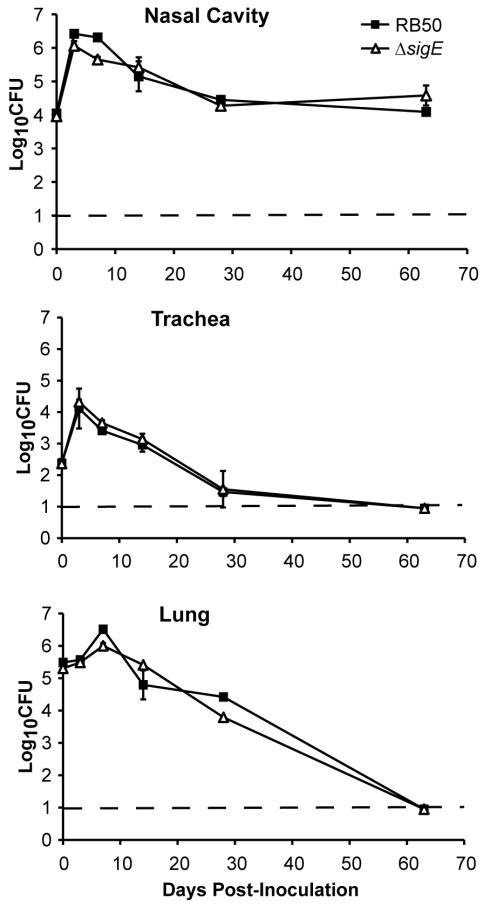


Figure 2-3: Colonization of the respiratory tract of C57BL/6 mice by RB50 and RB50Δ*sigE*. Groups of three 4-6 week-old C57BL/6 mice were inoculated with 5×10^5 CFU of RB50 (filled squares) and RB50ΔsigE (open triangles). Groups of three mice were sacrificed at each time point. The bacterial load in the indicated organ is expressed as $\log_{10}$ CFU ± SE. The dashed line indicates the limit of detection. The experiment was performed twice with similar results and a representative dataset is shown.

SigE contributes to systemic spread and lethal *B. bronchiseptica* infection in mice lacking B cells and T cells, but not in mice lacking TLR4 or TNF-α.

B. bronchiseptica has been observed to cause a range of disease including bronchitis, lethal pneumonia, and even systemic infection [136, 156]. Mice with defined

immune deficiencies are particularly susceptible to different forms of disease [149, 150, 162], facilitating assessment of the roles of specific bacterial factors/functions in interactions with different aspects of the host immune response.

Mice lacking key components of innate immunity, either TLR4 or TNF-α, were challenged with RB50 or RB50Δ*sigE* and signs of severe disease were monitored. Consistent with published studies, TLR4^{def} and TNF-α$^{-/-}$ mice inoculated with 10^5 CFU of RB50 quickly developed signs of lethal bordetellosis such as ruffled fur, hunched posture, decreased activity, and difficulty breathing, and succumbed 2 to 5 days post-inoculation [150, 151]. Mice challenged with RB50Δ*sigE* also showed similar signs of disease and time to death (Fig. 2-4). In a separate experiment, TLR4^{def} mice and TNF-α$^{-/-}$ mice infected with RB50 or RB50Δ*sigE* that were still alive by day 3 post-inoculation were dissected for bacterial enumeration in the respiratory as well as systemic organs. Both wild-type and *sigE*-deficient RB50 colonized the lungs of TLR4^{def} mice at 10^{7-8} CFU, which was almost 1000-fold higher than in the lungs of TLR4^{suf} mice. Moreover, both strains colonized the systemic organs in TLR4^{def}, but not TLR4^{suf} mice (Fig. 2-4). Both strains also grew to higher numbers in the lungs of TNF-α$^{-/-}$ mice than in the lungs of C57BL/6 mice and were recovered from systemic organs only in TNF-α$^{-/-}$ mice (Fig. 2-4). These data indicate that SigE is not required for *B. bronchiseptica* to cause lethal infection and colonize systemic organs in mice lacking TLR4 or TNF-α.

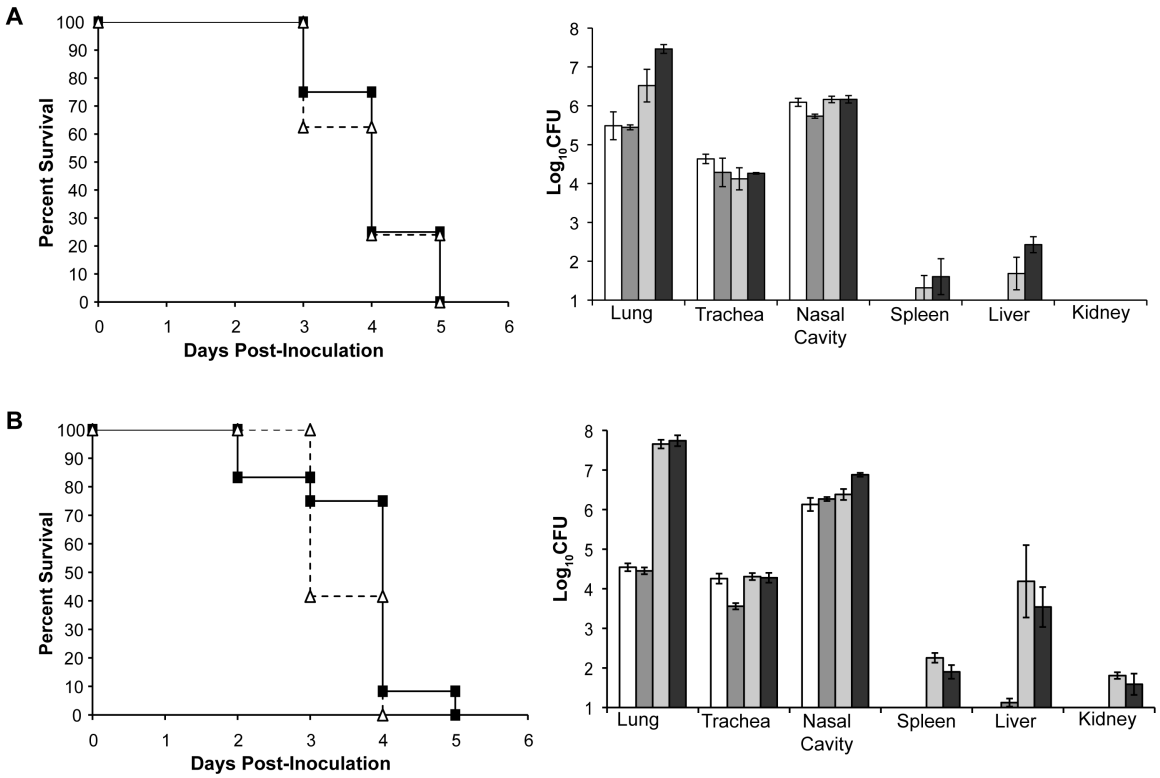


Figure 2-4: SigE is not required for survival or systemic colonization in TNF-α$^{-/-}$ or TLR4^{def} mice. (A) Groups of TNF-α$^{-/-}$ mice (n=8) were inoculated with 1x10^5 CFU of RB50 or RB50Δ*sigE*. *Left*, mice inoculated with RB50 (solid lines) and RB50Δ*sigE* (dashed lines) show similar time to death. *Right*, on day 3-post-inoculation, bacterial numbers were enumerated from the indicated organs in C57BL/6 mice inoculated with RB50 (white bars) or RB50Δ*sigE* (medium grey bars), and in TNF-α$^{-/-}$ mice inoculated with RB50 (light grey bars) or RB50Δ*sigE* (dark grey bars). (B) Groups of TLR4^{def} mice (n=12) were inoculated with 1x10^5 CFU of RB50 or RB50Δ*sigE*. *Left*, mice inoculated with RB50 (solid lines) and RB50Δ*sigE* (dashed lines) show similar time to death. *Right*, on day 3-post-inoculation, bacterial numbers were enumerated from the indicated organs in TLR4^{suf} mice inoculated with RB50 (white bars) or RB50Δ*sigE* (medium grey bars), and in TLR4^{def} mice inoculated with RB50 (light grey bars) or RB50Δ*sigE* (dark grey bars). The bacterial load is expressed as $\log_{10}$ CFU ± SE. Limit of detection is indicated as the bottom of the y-axis.

B and T cell-deficient Rag1$^{-/-}$ mice succumb to *B. bronchiseptica* infection, and death is associated with systemic spread of the infection [148]. To assess the role of SigE in the systemic spread of *B. bronchiseptica* in hosts deficient in adaptive immunity, groups of Rag1$^{-/-}$ mice were inoculated with 5×10^5 CFU of RB50 or RB50Δ*sigE*. Rag1$^{-/-}$ mice

inoculated with RB50 showed symptoms of lethal bordetellosis on day 13 post-inoculation and succumbed between days 14-35 post-inoculation (Fig. 2-5A). However, Rag1$^{-/-}$ mice inoculated with RB50Δ*sigE* survived without any overt signs of disease and were euthanized on day 122 post-inoculation. The nasal cavity, trachea, lung, spleen, liver and kidneys of these mice were excised to enumerate bacterial loads. Although 10^{5-7} CFU of RB50Δ*sigE* were recovered from the respiratory tract, this strain failed to colonize the spleen or kidney, and only 300 CFU were recovered from the liver (Fig 2-5B, dark gray bars). In a separate experiment, RB50 and RB50Δ*sigE*-inoculated Rag1$^{-/-}$ mice were sacrificed on day 28 post-inoculation, when some of the RB50-challenged mice were still alive. The bacterial loads of RB50 and RB50Δ*sigE* in the respiratory tract on day 28 post-inoculation were similar, about 10^{5-7} CFU. At this time, 10^{4-6} CFU of RB50 were recovered from liver, spleen and kidney (Fig. 2-5B, white bars). RB50Δ*sigE*, however, failed to colonize the spleen, kidney or liver (Fig. 2-5B, light gray bars). These results demonstrate that SigE is required for lethal systemic infection by *B. bronchiseptica* in Rag1$^{-/-}$ mice.

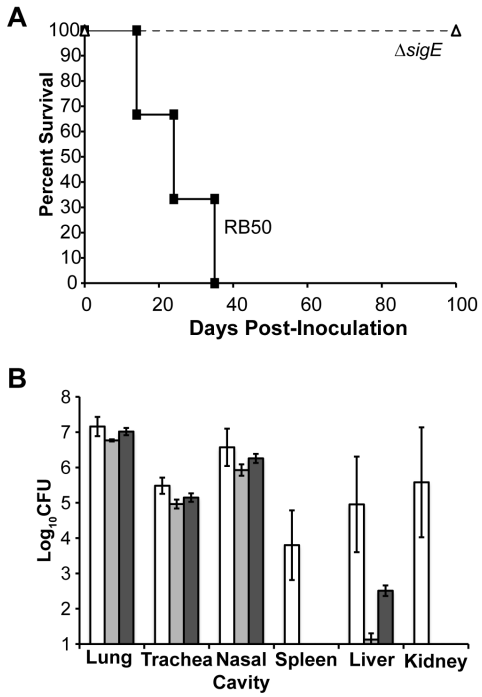


Figure 2-5: Survival and systemic colonization of Rag1$^{-/-}$ mice following infection with RB50 and RB50Δ*sigE*. (A) Groups of Rag1$^{-/-}$ (n=6) mice were inoculated with 5×10^5 CFU of RB50 (solid line with filled squares) or RB50Δ*sigE* (dashed line with open triangles) and monitored for survival. (B) Groups of four Rag1$^{-/-}$ mice were inoculated with 5×10^5 CFU of RB50 (white bars) or RB50Δ*sigE* (light grey bars) and dissected on day 28 post-inoculation for bacterial enumeration in the indicated organs. In a separate experiment, Rag1$^{-/-}$ mice inoculated with RB50Δ*sigE* were euthanized for bacterial numbers in the indicated organs on day 122 post-inoculation (dark grey bars). The bacterial load is expressed as $\log_{10}$ CFU ± SE. Limit of detection is indicated as the bottom of the y-axis.

The failure of RB50Δ*sigE* to colonize distal organs of Rag1$^{-/-}$ mice suggests that this mutant is defective in either getting into or survival in the bloodstream and/or systemic organs. Since we have previously shown that the ability of *B. bronchiseptica* to cause systemic infection is dependent on its resistance to complement-mediated killing [148], we hypothesized that RB50Δ*sigE* might be susceptible to complement-mediated killing. To test this, 500 CFU of RB50, RB50Δ*sigE*, or RB50Δ*wbm*, a strain lacking O-antigen, which is known to be susceptible to complement [148], were incubated at 37 °C for one hour in

PBS with 20% complement-active or complement-inactive serum from naïve mice. The survival of RB50Δ*sigE* and RB50 was not affected by the presence of either serum. In contrast, the RB50Δ*wbm* strain was almost completely killed by complement-active, but not complement-inactive serum (Fig. 2-6). The observation that RB50Δ*sigE* survived in the presence of serum without *B. bronchiseptica*-specific antibodies indicates that the defect in causing systemic infection in mice lacking B and T cells is not due to failure to survive the antimicrobial components in serum, including complement.

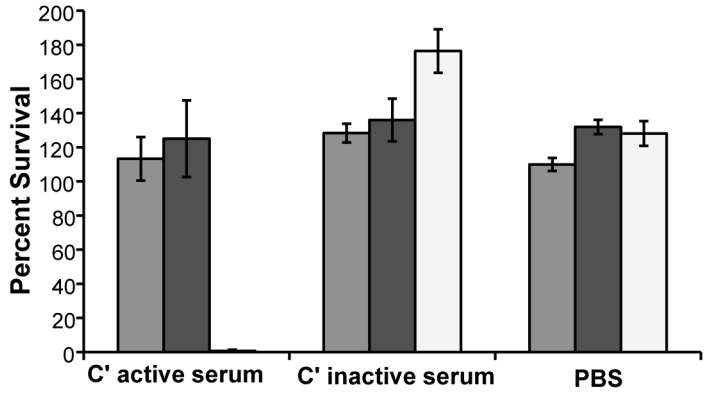


Figure 2-6: RB50Δ*sigE* is not more susceptible to complement-mediated killing. RB50 (medium grey bars) and RB50Δ*sigE* (dark grey bars) survive similarly during incubation with complement-active serum, complement-inactive serum, or PBS alone. RB50Δ*wbm* (white bars) is sensitive to complement-active serum.

SigE contributes to cytotoxicity to macrophages.

We further tested whether RB50Δ*sigE* interacts differently than RB50 with another major bactericidal component in the bloodstream, phagocytes. *B. bronchiseptica* is cytotoxic to macrophages, and this toxicity has been attributed to the activities of the type three secretion system (TTSS) and adenylate cyclase toxin [163]. To test the role of SigE in macrophage cytotoxicity, RAW264.7 murine macrophages were incubated for 4 hours at an MOI of 10 with RB50, RB50 lacking *sigE*, RB50 lacking a functional TTSS (WD3), or

RB50 lacking both TTSS and adenylate cyclase toxin (AVS). In this experiment, both the RB50 and RB50Δ*sigE* strains contained the empty cloning vector pEV to allow direct comparisons with the complemented strain, RB50Δ*sigE* pSigE. Cytotoxicity was determined by measuring lactate dehydrogenase (LDH) release from the treated macrophages. WD3 and AVS caused little cytotoxicity, similar to treatment with medium alone. RB50Δ*sigE* pEV caused approximately 50% less cytotoxicity than wild-type RB50 pEV (Fig. 2-7). This defect in cytotoxicity was complemented by supplying the *sigE* gene on the plasmid pSigE (Fig. 2-7), indicating that loss of *sigE* negatively impacts the ability of RB50 to kill macrophages.

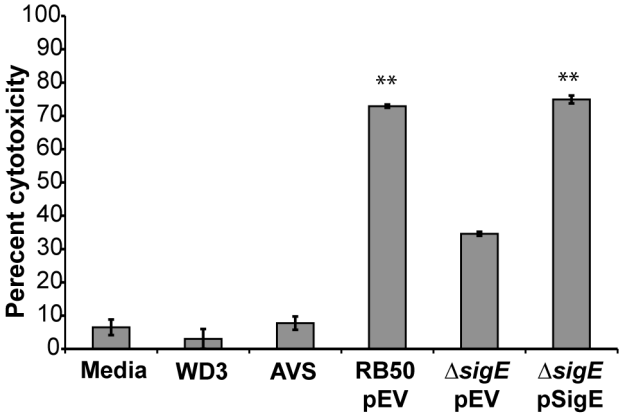


Figure 2-7: RB50Δ*sigE* is less cytotoxic to macrophages. RAW 264.7 cells were incubated at an MOI of 10 with medium containing RB50 pEV, RB50Δ*sigE* pEV, RB50Δ*sigE* pSigE, TTSS-deficient RB50 strain WD3, ACT and TTSS-deficient RB50 strain AVS, or medium alone for 4 hours in the presence of 1mM IPTG to induce expression of *sigE* from the pLac promoter of pSigE. The average percent cytotoxicity of four wells in four separate experiments as measured by (LDH release from a well/LDH release from the positive control well) x100 ± SE is shown. The differences in percent cytotoxicity between RB50Δ*sigE* pEV and either RB50 pEV or RB50Δ*sigE* pSigE are statistically significant (** indicates P value < 0.01), while the cytotoxicities of RB50 pEV and RB50Δ*sigE* pSigE are not significantly different.

RB50Δ*sigE* is more efficiently phagocytosed and killed by PMNs.

To test if RB50Δ*sigE* is more susceptible to another bactericidal mechanism, phagocytosis by peripheral blood polymorphonuclear leukocytes (PMNs), RB50 and RB50Δ*sigE* were incubated with freshly isolated human PMNs and attachment to, phagocytosis by, and killing by these cells were measured. PMNs bound RB50Δ*sigE* more efficiently than RB50 (Fig. 2-8A), and significantly more RB50Δ*sigE* than RB50 were phagocytosed by PMNs (Fig. 2-8B). However, the number of viable intracellular RB50Δ*sigE* was ~50% of the numbers of viable RB50 (Fig. 2-8C, left panel). When differences in attachment and phagocytosis were taken into consideration, significantly more internalized RB50Δ*sigE* were killed compared to RB50 (Fig. 2-8C, right panel). Together, these data indicate that SigE contributes to *B. bronchiseptica* resistance to phagocytosis and killing by PMNs.

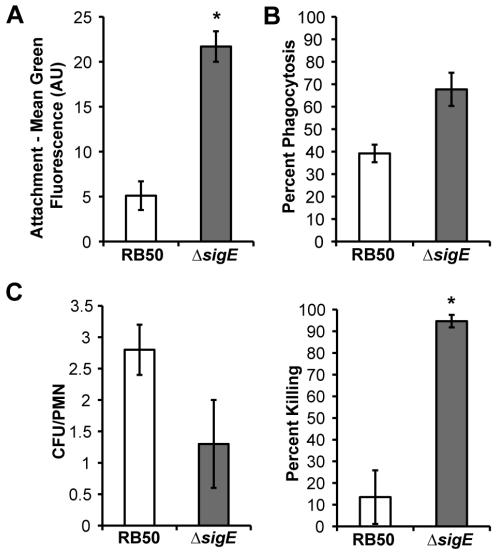


Figure 2-8: RB50Δ*sigE* is more efficiently phagocytosed and killed by PMNs. (A) GFP-expressing RB50 (white bars) and RB50Δ*sigE* (grey bars) were incubated with freshly isolated human peripheral blood PMNs for 20 min at an MOI of 50. Attachment levels were measured as mean intensities ± SE of green fluorescence associated with

PMNs. (B) Cell surface-bound bacteria were detected by incubation with RPE-labeled goat F(ab')2 fragments of anti-mouse IgG, after incubation with immune serum. Mean phagocytosis levels ± SE were calculated from the decrease in red fluorescence of GFP-positive cells incubated for an additional 30 min at 37 °C allowing for internalization (RPE2, 50 min total incubation time) compared to that of cells incubated for only 20 min (RPE1). Percent phagocytosis is (1-RPE2/RPE1) × 100%. (C) To determine killing of bacteria by PMNs, cells incubated with bacteria for 50 min were treated with antibiotics to kill extracellular bacteria. Viable bacteria per PMN (left) and percent killing of internalized bacteria (right) were expressed as mean ± SE. AU indicates arbitrary units; * indicates a P-value of < 0.05.

Discussion

The BvgAS system of the bordetellae plays a central role in regulating gene expression during pathogenesis [27, 143]. However, other regulators may be required during the infectious disease cycle, as *Bordetella* species have a large number of sensory systems encoded in their genomes [139]. In this study, we focused on cell envelope sensing systems and investigated the alternative sigma factor, SigE. We found that SigE of *B. bronchiseptica* does indeed mediate a protective cell envelope stress response and is required for systemic infection, suggesting that the role of SigE is to combat stresses to this compartment imposed by the immune system within a host and by harsh conditions in the environment outside a host. This work is the first demonstration of a cell envelope sensing system in the bordetellae. The σ^E system has been explored in the most depth in enteric pathogens belonging to the γ-proteobacteria [39, 47, 66]. The bordetellae, members of the β-proteobacteria, encounter distinctly different environments in the respiratory tract and therefore provide an excellent model to study how the SigE system has been adapted throughout evolution to serve the needs of diverse bacterial pathogens.

The entire *sigE* locus (BB3754-BB3749) is > 99% identical among the classical bordetellae, suggesting a conserved role for this system in the human pathogens *B. pertussis* and *B. parapertussis*. However, the lifestyles and, therefore, conditions encountered differ amongst the three bordetellae strains. *B. bronchiseptica* can live outside the host and primarily infects non-human mammals, although it can infect immunocompromised humans [137, 156]. In contrast, *B. pertussis* and *B. parapertussis* primarily infect humans and are directly transmitted between hosts [164, 165]. As we learn more about the role of SigE in the bordetellae, it will be of interest to determine whether

stresses that induce the SigE system and the SigE regulon members are as highly conserved as the *sigE* locus itself among the bordetellae.

Our results define roles for SigE in *B. bronchiseptica* that are only partially overlapping with those for σ^E in other pathogens. SigE was important for survival of *B. bronchiseptica* in the face of both global stresses to the cell envelope caused by heat shock, exposure to ethanol and detergent, and specific stresses caused by several beta-lactam antibiotics (Fig. 2-2). SigE was not required for survival during osmotic stress, as found for *B. cenocepacia* and *S. enterica* serovar Typhimurium, [50, 97]. SigE was also not required for response to oxidative stress or the antimicrobial peptide polymyxin B, unlike the *S. enterica* serovar Typhimurium σ^E ortholog [65, 97]. The variations among bacteria in their use of σ^E systems likely reflect both differences in stresses encountered in environmental reservoirs and during infection, as well as differences in the arrays of additional cellular stress responses possessed by each species. These other responses can act along with or in place of σ^E. The presence of other stress responses may be particularly pertinent to *B. bronchiseptica*, because its genome is predicted to encode seven ECF sigma factors in addition to SigE [40] that may have complimentary and redundant functions with SigE. Future studies defining conditions that activate other ECF sigma factors and their roles in *B. bronchiseptica* pathogenesis will provide a more comprehensive understanding of how *B. bronchiseptica* copes with extracytoplasmic stress.

Stress response systems, like the σ^E system, rapidly induce the expression of specialized sets of genes. These systems are often tightly regulated and expressed only when needed, because inappropriate expression of their regulons can interfere with other important cellular functions [6, 134, 166]. We found that SigE was not required for

colonization and persistence of RB50 within the respiratory tract of an immunocompetent host (Fig. 2-3), the primary niche of *B. bronchiseptica*. This result suggests that this pathogen does not encounter stresses in the respiratory tract that require a response by the SigE system. However, *B. bronchiseptica* encounters new challenges during systemic infection in $Rag1^{-/-}$ mice lacking B and T cells. The bloodstream is under greater immune surveillance and has a different arsenal of antimicrobial factors to attack invaders than the respiratory tract. The defect of RB50Δ*sigE* in systemic infection of $Rag1^{-/-}$ mice, therefore, reveals a specific function for SigE in response to an unknown stress particular to these systemic sites.

The inability of RB50Δ*sigE* to cause systemic lethal infections in $Rag1^{-/-}$ mice (Fig. 2-5) could be due to failure to enter or survive in the bloodstream and/or systemic organs of these mice. Since the mutation does not affect survival during incubation with serum in vitro, it is unlikely that the *sigE*-deficient strain is more susceptible to complement or other antimicrobial components in serum. The defect in infection of $Rag1^{-/-}$ mice may then be related to altered interactions of the mutant strain with phagocytic cells in the bloodstream. RB50Δ*sigE* is more susceptible to peripheral blood PMNs than RB50 (Fig. 2-8), and is also less cytotoxic to macrophages than RB50 (Fig. 2-7). Either or both of these defects could explain the failure to recover RB50Δ*sigE* from systemic organs of mice lacking adaptive immune responses and the decreased virulence in these mice.

Why does the RB50Δ*sigE* mutant cause systemic, lethal infection in $TLR4^{def}$ and TNF-$\alpha^{-/-}$ mice, but not $Rag1^{-/-}$ mice? The lower cytotoxicity of the *sigE* mutant and its increased sensitivity to phagocytic killing does not affect its virulence in mice lacking innate immune functions. This could be because bacterial numbers within the respiratory

tract of TLR4^{def} or TNF-$\alpha^{-/-}$ mice are nearly an order of magnitude higher than in the lungs of Rag1$^{-/-}$ mice. As such, the large number of bacteria in TLR4^{def} or TNF-$\alpha^{-/-}$ mice may overwhelm limiting host antimicrobial defense mechanisms that can contain the lower bacterial numbers in the lungs of Rag1$^{-/-}$ mice. Alternatively, although the cytotoxicity of the *sigE* mutant is reduced, it may still be sufficient to establish lethal infections in the absence of TLR4 or TNF-α. Thus TLR4- and TNF-α-dependent functions, such as efficient phagocytosis and killing, appear to be sufficient to prevent lethal systemic infection by RB50Δ*sigE* in Rag1$^{-/-}$ mice. Although the exact role remains to be elucidated, our results clearly indicate that SigE is required for systemic infection of mice lacking B and T cells.

Although the *B. bronchiseptica* strain RB50 causes asymptomatic infections in immunocompetent mice, other strains of *B. bronchiseptica* can cause a wide range of disease severity in other hosts [136, 138, 156]. In particular subsets of immunocompromised humans, such as those infected with HIV, severe systemic *B. bronchiseptica* infections have been observed [137]. This fact, along with the high degree of sequence conservation for the *sigE* locus in *B. pertussis* and *B. parapertussis*, highlights the importance of understanding the stressors that activate SigE and how the SigE system responds to them during systemic infection.

Materials and Methods

Strains and Media

A complete list of strains used in this study can be found in Table 1. *B. bronchiseptica* strains are derivatives of the previously described *B. bronchiseptica* strain RB50 [167]. *B. bronchiseptica* was maintained on Bordet-Gengou (BG) agar (Difco) containing 10% defibrinated sheep blood (Hema Resources) and 20 μg/ml streptomycin. In liquid culture, *B. bronchiseptica* was grown in Stainer-Scholte broth [168] with aeration. Chloramphenicol was used at 20 μg/ml and IPTG at 1 mM where noted. The RB50Δ*sigE* mutant was constructed as described below. *E. coli* strains used to measure SigE activity are derivatives of MG1655 that carry the σ^E-dependent *rpoH*P3::*lacZ* reporter (strain SEA001 [68]). *E. coli* strain BL21(DE3) pLysS was used to express constructs for protein purification. *E. coli* were grown in LB broth in a gyratory water bath with aeration. Ampicillin was used at 100 μg/ml, tetracycline at 20 μg/ml, and kanamycin at 15 μg/ml as needed for experiments with *E. coli*.

Plasmid construction

All plasmids used in this study are listed in Table 1 and oligonucleotide sequences are given in Table 2. Plasmid pSEB006 was constructed to express *sigE* in *E. coli*. The *sigE* gene was amplified from RB50 genomic DNA with the primers SigEF and SigER and cloned into the expression vector pTrc99a under the control of the IPTG-inducible trc promoter. To facilitate purification of SigE, the plasmid pXQZ001 was constructed by amplifying the *sigE* gene from RB50 genomic DNA using the primers HisSigEF and HisSigER. The resulting PCR product was cloned into the T7 expression vector pET-15b (Novagen), which adds a 6X-His tag to the N-terminus of recombinant proteins. To

express *sigE* in *B. bronchispetica*, *sigE* was amplified from RB50 genomic DNA using primers 72SigEF and 72SigER and ligated into the XbaI and XhoI sites downstream of the pLac promoter in pEV to create pSigE. The expression vector pEV was constructed from the broad host range vector pJS72 by replacing the spectinomycin resistance gene with the *cat* gene encoding chloramphenicol resistance amplified from pKD3 [169] using primers 72ChlorF and 72ChlorR. The exchange of drug markers was necessary because RB50 is naturally resistant to spectinomycin. pEV and pSigE were moved into RB50 and RB50Δ*sigE* through tri-parental mating on BG agar with $MgCl_2$. Transconjugants were selected on BG containing 60 μg/ml streptomycin and 20 μg/ml chloramphenicol. Plasmid pCW505 (kindly supplied by Dr. Alison Weiss, Cincinnati, Ohio), which induces cytoplasmic expression of GFP without affecting growth or antigen expression, was used to visualize RB50 and RB50Δ*sigE* in the phagocytosis assays described below [170].

Table 2-1: Strains and plasmids

	Strain name	Genotype	Source, Reference
E. coli	SEA001	MG1655 Φλ*rpoH*P3::*lacZ* Δ*lac*X74	[82]
	SEA5036	BL21(DE3) *slyD*::kan pLysS pPER76	[92]
	XQZ001	BL21(DE3) *slyD*::kan pLysS pXQZ001	This work
	SEA4114	CAG43113 *rpoE*::kan *nadB*::Tn*10*	[71]
	SEA008	SEA001 pTrc99a	[71]
	SEA5005	SEA001 pSEB006	This work
	XQZ003	DH5α pXQZ0003	This work
	SS1827	DH5α pSS1827	[171]
B. bronchiseptica	RB50	RB50	[167]
	SEA5516	RB50Δ*sigE*	This work
	MER001	RB50 pCW505	This work
	MER002	RB50Δ*sigE* pCW505	This work
	SEA5518	RB50 pEV	This work
	SEA5520	RB50Δ*sigE* pEV	This work
	SEA5526	RB50 pSigE	This work
	SEA5530	RB50Δ*sigE* pSigE	This work
	RB50Δwbm	RB50Δ*wbmBwbmCwbmDwbmE*	[172]

B. bronchiseptica (cont.)	WD3	RB50Δ*bscN*	[163]
	AVS	RB50Δ*cyaA*Δ*bscN*	[149, 173]
	Plasmid name	**Description**	**Source, Reference**
	pTrc99a	Vector, pBR322 ori, Ap^R	Pharmacia
	pSEB006	*sigE* in pTrc99a	This work
	pSEB015	isolated *rpoH*P3 promoter in pRLG770, Ap^R	[92]
	pPER76	*rpoE* in T7 expression vector pET15b, Kan^R	[70]
	pXQZ001	*sigE* in T7 expression vector pET15b, Kan^R	This work
	pXQZ002	Δ*sigE* in TOPO-TA vector	This work
	pSS1827	helper plasmid competent for mating, Ap^R	[171]
	pSS3962	*Bordetella*-specific allelic exchange vector, Kan^R	Stibitz, unpublished work
	pXQZ003	Δ*sigE* in pSS3962	This work
	pEV	Vector pJS72, Ω$Spec^R$ cassette replaced with Cm^R	This work
	pSigE	*sigE* in pEV	This work
	pCW505	cytoplasmic expression of GFP	[174]

Construction of RB50Δ*sigE*

The *sigE* gene was deleted from RB50 using a *Bordetella*-specific allelic exchange procedure to produce strain SEA5516. Primers used in the construction are listed in Table 2. A PCR product containing 637 bp proximal to the 5' end of *sigE* was amplified from RB50 genomic DNA using primers SigEKO_LeftF and SigEKO_LeftR. A non-overlapping PCR product containing 534 bp proximal to the 3' end of *sigE* was amplified with primers SigEKO_RightF and SigEKO_RightR. The two fragments were digested with BamHI and ligated. The resulting construct was amplified with primers SigEKO_LeftF and SigEKO_RightR, cloned into the TopoTA vector (Invitrogen), and verified by sequencing to give plasmid pXQ002. In this deletion construct, the 528 bp central region of the *sigE* gene is deleted leaving 66 bp at the 5' end and 6 bp at the 3' end of the *sigE*

gene. The deletion construct from pXQ002 was then cloned into the EcoRI site of the allelic exchange vector pSS3962 (Stibitz S., unpublished data) to generate pXQ003 and transformed into *E. coli* strain DH5α. Tri-parental mating with wild-type *B. bronchiseptica* strain RB50, *E. coli* strain DH5α harboring the pXQ003 vector (strain XQ003), and DH5α harboring the helper plasmid pSS1827 (strain SS1827) [175, 176] and selection of mutants were performed as previously described [175]. The deletion strain was verified by PCR using primers SigEKO_LeftF and SigEKO_RightR and by Southern blot analysis.

Table 2-2: Primer sequences

Primer name	Sequence (5′ - 3′)	Source or Reference
SigEF	GGCGGAGAATTCAGGAGGAGGCGTCATGAGCGAACGCGATG	This work
SigER	GGCCTAGGATCCTTACCAGCGACGCTCGGCAT	This work
HisSigEF	GGCCTGGCATATGAGCGAACGCGATGTCGA	This work
HisSigER	GGCCTAGGATCCTTACCAGCGACGCTCGGCAT	This work
72SigEF	GCGCGGTCTAGAAGGAGGAGGCGTCATGAGCGAACGCGATG	This work
72SigER	GCCCGGCTCGAGTTACCAGCGACGCTCGGCAT	This work
72ChlorF	GCGGCGGGATCCTGTGTAGGCTGGAGCTGCTTC	[169]
72ChlorR	GCCGCCGGATCCCATATGAATATCCTCCTTA	[169]
SigEKO_LeftF	GGGAATTCAAGATCGAGATCGGCCTGTCGAAT	This work
SigEKO_LeftR	AGGGATCCGAAGGCTTTCTTGTCGCCACGTTGTA	This work
SigEKO_RightF	AGGGATCCTGGTAAGGAGTGGCAGTCATGCAA	This work
SigEKO_RightR	GCGAATTCAAAGCAACGGTGTCATCAACGTCC	This work
PFamF	GGGCGGGAATTCTGCCGTTCGTGGATGTCCAG	This work
PFamR	GGGCGGAAGCTTGGGCCAACGAACTACTGGGT	This work

β-galactosidase assays

Overnight cultures were diluted into fresh medium and grown to an OD_{600} of 0.1-0.2 at 30 °C. Where indicated, IPTG was added to a final concentration of 1 mM. Samples were collected 2.5 hours later and β-galactosidase activity from the σ^E-dependent reporter was assayed as previously described [82, 177].

Complementation of *E. coli* Δ*rpoE* by *B. bronchiseptica sigE*

The ability of *B. bronchiseptica sigE* to suppress the lethality caused by deletion of *rpoE* in *E. coli* was determined using a co-transduction assay as described [71]. The *rpoE*::*kan nadB*::Tn*10* allele from strain SEA4114 was moved via P1transduction into strain SEA5005, which carries *sigE* on the plasmid pSEB006. Tet-resistant (tet^R) transductants were selected and then screened for kanamycin resistance (kan^R). Although the *nadB* and *rpoE* alleles are tightly linked (>99%), cotransduction resulting in tet^R kan^R colonies will only occur if *rpoE* is no longer essential for viability. In transductions with *E. coli* expressing *sigE* (strain SEA5005) as the recipient strain, 31 out of 32 tet^R transductants were also kan^R. In contrast, none of the 39 tet^R transductants were kan^R when *E. coli* carrying the empty cloning vector (strain SEA008) was the recipient strain.

Protein purification

N-terminally His-tagged *B. bronchiseptica* SigE and *E. coli* σ^E were purified from strain XQZ001 and SEA5036, respectively, as previously described for *E. coli* σ^E [92]. Briefly, cells were grown at 25 ºC to an OD_{600} of 0.5, at which point IPTG was added to induce protein production. Following 1.5-3 hours of induction, cells were harvested by centrifugation and resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 2.5 mM β-mercaptoethanol, 1 mM PMSF). Resuspended cells were then lysed by sonication, and the lysate cleared by centrifugation. The supernatant containing soluble His-SigE was loaded onto a Ni-NTA column (Qiagen). Bound proteins were eluted with a stepwise gradient of 20, 60, 100, and 200 mM imidazole in column buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2.5 mM β–mercaptoethanol). Fractions containing SigE

were pooled and dialyzed into 20 mM Tris-HCl pH 8.0, 50 mM NaCl, and 2.5 mM β-mercaptoethanol.

In vitro transcription

100 nM *E. coli* core RNA polymerase (Epicentre) was incubated with 400 nM His-SigE or His-σ^E in transcription buffer (40 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT, 0.1 mg/ml BSA) for 10 min at 30 °C to form holoenzyme. Multi-round transcription reactions were initiated by addition of holoenzyme at a final concentration of 40 nM sigma factor and 10 nM core RNA polymerase, to prewarmed (30 ºC) transcription mix containing 5.0 nM supercoiled plasmid template pSEB015 [92] or 5.0 nM linear p*fam* template, 5% glycerol, 200 μM ATP, 200 μM CTP, 200 μM GTP, 10 μM UTP, and 2.5 μCi [α-^{32}P]UTP in transcription buffer. After 10 min at 30 °C, reactions were stopped by the addition of stop solution (80% formamide, 20 mM EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue). Samples were electrophoresed on 6% polyacrylamide gels containing 7.5 M urea, and transcripts were visualized by phosphorimaging. The linear p*fam* template was generated by amplification of the promoter region of the gene encoding σ^{32} in RB50, *fam*, using the primers PFamF and PFamR (Table 2).

Disk diffusion assays

B. bronchiseptica cultures in mid-log phase were diluted to 6×10^8 CFU/ml and spread on Stainer-Scholte agar plates to generate a lawn of bacteria. Disks containing 300 iu polymyxin B, 10 μg ampicillin, 100 μg mecillinam, 750 μg sodium dodecyl sulfate (SDS) and 2.9 μg EDTA, 30 μg aztreonam, 10 μg imipenem, 10 μg meropenem, 30 μg chloramphenicol, 15 μg erythromycin, 30 μg kanamycin, 30 μg nalidixic acid, 150 μg rifampicin, 23.75 μg sulfamethoxazole and 1.25 μg trimethoprim, 30 μg tetracycline, 3.0

μg deoxycholate, 3% hydrogen peroxide, or 2% paraquat were applied to the plates and the zones of inhibition were measured after overnight incubation at 37 °C.

Temperature and Ethanol stress

For temperature stress experiments, mid-log phase cultures of RB50 and RB50Δ*sigE* were diluted to an OD_{600} of 0.01 in fresh Stainer-Scholte broth and incubated at 37 ºC in a gyratory water bath with shaking. At an OD_{600} of 0.1, cultures were either shifted to 40 °C for adaptation or kept at 37 ºC. After 90 minutes, all cultures were shifted to 50 °C, and survival was measured by plating and CFU counts. For ethanol stress experiments, mid-log-phase cultures of the pertinent strains were subcultured into fresh Stainer-Scholte broth with or without 3% ethanol and incubated at 37 °C in a gyratory water bath with aeration. Bacterial growth was measured by OD_{600}.

Complement killing assay

Complement killing assays were performed as previously described [178]. Briefly, freshly collected blood from C57BL/6 mice was pooled, incubated at 4 °C for 1 hour and centrifuged at 2,000 x g for 10 min. Serum was collected and diluted 1:5 in PBS. Complement-inactive serum was prepared by heating an aliquot of the diluted sample at 56 °C for 30 min. Approximately 500 CFU of RB50, RB50Δ*sigE*, and RB50Δ*wbm* from mid-log phase cultures were incubated with 45 μl of diluted serum or PBS (final volume for incubation was 50 μl) for 1 hour at 37 °C. Bacterial numbers before and after incubation were determined by plating and CFU counts. Each strain was assayed in triplicate.

Cytotoxicity assay

Cytotoxicity assays were performed as previously described [149]. Briefly, RAW 264.7 murine macrophage cells were cultured in Dulbecco modified Eagle medium

(DMEM) (HyClone) with 10% fetal bovine serum (FBS) (HyClone). Cells were grown to 80% confluency in 96-well plates at 37 °C in 5% CO_2 and then washed with RPMI (Mediatech) containing 5% FBS. Bacteria were added in 100 μl RPMI at a multiplicity of infection (MOI) of 10. After a 5 min centrifugation at 250 x g, the RAW 264.7 cells and bacteria were incubated for four hours. Percent lactate dehydrogenase (LDH) release, a measure of cytotoxicity, was determined by using Cytotox96 Kit (Promega) according to the manufacturer's protocol.

Phagocytosis and killing by polymorphonuclear leukocytes

Attachment and phagocytosis of the *B. bronchiseptica* strains by peripheral blood polymorphonuclear leukocytes (PMNs) were evaluated as previously described with a few modifications [179]. Briefly, GFP-expressing bacteria were incubated with PMNs at an MOI of 50 for 20 min at 37 °C to allow binding. After extensive washing to remove non-attached bacteria, an aliquot was maintained on ice to be used as a bacterial attachment control. The remaining PMNs were further incubated for 30 min at 37 °C to allow internalization, at which point phagocytosis was stopped by placing PMNs on ice. Bacteria bound to the cell surface in both aliquots were detected by incubation with RB50 immune serum for 30 min at 4 °C, followed by incubation with R-phycoerythrin (RPE)–labeled goat $F(ab')_2$ fragments of anti-mouse IgG at 4 °C for 30 min. All incubations were done in the presence of 25% heat-inactivated human serum to prevent nonspecific binding of antibodies. After washing, ten thousand cells per sample were analyzed by flow cytometry. Attachment control samples were also analyzed by fluorescence microscopy using a DMLB microscope coupled to a DC 100 camera (Leica Microscopy Systems Ltd.). Green fluorescence intensity associated with PMNs maintained at 37 °C for 20 min has

previously been shown to represent bacterial attachment [179]. Phagocytosis was calculated from the decrease in mean red fluorescence intensity of GFP-positive PMNs after the 30 min incubation allowing for internalization, as previously described [180]. Percent phagocytosis was calculated as follows: 100 ×(1-RPE2/RPE1), where RPE1 is the mean RPE-fluorescence of the GFP-positive cells after 20 min at 37 °C (attachment control) and RPE2 is the mean RPE-fluorescence of the GFP-positive cells after 50 min (internalized bacteria) at 37 °C.

Killing of bacteria by PMNs was assessed as follows: after phagocytosis of the bacteria, 400 μg/ml polymyxin B and 350 μg/ml chloramphenicol were added to the PMNs for 1 hour to kill the remaining extracellular bacteria and assess intracellular survival. Serial dilutions of samples were plated to determine the number of viable intracellular bacteria per PMN. The relative percent survival of internalized bacteria was calculated from the relative phagocytosis index and taking into account the initial attachment level of each strain, as follows: percent bacterial killing = [1-N/(A×P)]×100, where A= number of bacteria associated with PMN after 20 min at 37 °C (determined by fluorescent microscopy), P= phagocytosis index (1-RPE2/RPE1), N= number of viable bacteria per cell after incubation with antibiotics. Control experiments to assess the efficacy of antibiotic bactericidal activity were performed in parallel. Briefly, samples of 5×10^8 bacteria were incubated with antibiotics for 30 min at 37 °C and plated. This resulted in a >99% decrease in CFU.

Animal experiments

C57BL/6J, B6.129S-*Tnf*^{tm1Gkl}/J (TNF-$\alpha^{-/-}$), B6 129S7-*Rag*1^{tm1Mom}/J (Rag1$^{-/-}$), C3H/HeOuJ (TLR4^{suf}) and C3H/HeJ (TLR4^{def}) mice were obtained from Jackson

laboratories (Bar Harbor). All mice were bred in our *Bordetella*-free, specific pathogen-free breeding rooms at The Pennsylvania State University. For inoculation, mice were sedated with 5% isoflurane (Abbott laboratory) in oxygen and 50 μl of PBS containing 10^5 or 5×10^5 CFU of the indicated bacteria were pipeted onto the external nares [144, 181]. This method reliably distributes the bacteria throughout the respiratory tract [181]. Survival curves were generated by inoculating TLR4^{def}, TNF-$\alpha^{-/-}$ and Rag1$^{-/-}$ mice with either RB50 or RB50Δ*sigE*. Mice suffering from lethal bordetellosis as determined by severe hunched posture, ruffled fur, extremely labored breathing and apathy were euthanized to prevent unnecessary suffering [151]. For quantifying bacterial load, mice were euthanized via CO_2 inhalation, and lung, trachea, nasal cavity, spleen, liver and/or kidneys were excised. Tissues were homogenized in PBS, aliquots were serially diluted, plated, incubated at 37 °C for 2 to 3 days, and CFU were determined. All protocols were reviewed by the university IACUC and all animals were handled in accordance with institutional guidelines (IACUC approval number: 31297).

Statistical Analysis

The mean +/- standard error (SE) of the geometric mean was determined when appropriate and expressed as error bars. Two-tailed, unpaired Student's T-tests were used to determine statistical significance between groups. All experiments were performed at least twice with similar results.

Chapter 3

RseA and RseB negatively regulate SigE and contribute to colonization and virulence in mice

Abstract

Bacterial pathogens, such as the mammalian respiratory pathogen *Bordetella bronchiseptica*, encounter a multitude of stress conditions during infection. One method many bacteria have developed to cope with these stresses is activation of the extracytoplasmic function sigma factor σ^E, which redirects transcription to a subset of genes that will allow the cell to respond to the stress. We recently described the role of the σ^E ortholog SigE in *B. bronchiseptica* in response to cell envelope stress and in virulence. In this chapter, I demonstrate that the genes downstream of *B. bronchiseptica sigE, rseA* and *rseB*, encode negative regulators of SigE activity. A strain with high SigE activity due to deletion of *rseA* and *rseB* (RB50Δ*rseAB*) is more resistant to treatment with many cell wall-active antibiotics and SDS+EDTA, but is not more resistant to heat shock or ethanol stress, which require *sigE* for survival. Cells with high, constitutive SigE activity are deficient in colonizing the lower respiratory tract of immunocompetent mice compared to either wild-type RB50 or RB50Δ*sigE*. RB50Δ*rseAB* is also unable to cause lethal infection in mice lacking key components of innate immunity ($TLR4^{def}$ or $TNF\text{-}\alpha^{-/-}$), or in mice lacking the ability to make B-cells or T-cells ($Rag1^{-/-}$). These data suggest that constitutive activation of SigE could directly or indirectly interfere with regulation of some virulence

mechanisms, and that proper regulation of SigE activity is required to fully respond to cell envelope stress, efficiently colonize the lower respiratory tract, and cause lethal disease.

*Infection studies were performed by Xuqing Zhang, and cytotoxicity assays were performed by Sara Hester.

Introduction

Bordetella species are important Gram-negative respiratory pathogens. *Bordetella bronchiseptica* causes disease ranging from asymptomatic colonization to fatal pneumonia in a wide variety of non-human mammals, and can also infect immunocompromised humans [136-138]. The related *Bordetella pertussis* and *Bordetella parapertussis* are the etiological agents of whooping cough in humans, and are thought to have independently evolved from a *B. bronchiseptica*-like progenitor [135, 139]. In order to survive both changing conditions within a host and direct attack by a host immune system, bacterial pathogens such as the bordetellae must be able to sense and respond to external stress. In *Bordetella* species, the BvgAS two-component system senses environmental cues that are not well-understood, and responds by regulating expression of most of the major virulence factors, such as the adhesins, filamentous hemagglutinin, fimbriae, and pertactin, as well as the type three secretion system (T3SS) [27, 141, 182].

These important virulence factors are primarily expressed at the cell envelope, which is the first point of contact between a bacterial cell and the host environment. In order to maintain the integrity of this important compartment, bacteria have developed an array of additional systems dedicated to sensing stress at the envelope and relaying that information to the transcriptional machinery. Included among these cell envelope stress response systems are the extracytoplasmic function (ECF) sigma factors, the largest and most diverse group of alternative sigma factors [39]. ECF sigma factors, such as σ^E, are widely distributed across bacterial phyla; at least 112 bacterial genomes are predicted to encode a σ^E-like sigma factor [40]. σ^E systems have been implicated in cell envelope stress response and virulence in a variety of bacterial pathogens, including *Salmonella enterica*

serovar Typhimurium, *Vibrio cholerae*, *Pseudomonas aeruginosa*, *Haemophilus influenzae,* and *Burkholderia* species [52, 64, 65, 97, 98, 125, 183].

While regulation of the σ^E system has been studied in the most molecular detail in *Escherichia coli*, this paradigm is conserved in most bacteria encoding an σ^E-like sigma factor. The gene encoding σ^E is often co-transcribed with the genes encoding its negative regulators, a transmembrane anti-sigma factor, RseA, and a periplasmic protein RseB [40, 57]. Under non-stress conditions, most σ^E is held at the inner membrane by the cytoplasmic domain of RseA, preventing σ^E from binding to core RNA polymerase. RseB binds RseA in the periplasm and enhances inhibition. Regulated proteolysis of RseA by the membrane-associated serine proteases DegS and RseP releases the cytoplasmic domain of RseA bound to σ^E into the cytoplasm, where RseA is then degraded by cytoplasmic proteases such as ClpXP [184, 185]. This frees σ^E to bind to core RNA polymerase (RNAP) and direct transcription of its regulon. We have recently identified an RpoE-like sigma factor, SigE, in the respiratory pathogen *Bordetella bronchiseptica*, and demonstrated a role for this sigma factor in both stress response and virulence (Chapter 2).

In this work, I show that the genes downstream from *sigE*, *rseA* (originally annotated *bb3751*) and *rseB* (originally annotated *mucB*), encode negative regulators of SigE activity. We have generated a strain lacking these negative regulators that has high, constitutive SigE activity (RB50Δ*rseAB*). This strain, while more resistant than its isogenic parent RB50 to some stress conditions, exhibits defects in colonization and causing disease, indicating that proper regulation of this sigma factor system is important for pathogenesis. The role of RseA and RseB during infection has only previously been investigated in a handful of bacteria, and this role varies in each bacterial species. This

work is one of the first comprehensive descriptions of constitutive activation of an ECF sigma factor leading to a decreased ability to colonize a host and cause disease, despite conferring increased resistance to some isolated stresses.

Results

B. bronchiseptica RseA and RseB share similarity with other negative regulators of RpoE-like sigma factors

In Chapter 2, I showed that *sigE* encodes an RpoE-like sigma factor with significant homology to similar sigma factors from a variety of organisms (Chapter 2). In *B. bronchiseptica*, the genes downstream from *sigE, rseA* (previously annotated *bb3751*) and *rseB* (previously annotated *mucB*), are predicted to encode proteins similar to the negative regulators RseA and RseB in *E. coli* or MucA and MucB in *P. aeruginosa* (Fig. 3-1A, 3-1C). The amino acid sequences of *E. coli* RseA ($RseA_{Ec}$) and *B. bronchiseptica* RseA ($RseA_{Bb}$) are only 20.5% identical/33.3% similar overall. However, there is greater sequence similarity between the cytoplasmic domains that interact with the sigma factor (22.5% identity/35% similarity), than between the periplasmic domains (12.4% identity/20.4% similarity). $RseB_{Ec}$ and $RseB_{Bb}$ share slightly more sequence similarity (23.8% identity/40.3% similarity) than $RseA_{Ec}$ and $RseA_{Bb}$. The *B. bronchiseptica* and *P. aeruginosa* systems share roughly the same amount of sequence similarity as the *B. bronchiseptica* and *E. coli* systems; $RseA_{Bb}$ and MucA share 24.3% identity/36% similarity, and $RseB_{Bb}$ and MucB share 26.7% identity/38.8% similarity. Although the primary sequences of $RseA_{Ec}$ and $RseA_{Bb}$ are only modestly similar, many of the residues known to be important for $RseA_{Ec}$:σ^E binding are conserved. For instance, $RseA_{Ec}$-D11 and W33 are required for interaction with σ^E (Fig. 3-1, star); σ^E activity increases upon mutation of either, and a D11H mutation abolishes binding [56, 71].

Because many of the residues and regions important for $RseA_{Ec}$:σ^E interaction are conserved in *B. bronchiseptica*, I predicted that $RseA_{Bb}$-cyto might have a similar structure to *E. coli* RseA. I submitted the amino acid sequence for the cytoplasmic domain of *B. bronchiseptica* RseA to SwissModel, which searches the Protein DataBase (PDB) for predicted structural similarity to published crystal structures [186-188]. The only closely related structure identified was that of the complex of *E. coli* σ^E with the cytoplasmic domain of $RseA_{Ec}$ ($RseA_{Ec}$-cyto) [189]. An alignment of the predicted structure of the SigE:$RseA_{Bb}$-cyto interaction, based on the structure of σ^E:$RseA_{Ec}$-cyto, is shown in Figure 3-1B. Many of the residues identified above as important for σ^E-binding in *E. coli* are predicted to adopt a similar orientation in this predicted structure for *B. bronchiseptica* RseA. While these putative structural similarities between $RseA_{Ec}$ and $RseA_{Bb}$ suggest that they may function similarly in their role as negative regulators of σ^E activity, sequence and minor structural differences between these systems suggest the possibility that some of the sigma:anti-sigma interactions could be species-specific.

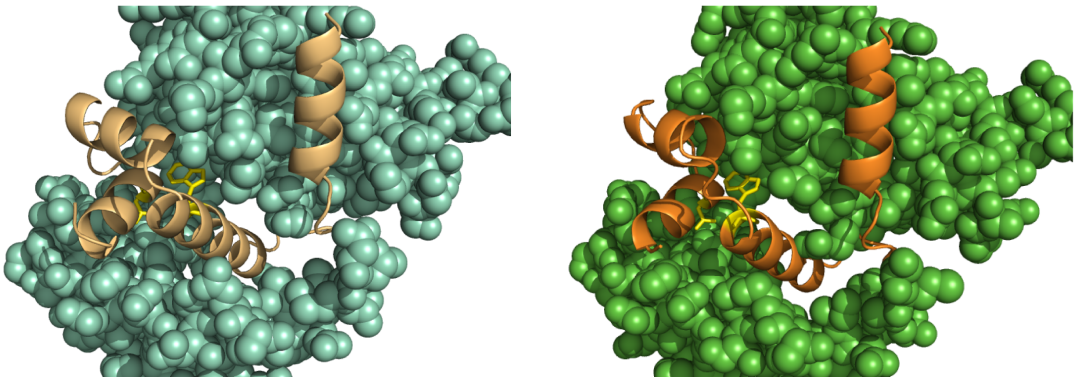
A

```
Ecoli   ------------MQKEQLSALMDGETLDSELLNELAHN---PEMQKTWESYHLIRDSMRGDTPEVLHFDISSRVMAAIEEEPVR 69
Paer    --------MSREALQETLSAVMDNEADELELRRVLAACGEDAELRSTWSRYQLARSVMH-REPTLPKLDIAAAVSAALADEAAP 75
Bbron   MQTAAKSLTAESSWEESVSAWMDGEGSE-DFFPEMSHP----HGRQTWDTYHLIGDALR-------NADLAVSPSAQFQARLAR 72
                      :* :** **.*  : ::   ::      . :.**. *:*  . ::       : *::    * :  . .

Ecoli   QPATLIPEAQPAPHQWQKMPFWQKVRPWAAQLTQMGVAACVSLAVIVGVQHYNGQSETSQQPETPVFNTLPMMGKASPVSLGVP 153
Paer    ------PKAEKGP--------WRMVG-------RLAVAASVTLAVLAGVRLYNQNDALPQMAQQGTTPQIALPQVKGPAVLAGY 139
Bbron   ------------------------------------ALEAELPIVAAPRRRSPLR-LGLSGLAVAAAVATVAWVAQPYLTGAP 128
                                           *  . *.::.. :  .            .    .:     *   .

Ecoli   SEATANNGQQQQVQEQRRRINAMLQDYELQRRLHSEQLQFEQAQTQQAAVQVPGIQTLGTQSQ 216
Paer    SEEQG----APQVITNSSSSDTRWHEQRLPIYLR-QHVQQSAVSGTESALPYARAASLENR-- 194
Bbron   TGETR--------VLADATPASAGDDAGLRDYLEAHRQVAGPSAVRQVSFDAGAGR------- 166
        :                   :  .:  *   *. .:         : :.
```

B



E. coli σ^E:RseA *B. bronchiseptica* SigE:RseA

C

```
Ecoli   ----------------------MKQLWFAMSLVTGSLLFSANASATPASG----ALLQQMNLASQSLNYELSFISINKQGVES 57
Paer    ------------------------MRTTSLLLLLGSLMAVPATQAADAS-----DWLNRLAEADRQNSFQGTFVYERNGSFST 54
Bbron   MMTPLLYMRRPALGRMTWRQRQTASLAAMAVLMGLMAIYARPVRAAAPAVPEDPVLLLSSIQQAARKQDYAGVFMYQQGESIQS 83
                                  :: :   ::   .   *: *        *. :  * :. .:   *:  .  ...:

Ecoli   LRYRH-ARLDNRPLAQLLQMDGPRREVVQRGNEISYFEPGLEPFTLNGDYIVDSLPSLIYTDFKRLSPYYDFISVGRT-RIADR 139
Paer    HEIWHRVESDGAVRERLLQLDGARQEVVRVDGRTQCISGGLADQLAD-----AQLWPVRKFDPSQLASWYDLRLVGES-RVAGR 132
Bbron   SRLVHVLDGTGE-RERLEILDGQPREYLRHNEDVQCLIPERKTVLLERRR-GDRFPGLLLGDPAGLAENYHIRAESQLHRVADR 165
         .  *     .    :*  :**  :* :: .   . :          :       :  :   *   *:  *.:   ..  *:*.*

Ecoli   LCEVIRVVARDGTRYSYIVWMDTESKLPMRVDLLDRDGETLEQFRVIAFNVNQDISSSMQTL--AKANLPPLLSVPVGEKAKFS 221
Paer    PAVVLAVTPRDQHRYGFELHLDRDTGLPLKSLLLNEKGQLLERFQFTQLNT-GAAPAEDQLQ--AGAECQVVGPAKADGEKTVA 214
Bbron   QCRLITVEPRDALRYGYRLCADVATNLLLKAQTLDAARRVVEQVSFTSIRLGGDVDPAMLTSRWSTRDWRVLEPVMKPVDLAAQ 249
         . :: * .**  **.: :  *  : * ::   *:   . :*:. .  :.      .       :  :   : ..    .

Ecoli   WTPTWLPQGFSEVSSSRRPLPTMDNMPIESRLYSDGLFSFSVNVNRATPSSTDQMLRTGRRTVSTSVRDNAE-----ITIVGEL 300
Paer    WRSEWLPPGFTLTRSFMRRSPVTP-DPVACLTYGDGLARFSVFIEPLHGAMVGDARSQLGPTVVVSKRLQTDDGGQMVTVVGEV 295
Bbron   GWRITAPKGFSAVLQVAR---SMGPATVNQLVLSDGLAAISVFIEPYDSRRHAHPPEGLARRGAINIYGTRIAD-FWLTALGEV 329
              * **: . .  *        .:     .***  :** ::        .           .           :* :**:

Ecoli   PPQTAKRIAEN---IKFGAAQ-- 318
Paer    PLGTAERVALS---IRPEAAAQK 316
Bbron   PAGTLEQLAETTEYVPVDAAKQQ 353
        *  * :::* .   :   **
```

Figure 3-1: Sequence and structural similarity in the negative regulators of the SigE system. (A) Sequence alignment of *E. coli* RseA (Ecoli), *P. aeruginosa* MucA (Paer), and *B. bronchiseptica* RseA (Bbron). Asterisks denote identity in all three organisms; two dots denote strong conservation, and one dot denotes weak conservation. Red stars indicate

residues particularly important for the *E. coli* σ^E:RseA interaction (B) Structural alignment of the *E. coli* σ^E:RseA-cyto complex (left, accession number 1OR7) with the predicted structure of the *B. bronchiseptica* SigE:RseA-cyto complex (right). σ^E_{Ec} or modeled $SigE_{Bb}$ regions 2 (top) and 4 (bottom) are shown in green spheres; RseA-cyto is shown in orange ribbons. Yellow residues indicate two of the conserved residues in RseA identified as being particularly important for σ^E-RseA-cyto interaction (*E. coli* RseA-D11 and W33) [56, 71, 189]. Structures were visualized using PyMOL [190]. (C) Sequence alignment of *E. coli* RseB (Ecoli), *P. aeruginosa* MucB (Paer), and *B. bronchiseptica* RseB (Bbron). Asterisks denote identity in all three organisms; two dots denote strong conservation, and one dot denotes weak conservation.

RseA and RseB negatively regulate SigE

I have demonstrated that *B. bronchiseptica* SigE is an RpoE-like sigma factor that directs transcription from the promoter region of *B. bronchiseptica fam* (σ^{32}) in vitro (Chapter 2). To test whether $RseA_{Bb}$ and $RseB_{Bb}$ are negative regulators of $SigE_{Bb}$, we first generated a mutant of the *B. bronchiseptica* strain RB50 lacking both *rseA* and *rseB* (RB50Δ*rseAB*) using a previously described allelic exchange strategy [191]. I then amplified the *B. bronchiseptica fam* promoter region, cloned it upstream of a promoterless *lacZ* gene in the plasmid pMP220 [192], and moved this construct into RB50, RB50Δ*sigE* and RB50Δ*rseAB*. Reporter activity was reduced to nearly basal levels in a mutant lacking *sigE*, indicating that this reporter is responsive to SigE activity (Fig. 3-2A). Reporter activity was increased in RB50Δ*rseAB* when compared with RB50 (Fig. 3-2A), indicating that RseA and RseB function as negative regulators of SigE activity in *B. bronchiseptica*.

To examine the inhibition of *B. bronchispetica* SigE by $RseA_{Bb}$ and $RseB_{Bb}$ in more detail, I cloned the genes encoding *B. bronchiseptica sigE, rseA,* and *rseB* into a plasmid under the control of an IPTG-inducible promoter, and transformed each construct into an *E. coli* strain carrying the σ^E-dependent *rpoH*P3::*lacZ* fusion to monitor SigE activity by β-galactosidase assay [82]. I previously demonstrated that $SigE_{Bb}$ directs

transcription of *rpoH*P3 in vitro and in *E. coli* (Chapter 2). A strain over-expressing *sigE* alone increased reporter activity about 6-fold above basal levels (Chapter 2 and Fig. 3-2B). There was a basal level of SigE activity in a strain harboring only an empty vector because *E. coli* σ^E was still present in these strains.

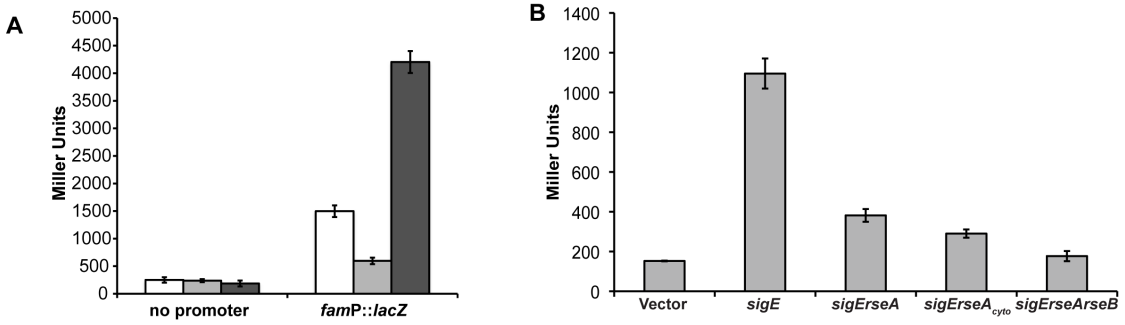


Figure 3-2: *B. bronchiseptica* RseA and RseB are negative regulators of SigE activity. (A) β-galactosidase activity of the plasmid-encoded *B. bronchiseptica* SigE-dependent reporter *fam*P::*lacZ* in wild-type RB50 (white bars), RB50Δ*sigE* (light grey bars), and RB50Δ*rseAB* (dark grey bars). A control plasmid with promoterless *lacZ* showed basal β-galactosidase activity in each strain (no promoter). The average of two to five independent experiments ± SE is shown. (B) β-galactosidase activity of the *E. coli* σ^E-dependent reporter *rpoH*P3::*lacZ* following overexpression of the various components of the *B. bronchiseptica* SigE system. A control plasmid, pTrc99a, has basal β-galactosidase activity, as *E. coli* σ^E is still encoded on the chromosome in these strains. The average of at least two independent experiments ± SE is shown.

In a strain over-expressing both *sigE* and *rseA,* reporter activity was significantly decreased, indicating that RseA negatively regulates SigE. Reporter activity returned to basal levels in a strain over-expressing *sigE*, *rseA*, and *rseB*, indicating that both RseA and RseB are important negative regulators of SigE activity (Fig. 3-2B). In *E. coli,* the gene encoding σ^E, *rpoE*, is essential [61]. *rpoE* could be deleted from a strain expressing either $sigE_{Bb}$ or both $sigE_{Bb}$ and $rseA_{Bb}$ from a plasmid, but not from a strain where $sigE_{Bb}$, $rseA_{Bb}$,

and *rseB*$_{Bb}$ were all expressed together (Table 3-1). This provides further evidence that both proteins are important for full inhibition of SigE$_{Bb}$ activity.

Table 3-1: Co-transduction of *nadB*::Tn10 and *rpoE*::Kan^R in *E. coli* cells expressing components of the SigE system. Number of colonies carrying each allele are indicated. Vector alone was a negative control, and the unmapped suppressor strain CAG43113 was a positive control for deletion of *rpoE* [71].

	unmapped suppressor	vector	*sigE*	*sigErseA*	*sigErseArseB*
nadB::Tn10 (Tet^R)	116	39	32	26	38
rpoE::Kan^R	116	0	31	25	0

The cytoplasmic domain of *E. coli* RseA (RseA-cyto) is sufficient to bind σ^E and prevent it from interacting with core RNAP [189]. In a strain of *E. coli* over-expressing both SigE$_{Bb}$ and the cytoplasmic domain of *B. bronchiseptica* RseA (RseA$_{Bb}$-cyto), *E. coli* σ^E reporter activity was similar to that of a strain over-expressing SigE$_{Bb}$ and full-length RseA$_{Bb}$, indicating that RseA$_{Bb}$-cyto is sufficient to inhibit SigE$_{Bb}$ activity (Fig. 3-2B).

To further confirm a direct inhibitory interaction between SigE$_{Bb}$ and RseA$_{Bb}$-cyto, I purified C-terminally His-tagged RseA$_{Bb}$-cyto, and performed in vitro transcription using holoenzyme formed with His-SigE$_{Bb}$ and *E. coli* core RNAP (ESigE$_{Bb}$). Addition of a 10-fold molar excess of RseA$_{Bb}$-cyto-His decreased transcription by ESigE$_{Bb}$ by nearly 70% (Fig. 3-3). For comparison, a 10-fold molar excess of RseA$_{Ec}$-cyto-Strep decreased transcription by Eσ^E_{Ec} by more than 90%. While RseA$_{Ec}$-cyto-Strep was able to inhibit SigE$_{Bb}$, RseA$_{Bb}$-cyto-His did not decrease σ^E_{Ec}-dependent transcription (Fig. 3-3). Overexpression of RseA$_{Bb}$ in an *E. coli* strain lacking *rseA*$_{Ec}$ also did not decrease *E. coli* σ^E activity (data not shown), providing further evidence that there are differences in how RseA$_{Bb}$ interacts with SigE$_{Bb}$ compared to how RseA$_{Ec}$ interacts with either σ^E_{Ec} or SigE$_{Bb}$.

Neither $RseA_{Bb}$-cyto-His nor $RseA_{Ec}$-cyto-Strep decreased transcription by $E\sigma^{70}_{Ec}$, indicating that these proteins do not generally inhibit all transcription (Fig. 3-3).

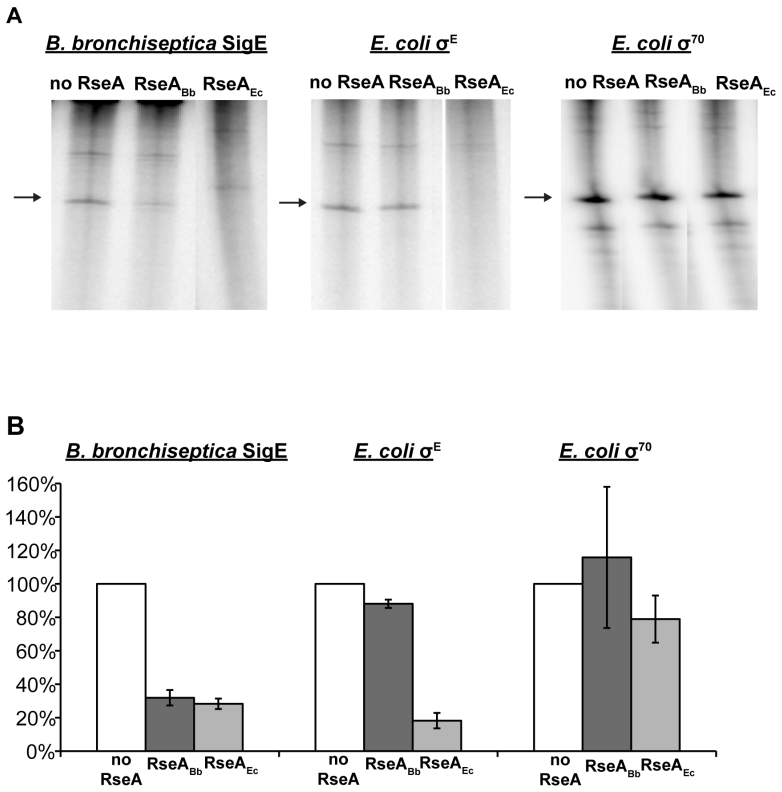


Figure 3-3: *B. bronchiseptica* RseA inhibits $SigE_{Bb}$, but not σ^{E}_{Ec}; *E. coli* RseA inhibits both $SigE_{Bb}$ and σ^{E}_{Ec} (A) The indicated sigma factor ($SigE_{Bb}$, σ^{E}_{Ec}, or σ^{70}_{Ec}) was incubated with the cytoplasmic domain of $RseA_{Bb}$ or $RseA_{Ec}$, or buffer (no RseA), allowed to form holoenzyme with *E. coli* core RNAP, then added to a transcription mix containing 50 mM NaCl and 5 nM supercoiled template pSEB015 (*rpoH*P3). (B) Quantification of at least three independent experiments ± SE is shown. Bb, *B. bronchiseptica*; Ec, *E. coli*

RB50Δ*rseAB* is more resistant to some cell envelope stress

RB50Δ*sigE* grew similarly to wild-type RB50 under non-stress conditions, but is more sensitive to some stress conditions, such as heat shock, ethanol stress, and some agents that attack the cell envelope – ampicillin, mecillinam, and SDS+EDTA (Chapter 2). A strain with high SigE activity (RB50Δ*rseAB*) grew slightly more slowly under non-stress

conditions, at 37 °C in Stainer-Scholte broth, than its isogenic parent or RB50Δ*sigE* (Table 3-2, Chapter 2). Expression of *rseA* and *rseB* from a plasmid (pRseAB), but not an empty vector (pEV), partially complemented this growth defect (Table 3-2).

Table 3-2: Growth rate of strains under non-stress conditions. The average doubling time of the indicated strains in at least three independent experiments ± SE is shown.

Strain	Doubling time (min)
RB50	64.9 ± 0.8
RB50Δ*sigE*	65.0 ± 0.6
RB50Δ*rseAB*	75.4 ± 1.6
RB50Δ*rseAB* pEV	77.7 ± 0.6
RB50Δ*rseAB* pRseAB	71.5 ± 1.0

E. coli σ^E was first identified for its role in transcribing the heat shock sigma factor σ^{32} under conditions of extreme heat stress [67]. I have shown that SigE transcribes the gene encoding σ^{32} in *B. bronchiseptica* (*fam*), and that cells lacking *sigE* are more sensitive to heat shock (Chapter 2). RB50Δ*sigE* also exhibits reduced thermotolerance, another aspect of the heat shock response (Chapter 2). Cells exposed to a high but non-lethal temperature often accumulate heat shock proteins that allow them to better survive a subsequent shift to a lethal temperature. RB50Δ*rseAB* responded to heat shock from 37 °C to 50 °C similarly to wild-type RB50 (Fig. 3-4A, closed symbols). When the cells were first shifted from 37 °C to 40 °C, an elevated but sublethal temperature, then shifted from 40 °C to 50 °C, RB50Δ*rseAB* survived slightly less well than RB50 (Fig. 3-4A, open symbols). Cells expressing a plasmid-encoded copy of *sigE* in the RB50Δ*sigE* background also showed decreased thermotolerance compared to wild-type (Chapter 2); in other organisms, such as *Burkholderia cenocepacia*, expressing *sigE* in trans without its native regulation does not complement the phenotype tested, but rather exacerbates the survival

defect under these conditions [50]. These results suggest that proper regulation of SigE activity is crucial for adapting to and surviving heat stress, and that simply increasing expression of a particular stress factor, such as SigE, does not necessarily increase resistance to a stress condition.

A strain lacking *sigE* is more sensitive to treatment with mecillinam, ampicillin, and the detergent SDS+EDTA (Chapter 2). RB50Δ*rseAB* was much more resistant to these agents, as well as the β-lactam meropenem and the cell-wall active antibiotic aztreonam. RB50Δ*sigE* is not more sensitive to meropenem or aztreonam (Chapter 2, data not shown). These results indicate that high SigE activity contributes to a cell's ability to cope with cell envelope stress. Plasmid-encoded *rseA* and *rseB* restored sensitivity to SDS+EDTA, aztreonam, mecillinam, and meropenem, and may have restored some sensitivity to ampicillin, indicating that resistance was conferred by deletion of these two regulators (Fig. 3-4B). RB50Δ*rseAB* and RB50 were equally sensitive to the cationic antimicrobial peptides polymyxin B and polymyxin E (colistin), which target the membrane (Fig. 3-4B and data not shown), the β-lactam imipenem, and other antibiotics that inhibit translation (chloramphenicol, erythromycin, kanamycin, tetracycline) and transcription (rifampicin). This suggests that the high SigE activity resulting from deletion of *rseA* and *rseB* does not confer general antibiotic resistance, but allows the cells to survive in the presence of specific cell envelope stresses.

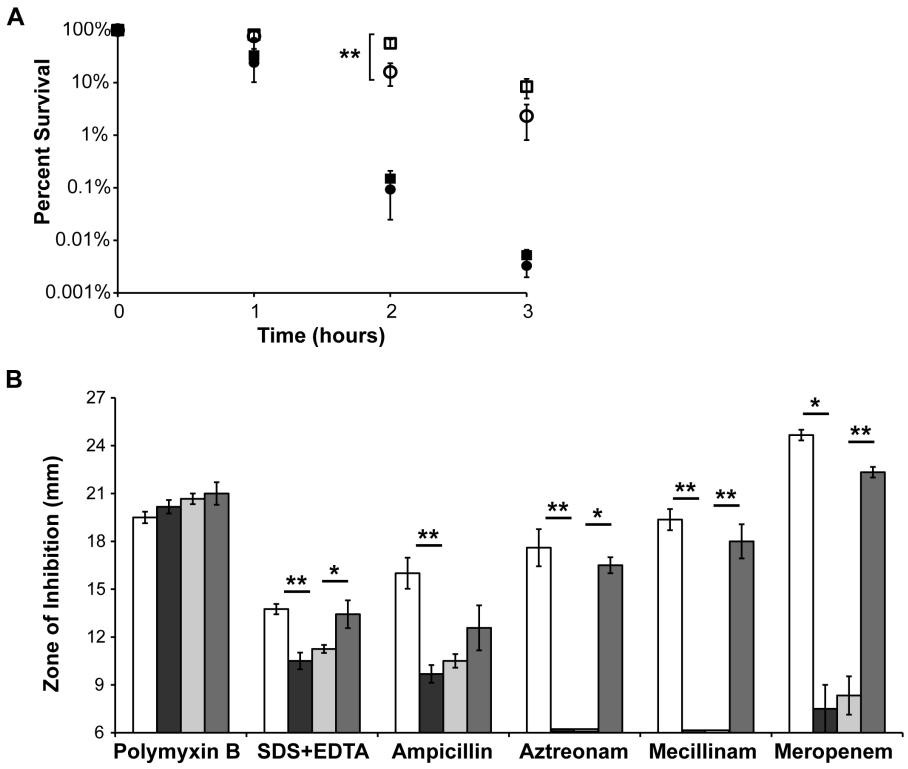


Figure 3-4: RB50Δ*rseAB* exhibits reduced thermotolerance, but increased resistance to some cell envelope stress. (A) RB50Δ*rseAB* (circles) is similarly sensitive to a shift from 37 °C to 50 °C (filled symbols) as RB50 (squares), but exhibits slightly reduced thermotolerance after adaptation at 40 °C before a shift to 50 °C (open symbols). An average of at least three independent experiments ± SE is shown. (B) Cells with constitutive SigE activity are more resistant to treatment with SDS+EDTA and to antibiotics that act on the cell wall. Disk diffusion assays were performed with RB50 (white bars), RB50Δ*rseAB* (dark grey bars), RB50Δ*rseAB* pEV (light grey bars), or RB50Δ*rseAB* pRseAB (medium grey bars). The disk diameter is 6 mm (bottom of y-axis). The average of at least two independent experiments ± SE is shown. * denotes a statistically significant difference with a p-value <0.05, ** denotes p-value <0.01.

Constitutive SigE activity leads to decreased colonization in the lower respiratory tract

To determine whether high SigE activity affects virulence in *B. bronchiseptica*, C57BL/6 mice were inoculated with $5x10^5$ CFU RB50 or RB50Δ*rseAB* in 50 µL PBS and

euthanized on days 0, 1, 3, 17, 14, and 28 post-inoculation to determine bacterial numbers in the nasal cavity, trachea, and lungs. RB50 colonized all respiratory organs efficiently. We have previously shown that RB50Δ*sigE* colonizes the respiratory tract of C57BL/6 mice similarly to wild-type RB50 (Chapter 2). While RB50Δ*rseAB* still colonizes and persists in the nasal cavity, significantly lower numbers of bacteria were recovered from the lower respiratory tract of animals inoculated with this mutant (Fig. 3-5). The colonization defect in the lower respiratory tract is not likely due to the slower growth observed in culture (Table 3-1), as the bacteria can colonize and persist efficiently in the nasal cavity. This indicates that constitutively high SigE activity is more detrimental to colonization and persistence than no SigE activity.

While defects in colonization of the lower respiratory tract begin early during infection, the difference between RB50 and RB50Δ*rseAB* persistence is exacerbated at later timepoints. RB50Δ*rseAB* is cleared from the lungs and trachea by day 28 post-inoculation, whereas RB50 is not cleared until about day 63 post-inoculation (Fig. 3-5 and Chapter 2). This suggests that RB50Δ*rseAB* could interact differently with the adaptive immune system than RB50.

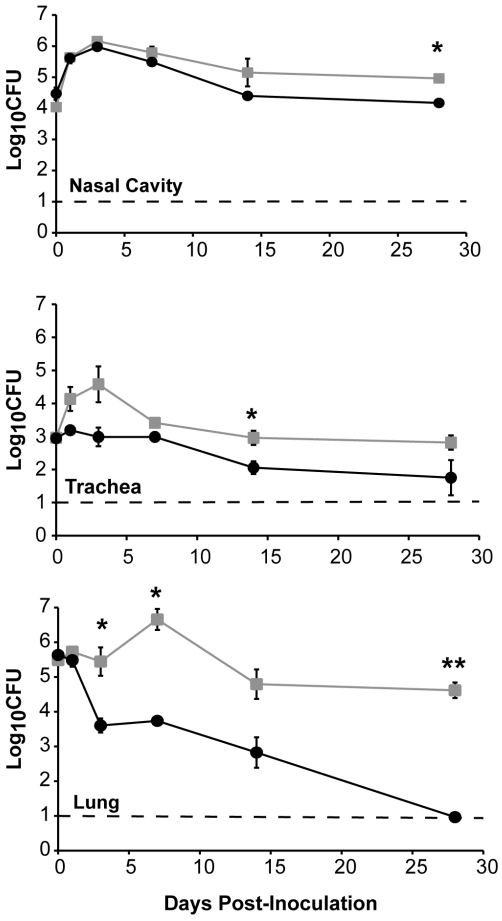


Figure 3-5: RB50Δ*rseAB* exhibits decreased colonization in the lower respiratory tract. CFU of either RB50 (grey squares) or RB50Δ*rseAB* (black circles) in the indicated organs of C57BL/6 mice were determined. Mean CFU ± SE of bacteria from at least 3 mice is shown. Dashed line indicates threshold of detection, * denotes p-value<0.05, and ** denotes p-value <0.01

Early clearance from the lower respiratory tract is not due to altered interactions with adaptive immunity

One possible explanation for the early clearance of RB50Δ*rseAB* from the lower respiratory tract is that more or different antibodies are being produced during infection with RB50Δ*rseAB* than during infection with RB50, promoting earlier clearance. To test this, we adoptively transferred immune serum from naïve mice, RB50- or RB50Δ*rseAB*-

vaccinated mice to C57BL/6 mice, and simultaneously inoculated these mice with RB50Δ*rseAB*. Mice were sacrificed after 12 hours to enumerate bacterial loads in the respiratory tract. Compared to naïve serum, both RB50 and RB50Δ*rseAB* immune serum had no effect in bacterial numbers in the nasal cavity (Fig. 3-6A). Immune serum from mice vaccinated with either bacterial strain rapidly reduced RB50Δ*rseAB* numbers in the lower respiratory tract to a similar extent. These data indicate that antibodies generated in RB50Δ*rseAB*-vaccinated animals are similarly effective at clearing RB50Δ*rseAB* when compared to antibodies generated in RB50-vaccinated animals, and provide preliminary evidence that the quicker clearance of RB50Δ*rseAB* may not be attributable to differential antibody functions.

Although antibodies generated against heat-killed RB50 and RB50Δ*rseAB* cleared RB50Δ*rseAB* from the respiratory tract to the same extent, constitutive activation of SigE could alter other interactions with the adaptive immune response during infection. To further investigate this possibility, we examined the ability of RB50Δ*rseAB* to cause systemic, lethal infection in mice lacking B cells and T cells. $Rag1^{-/-}$ mice were inoculated with $5x10^5$ CFU of RB50 or RB50Δ*rseAB* and euthanized 21 days post-inoculation to enumerate bacterial numbers (Fig. 3-6B, *left*). If the defect in lower respiratory tract colonization observed in immunocompetent (C57BL/6) mice was due to interactions with adaptive immunity, similar numbers of RB50 and RB50Δ*rseAB* should be recovered from the lower respiratory tract of mice lacking an adaptive immune response ($Rag1^{-/-}$). While similar numbers of RB50Δ*rseAB* and RB50 were recovered from the nasal cavity of these mice, lower numbers of RB50Δ*rseAB* were recovered from the lower respiratory tract of $Rag1^{-/-}$ mice than RB50, similar to the defect observed in C57BL/6 mice, indicating that

the colonization defect in the lower respiratory tract is likely not due to altered interactions with adaptive immunity.

In Rag1$^{-/-}$ mice, *B. bronchiseptica* strain RB50 escapes from the respiratory tract, colonizes systemic organs, and ultimately causes lethal disease [148]. In a separate experiment, Rag1$^{-/-}$ mice were inoculated with either RB50 or RB50Δ*rseAB* and monitored for survival. RB50-inoculated Rag1$^{-/-}$ succumbed to infection by day 35 post-inoculation. Rag1$^{-/-}$ mice inoculated with RB50Δ*rseAB*, however, remained healthy (Fig. 3-6B, *right*). The inability of RB50Δ*rseAB* to cause lethal infection in the absence of a functional adaptive immune response indicates that an important interaction with the innate immune system is disrupted when SigE is constitutively activated. We therefore examined how constitutive activation of SigE affected interactions of *B. bronchiseptica* with components of the innate immune system.

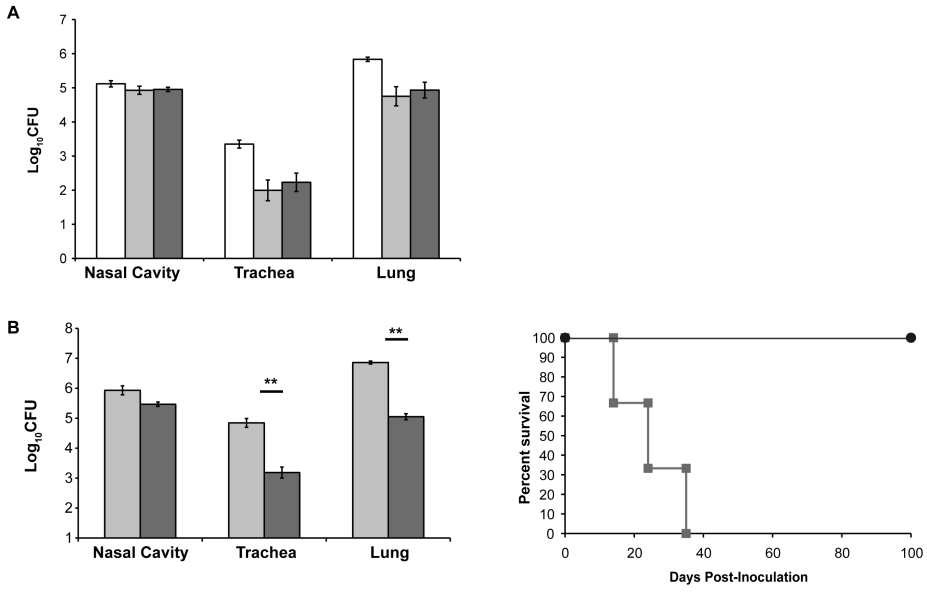


Figure 3-6: RB50Δ*rseAB* is equally susceptible to antibodies generated against RB50 and RB50Δ*rseAB*, and fails to cause lethal infection in Rag1$^{-/-}$ mice. (A) Serum from naïve mice (white bars), mice vaccinated with RB50 (light grey bars), or mice vaccinated

with RB50Δ*rseAB* (dark grey bars), was adoptively transferred to C57BL/6 mice simultaneously inoculated with RB50Δ*rseAB*. Bacterial numbers in respiratory organs were enumerated 12 hours after inoculation. (B) *Left,* Log_{10}CFU of RB50 (light grey bars) or RB50Δ*rseAB* (dark grey bars) recovered from the indicated organs on day 21 post-inoculation. Average CFU ± SE is shown; * denotes p-value<0.05, ** denotes p-value<0.01. *Right,* While $Rag1^{-/-}$ mice inoculated with RB50 succumb to lethal infection by about 35 days post-inoculation, $Rag1^{-/-}$ mice inoculated with RB50Δ*rseAB* survive.

RB50Δ*rseAB* does not cause systemic, lethal infection in $TLR4^{def}$ mice and has altered interactions with components of innate immunity

In mice lacking a key component of the innate immune response, TLR4, *B. bronchiseptica* quickly escapes the respiratory tract and causes systemic, lethal infection [149-151]. Cells lacking *sigE* still colonize systemic organs and cause lethal disease in TLR4-deficient ($TLR4^{def}$) mice (Chapter 2). $TLR4^{def}$ mice were challenged with $5x10^5$ CFU of RB50 or RB50Δ*rseAB* and monitored for survival. $TLR4^{def}$ mice inoculated with RB50 developed severe bordetellosis and succumbed two to five days post-inoculation. Mice inoculated with RB50Δ*rseAB,* however, did not display any disease symptoms during a 100-day period, and were euthanized at the end of the experiment (Fig. 3-7A). In a separate experiment, groups of TLR4-sufficient ($TLR4^{suf}$) or $TLR4^{def}$ mice challenged with RB50 or RB50Δ*rseAB* were euthanized three days post-inoculation, when some $TLR4^{def}$ mice inoculated with RB50 were still alive, and bacterial numbers were enumerated in respiratory and systemic organs. While bacterial numbers of both RB50 and RB50Δ*rseAB* were statistically similar in the nasal cavity and trachea of both $TLR4^{suf}$ and $TLR4^{def}$ mice, significantly fewer numbers of RB50Δ*rseAB* were recovered from the lungs, spleen and kidney of $TLR4^{def}$ mice than wild-type RB50 (Fig 3-7C).

Although fewer RB50Δ*rseAB* cells than RB50 cells were recovered from the lungs of TLR4^{def} mice, RB50Δ*rseAB* numbers were still higher in TLR4^{def} mice than in TLR4^{suf} control mice. During enumeration of bacterial numbers, we noted that the lungs of TLR4^{def} mice inoculated with RB50Δ*rseAB* did not appear as damaged as those inoculated with RB50. To quantify this, samples of lung tissue were taken from TLR4^{suf} and TLR4^{def} mice inoculated with either RB50 or RB50Δ*rseAB*, and scored for histopathology at the Animal Diagnostic Laboratories Facility at Penn State University. Decreased tissue damage was observed in the lungs of RB50Δ*rseAB*-challenged TLR4^{def} mice compared to RB50-challenged TLR4^{def} mice (Fig. 3-7B). These results suggest that constitutive SigE activity impairs the ability of *B. bronchiseptica* to cause tissue damage, even when it can colonize respiratory tract organs to high numbers.

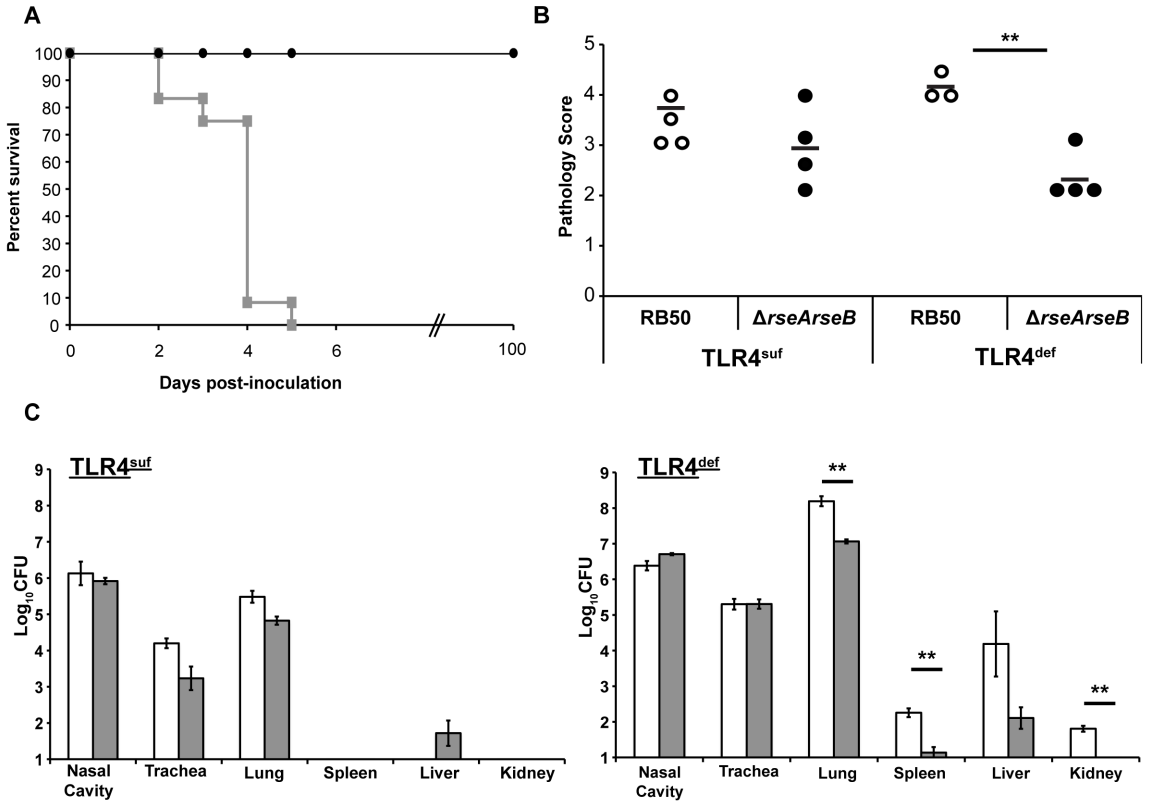


Figure 3-7: RB50Δ*rseAB* does not cause lethal, systemic infection in TLR4^{def} mice: (A) TLR4^{def} mice inoculated with either RB50 (grey squares) succumbed to infection two to five days post-inoculation, while mice inoculated with RB50Δ*rseAB* (black circles) survived. (B) Sections were taken from the lungs of TLR4^{suf} or TLR4^{def} mice inoculated with either RB50 or RB50Δ*rseAB* on day 3 post-inoculation, and pathology was scored. Lines indicate average pathology score. (C) Bacterial numbers of RB50, white, and RB50Δ*rseAB*, grey, recovered from the indicated organs of TLR4^{suf} mice, *left,* or TLR4^{def} mice, *right,* on day 3 post-inoculation.

To determine whether constitutive SigE activity also affects the interaction of *B. bronchiseptica* with other components of the innate immune response, we examined whether RB50Δ*rseAB* is more sensitive to complement, a major bactericidal component of the bloodstream. 500 CFU of RB50, RB50Δ*rseAB*, or RB50Δ*wbm* (a strain lacking O-antigen known to be susceptible to complement-mediated killing [148]) were incubated with 20% complement-active or complement-inactive serum for one hour at 37 °C. RB50Δ*wbm* was completely killed by complement-active, but not complement-inactive serum. RB50 and RB50Δ*rseAB*, however, were both equally resistant to complement-active serum (data not shown). This result indicates that the inability of RB50Δ*rseAB* to cause systemic infection is not due to increased susceptibility to complement-mediated killing.

We also examined whether cells with constitutively high SigE activity interacted differently with phagocytes, another bactericidal component of the innate immune response. *B. bronchiseptica* is cytotoxic to macrophages, and this is dependent on the activity of the type three secretion system (T3SS) and adenylate cyclase toxin [163, 193]. We previously showed that a strain lacking *sigE* is less cytotoxic to macrophages (Chapter 2). RAW264.7 murine macrophages were incubated at a multiplicity of infection (MOI) of 10 with either RB50, RB50Δ*rseAB*, a strain lacking a functional T3SS, Δ*bscN* (WD3), a

strain lacking adenylate cyclase toxin and a functional T3SS, Δ*bscN*Δ*cyaA* (8W1), or media alone. After four hours, cytotoxicity was measured by assay for lactate dehydrogenase (LDH) release. Cells lacking a functional T3SS and/or adenylate cyclase toxin caused little cytotoxicity, comparable to media alone. RB50Δ*rseAB* caused about three-fold less cytotoxicity than RB50 (Fig. 3-8), demonstrating that both high SigE activity and no SigE activity are detrimental to the ability of *B. bronchiseptica* to be cytotoxic to macrophages.

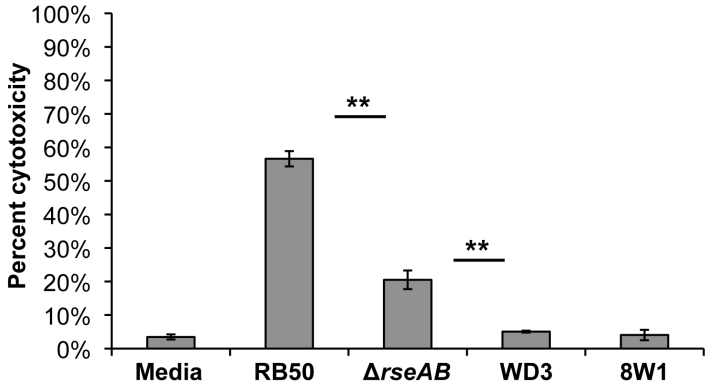


Figure 3-8: Cells lacking *rseA* and *rseB* are less cytotoxic to macrophages. RAW264.7 murine macrophages were incubated at an MOI of 10 with media containing RB50, RB50Δ*rseAB*, the T3SS-deficient strain WD3, T3SS and ACT-deficient strain 8W1, or media alone for 4 hours. The average percent cytotoxicity of 4 wells in at least two independent experiments as measured by (LDH release from a well/LDH release from the positive control well) x 100 ± SE is shown.

Discussion

Extracytoplasmic function (ECF) sigma factors are important mechanisms by which bacteria relay information about external changes to the transcriptional machinery in order to respond [43, 194]. Because bacteria must constantly sense and respond to changing conditions, it is important to understand how these systems are regulated. In the case of the RpoE-like group of ECF sigma factors, the sigma factor is held at the inner membrane by a transmembrane anti-sigma factor and a periplasmic protein that contributes to inhibition [40, 75]. Once a stress is encountered, the anti-sigma factor is degraded, releasing the sigma factor to redirect transcription to its regulon and respond to the stress. We previously identified an RpoE-like sigma factor, SigE, in the respiratory pathogen *Bordetella bronchiseptica* (Chapter 2).

Downstream of *B. bronchiseptica sigE* are the genes *bb3751* (*rseA*) and *mucB* (*rseB*), which encode proteins with homology to the negative regulators of *E. coli* σ^E, RseA and RseB, or of *P. aeruginosa* AlgU, MucA and MucB. In this work, I found that *rseA* and *rseB* indeed encode negative regulators of *B. bronchiseptica* SigE activity, and that some, but not all, of the predicted interactions between *E. coli* σ^E and RseA may be conserved in the *B. bronchiseptica* SigE system (Fig. 3-2 and 3-3). In *E. coli,* RseA contributes more significantly to σ^E inhibition than RseB; however, I demonstrated that both $RseA_{Bb}$ and $RseB_{Bb}$ make substantial contributions to inhibition of SigE activity (Fig. 3-2). This may be more similar to the AlgU (σ^E) system *P. aeruginosa*, where both MucA (RseA) and MucB (RseB) contribute nearly equally to inhibiting AlgU activity [58, 59]. I also showed that the cytoplasmic domain of RseA is sufficient to inhibit SigE activity in vitro and in vivo, similar to other characterized σ^E systems.

B. bronchiseptica lacking *sigE* (RB50Δ*sigE*) is more sensitive than wild-type RB50 to temperature stress, ethanol, and treatment with SDS+EDTA and various cell wall-active antibiotics (Chapter 2). Here, I demonstrate that cells with high SigE activity because they lack the negative regulators *rseA* and *rseB* (RB50Δ*rseAB*) are more resistant to treatment with either SDS+EDTA or some antibiotics that target the peptidoglycan layer, and that a plasmid-encoded copy of *rseA* and *rseB* restored sensitivity to these treatments (Fig. 3-4B). However, RB50Δ*rseAB* is not more resistant to high temperature or to ethanol stress than RB50 (Fig. 3-4A and data not shown). An *rseB* mutant in *Vibrio vulnificans* is not more resistant to treatment with ethanol, hydrogen peroxide, or SDS than wild-type [101], and expression of a plasmid-encoded copy of *sigE* in *Burkholderia cenocepacia* actually results in greater sensitivity to heat stress [50]. These, together with our results, suggest that proper regulation of σ^E systems is required to appropriately respond to stress, and that constitutive, high *B. bronchiseptica* SigE activity does not confer resistance to all isolated stress conditions that require SigE for survival.

In *P. aeruginosa,* strains lacking the *rseA* homolog *mucA* have decreased expression of the T3SS and many other virulence factors, including some tissue-damaging proteases (LasA, elastase), endotoxin A, type IV pili, and flagella via downregulation of the cyclic AMP-dependent transcription factor Vfr. This leads to a decrease in virulence, and increased ability to establish chronic infection [114, 117]. In *Actinobacillus pleuropneumoniae*, while a Δ*rseA* strain was identified as having increased adherence, it was found to be attenuated for virulence [60, 134]. In *B. bronchiseptica*, a strain lacking *rseA* and *rseB* does not colonize the lungs and trachea of wild-type (C57BL/6) mice as efficiently as RB50, but persists in the nasal cavity similarly to RB50 (Fig. 3-6). The defect

in RB50Δ*rseAB* colonization of the lower respiratory tract begins early, and is exacerbated later in infection. Through our work examining the effect of high SigE activity in immunocompromised mouse models, we hypothesize a role for the SigE system in interactions with the innate immune system. RB50Δ*rseAB* is still defective in colonizing the lower respiratory tract of $Rag1^{-/-}$ mice, which lack an adaptive immune response (Fig. 3-6B). Interestingly, while RB50Δ*rseAB* was still defective in colonizing the lungs of $TLR4^{def}$, it could efficiently colonize the trachea of these mice (Fig. 3-7C). This suggests that the defect of RB50Δ*rseAB* in tracheal colonization may be due to high SigE activity interfering with some important interaction of *B. bronchiseptica* with the innate immune system.

RB50Δ*rseAB* failed to cause systemic, lethal infection in both $TLR4^{def}$ mice and $Rag1^{-/-}$ mice, suggesting that this strain either cannot leave the respiratory tract to colonize systemic organs, or cannot survive in the bloodstream. Ongoing work will differentiate between these two possibilities. RB50Δ*rseAB* is not more susceptible to one bactericidal component of the bloodstream, complement, but does cause less cytotoxicity to macrophages. *B. bronchiseptica* lacking either *sigE* or with high SigE activity is less cytotoxic to macrophages, but not to the same extent as strains lacking a functional T3SS or adenylate cyclase toxin (Fig. 3-8) [163]. This indicates that both high and no SigE activity affect the cytotoxicity of *B. bronchiseptica* to macrophages, either by directly regulating some factor important for cytotoxicity or by indirectly preventing a factor, such as the T3SS, from being properly expressed or inserted into the envelope. Further work is necessary to determine the specific role the SigE system has in interactions with macrophages.

σ^E is known to regulate the transcription of many cell envelope components in other bacteria, such as *E. coli*. Perhaps with SigE constitutively activated, the cell envelope protein composition has been altered to either mask or expose something differently from wild-type. Future studies determining the *B. bronchiseptica* SigE regulon could provide insight into the role this system plays in pathogenesis of this respiratory pathogen.

Materials and Methods

Strains and Media

A complete list of strains used in this study can be found in Table 3-2. *B. bronchiseptica* strains are derivatives of the previously described *B. bronchiseptica* strain RB50 [29]. *B. bronchiseptica* was maintained on Bordet-Gengou (BG) agar (Difco) containing 10% defibrinated sheep blood (Hema Resources) and 20 μg/ml streptomycin. In liquid culture, *B. bronchiseptica* was grown in Stainer-Scholte broth [30] with aeration. Tetracycline was used at 20 μg/mL, and kanamycin at 30 μg/mL as needed. The RB50Δ*rseAB* mutant was constructed as described below.

E. coli strains used to measure SigE activity are derivatives of MG1655 that carry the σ^E-dependent *rpoH*P3::*lacZ* reporter (strain SEA001 [31]). *E. coli* strain BL21(DE3) pLysS was used to express constructs for protein purification. *E. coli* were grown in LB broth in a gyratory water bath with aeration. Ampicillin was used at 100 μg/ml, and IPTG was used at 1 mM as needed for experiments with *E. coli*.

Table 3-3: Strains

	Strain	Genotype	Source, Reference
E. coli	SEA001	MG1655 ΦλrpoHP3::lacZ ΔlacX74	[82]
	SEA5036	BL21(DE3) slyD::kan pLysS pPER76	[92]
	XQZ001	BL21(DE3) slyD::kan pLysS pXQZ001	Chapter 2
	SEA5115	BL21(DE3) slyD::kan pLysS pSEB018	This work
	SEA5088	BL21(DE3) pLysS pSEB019	This work
	SEA4114	CAG43113 *rpoE::kan*	[71]
	SEA008	SEA001 pTrc99a	[71]
	SEA5005	SEA001 pSEB006	This work
	SEA5043	SEA001 pSEB007	This work
	SEA5044	SEA001 pSEB008	This work

	Strain	Genotype	Source, Reference
***E. coli* (cont.)**	SEA5103	SEA001 pSEB009	This work
	XQZ004	DH5α pXQZ004	This work
	SS1827	DH5α pSS1827	[171]
B. bronchiseptica	RB50	RB50	[167]
	SEA5516	RB50Δ*sigE*	(Chapter 2)
	SEA5517	RB50Δ*rseArseB*	This work
	RB50Δwbm	RB50Δ*wbmBwbmCwbmD*	[172]
	WD3	RB50Δ*bscN*	[163]
	AVS	RB50Δ*cyaA*Δ*bscN*	[149, 173]
	SEA	RB50 pMP220	This work, [192]
	SEA	RB50Δ*sigE* pMP220	This work
	SEA	RB50Δ*rseArseB* pMP220	This work
	SEA	RB50 pFam	This work
	SEA	RB50Δ*sigE* pFam	This work
	SEA	RB50Δ*rseArseB* pFam	This work
	SEA	RB50Δ*rseArseB* pBBR1BAD	This work
	SEA	RB50Δ*rseArseB* pRseAB	This work

Plasmid constructions

All plasmids used in this study are listed in Table 3-2 and oligonucleotide sequences are given in Table 3-3. Plasmids pSEB007, pSEB008, and pSEB009 were constructed to express various components of the *B. bronchiseptica* SigE system in *E. coli*. The *sigE* and *rseA* genes were amplified together from RB50 genomic DNA with the primers SigEF and RseAR; *sigE*, *rseA*, and *rseB* were amplified together using the primers SigEF and RseBR; *sigE* and $rseA_{Bb}$-*cyto* were amplified using the primers SigEF and RseAcytoR. These constructs were then ligated into the expression vector pTrc99a under the control of the IPTG-inducible trc promoter.

To construct the *fam*P::*lacZ* fusion in *B. bronchiseptica*, the SigE-dependent promoter was encoded on the oligos PFamLacZF and PFamLacZR. These oligos were mixed at a concentration of 1 μg, placed at 80 °C and let to slowly cool to room temperature. These were then ligated into the EcoRI and PstI sites of pMP220 to create the

plasmid pFam. pFam was moved into RB50, RB50Δ*sigE*, and RB50Δ*rseArseB* through tri-parental mating on BG agar with $MgCl_2$. Transconjugants were selected on BG containing 20 μg/ml streptomycin and 20 μg/ml tetracycline.

Table 3-4: Plasmids

Plasmid name	Genotype	Source, Reference
pTrc99a	Vector, pBR322 ori, Ap^R	Pharmacia
pSEB006	*sigE* in pTrc99a	(Chapter 2)
pSEB007	*sigE* and $rseA_{Bb}$ in pTrc99a	This work
pSEB008	*sigE,* $rseA_{Bb}$ and $rseB_{Bb}$ in pTrc99a	This work
pSEB009	*sigE* and $rseA_{Bb}$-*cyto* in pTrc99a	This work
pSEB015	isolated *rpoH*P3 promoter in pRLG770, Ap^R	[92]
pPER76	*rpoE* in T7 expression vector pET15b, Kan^R	[70]
pXQZ001	*sigE* in T7 expression vector pET15b, Kan^R	(Chapter 2)
pSEB017	$rseA_{Ec}$-*cyto-strep* in vector pASKIBA3plus	This work
pSEB018	$rseA_{Ec}$-*cyto-strep* in T7 expression vector pET24b	This work
pSEB019	$rseA_{Bb}$-*cyto-His* in T7 expression vector pET21a	This work
pSS1827	helper plasmid competent for mating, Ap^R	[171]
pSS4245	Bordetella-specific allelic exchange vector, Kan^R	Stibitz
pXQZ004	Δ*rseArseB* allele in pSS4245	This work
pMP220	Promoterless *lacZ*, Tet^R	[192]
pFam	SigE-dependent *fam* promoter in pMP220, Tet^R	This work
pBBR-MCS1	Broad host-range vector, Kan^R	[195]
pBAD18	Vector, pBR322 ori, $Chlor^R$	[196]
pBBR1BAD	AraC and P_{BAD} in pBBR1-MCS1, Kan^R	This work
pRseAB	*rseA* and *rseB* in pBBR1BAD	This work

Construction of RB50Δ*rseAB*

rseA and *rseB* were deleted from RB50 using a *Bordetella*-specific allelic exchange procedure to produce strain SEA5517. The sequences of the primers used in the construction are listed in Table 3-3. A PCR product containing 636 bp proximal to the 5' end of *rseA* was amplified from RB50 genomic DNA using primers RseABKO_LeftF and RseABKO_LeftR. A non-overlapping PCR product containing 767 bp proximal to the 3' end of *rseB* was amplified with primers RseABKO_RightF and RseABKO_RightR. The two fragments were digested with BamHI and ligated. The resulting construct was

amplified with primers SigEKO_LeftF and SigEKO_RightR, cloned into the TopoTA vector (Invitrogen), and verified by sequencing to give plasmid pXQZ004. In this deletion construct, the 1544 bp central region of *rseA* and *rseB* is deleted, leaving 6 bp at the 5' end of *rseA* and 9 bp at the 3' end of *rseB*. The deletion construct from pXQZ004 was then cloned into the EcoRI site of the allelic exchange vector pSS4245 to generate pXQZ005 and transformed into *E. coli* strain DH5α. Tri-parental mating with wild-type *B. bronchiseptica* strain RB50, *E. coli* strain DH5α harboring the pXQZ005 vector (strain XQZ005), and DH5α harboring the helper plasmid pSS1827 (strain SS1827) [34,35] and selection of mutants were performed as previously described [34]. The deletion strain was verified by PCR using primers RseABKO_LeftF and RseABKO_RightR and by Southern blot analysis.

Table 3-5: Primer Sequences

Primer name	Sequence (5' to 3')	Source, Reference
SigEF	GGCGGAGAATTCAGGAGGAGGCGTCATG AGCGAACGCGATG	Chapter 2
RseAF	GAATTCAGGAGTGGCAGTCATGCAAACC	This work
RseAR	GGATCCTCAACGTCCTGCTCCGGCGTCG	This work
RseBR	GGATCCTTATTGTTGTTTCGCCG	This work
RseAcytoR	GAATTCAGGAGGATCCCGTCAATGAGTA	This work
RseAEcStrepF	ATGGTAGGTCTCAAATGCAGAAAGAACA ACTTTCCGCTTTAA	This work
RseAEcStrepR	ATGGTAGGTCTCAGCGCTCTGTGCCGCCC ACGGACGTA	This work
RseABbHisF	CATATGATGCAAACCGCAGCCAAGTCC	This work
RseABbHisR	CTCGAGCTCGGCCTCCAGCGCGCGGGC	This work
PFamLacZF	AATTCCGGAACTTTGGCGCGGCGCAGCA	This work
PFamLacZR	GGCCTTGAAACCGCGCCGCGTCGTCAGC	This work
RseABKO LeftF	AGTGAATTCCCCCTTCTTCAGCGTCTTG	This work
RseABKO LeftR	GATGGATCCCAACAATAAGCCCGCAAAG	This work
RseABKO RightF	ATTGGATCCTTGCATGACTGCCACTC	This work
RseABKO RightR	GGAGAATTCGGGTCTCGGGTTACAAATA	This work
RseABcompF	GAATTCAGGAGGAGGCGTCATGAGCGAA	This work
RseABcompR	AAGCTTTTACCAGCGACGCTCGGCAT	This work

β-galactosidase assays

For assays with the *rpoH*P3::*lacZ* reporter in *E. coli,* cultures were grown to an OD_{600} of 0.1-0.2, then IPTG was added to 1mM final concentration. Samples were taken 2.5 hours after addition of IPTG, and β-galactosidase activity was determined by standard assay [177]. For assays with the *fam*P::*lacZ* fusion in *B. bronchiseptica*, cultures started from a single resuspended colony were grown for 14-16 hours at 37 °C, and β-galactosidase activity was measured [177]. A minimum of 3 experiments was performed for each strain.

Protein purification

His-SigE and His-σ^E were purified from strains as previous described (Chapter 2, [92]). To facilitate purification of $RseA_{Ec}$-cyto, the plasmid pSEB018 was constructed by amplifying the first 300 nucleotides of $RseA_{Ec}$ from MG1655 genomic DNA using the primers RseAEcStrepF and RseAEcStrepR. The resulting PCR product was first cloned into pASKIBA3plus (IBA), which adds a C-terminal Strep tag to recombinant proteins. Due to problems with constitutive expression of this construct, which is toxic in *E. coli*, we then subcloned $RseA_{Ec}$-cyto-Strep into pET24b to create plasmid pSEB018. C-terminally Strep-tagged $RseA_{Ec}$-cyto was purified from strain BL21(DE3) *slyD*::*kan* p*LysS* pSEB001 as follows. Briefly, cells were grown at 30°C to an OD_{600} of 0.3, at which point IPTG was added to induce protein production. Following 2 hours of induction, cells were harvested by centrifugation and resuspended in lysis buffer (100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol). Cells were then lysed by sonication and the lysate was cleared by centrifugation. The supernatant containing soluble $RseA_{Ec}$-cyto-Strep was loaded onto a StrepTactin column (IBA). Bound proteins were eluted with 20

mM Tris-HCl pH 8.0, 500 mM NaCl, and 2.5 mM β-mercaptoethanol. Fractions containing $RseA_{Ec}$-cyto-Strep were pooled and dialyzed into 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 15% glycerol and 2.5 mM β-mercaptoethanol.

For purification of $RseA_{Bb}$-cyto, the first 231 nucleotides of *rseA* were amplified from RB50 genomic DNA with the primers RseABbHisF and RseABbHisR, and the resulting PCR product was ligated into pET24b (Novagen). C-terminally His-tagged $RseA_{Bb}$-cyto was purified from strain BL21(DE3) *slyD*::*kan* p*LysS* pSEB as follows. Briefly, cells were grown at 30°C to an OD_{600} of 0.3, at which point IPTG was added to induce protein production. Following 2 hours of induction, cells were harvested by centrifugation and resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 2.5 mM β-mercaptoethanol, 1 mM PMSF). Cells were then lysed by sonication and the lysate was cleared by centrifugation. The supernatant containing soluble $RseA_{Bb}$-cyto-His was loaded onto a Ni-NTA column (Qiagen). Bound proteins were eluted with a stepwise gradient of 20, 60, 100, and 200 mM imidazole in column buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2.5 mM β-mercaptoethanol). Fractions containing $RseA_{Bb}$-cyto-His were pooled and dialyzed into 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 15% glycerol and 2.5 mM β-mercaptoethanol.

In vitro transcription

80 nM His-$SigE_{Bb}$, His-σ^{E}_{Ec}, or His-σ^{70}_{Ec} (gift from K. Murakami) was incubated with either 800 nM $RseA_{Bb}$-cyto-His, $RseA_{Ec}$-cyto-strep, or transcription buffer (40 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT, 0.1 mg ml^{-1} BSA) at 30 °C for 15 minutes. *E. coli* core RNA polymerase (Epicentre) was added, and the mixture incubated at 30 °C for 10 additional minutes. Transcription was initiated by addition of

ESigEBb, $E\sigma^{E}_{Ec}$, or $E\sigma^{70}_{Ec}$ with or without the appropriate anti-sigma factor, to a final concentration of 40 nM sigma, 10 nM core RNA polymerase, and 400 nM RseA-cyto, to prewarmed transcription mix containing 5.0 nM supercoiled plasmid template pSEB015 [92], 5% glycerol, 200 μM ATP, 200 μM CTP, 200 μM GTP, 10 μM UTP, and 2.5 μCi [α-^{32}P]UTP in transcription buffer. After 10 minutes at 30 °C, reactions were stopped by the addition of stop solution (80% formamide, 20 mM EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue). Samples were electrophoresed on 6% polyacrylamide gels containing 7.5 M urea, and visualized by phosphorimaging.

Temperature stress

For heat shock and thermotolerance experiments, mid-log phase cultures of RB50 and RB50Δ*rseAB* were diluted to an OD_{600} of 0.01 in fresh Stainer-Scholte broth and incubated at 37 °C in a gyratory water bath with shaking. At an OD_{600} of 0.1, cultures were either shifted to 40 °C for adaptation, or kept at 37 °C. After 90 minutes, all cultures were shifted to 50 °C, and survival was measured by plating and CFU counts.

Disk diffusion assays

Disk diffusion assays were performed as previous described (Chapter 2). Briefly, *B. bronchiseptica* cultures in mid-log phase were diluted to 6×10^{8} CFU/ml and spread on Stainer-Scholte agar plates to generate a lawn of bacteria. Disks containing 300 iu polymyxin B, 10 mg ampicillin, 100 mg mecillinam, 750 mg sodium dodecyl sulfate (SDS) and 2.9 mg EDTA, 30 μg aztreonam, 10 μg imipenem, 10 μg meropenem, 30 μg chloramphenicol, 15 μg erythromycin, 30 μg kanamycin, 150 μg rifampicin, or 30 μg tetracycline were applied to the plates and the zones of inhibition were measured after overnight incubation at 37 °C.

Cytotoxicity assay

Cytotoxicity assays were performed as previously described (38). Briefly, RAW 264.7 murine macrophage cells were cultured in Dulbecco modified Eagle medium (DMEM) (HyClone) with 10% fetal bovine serum (FBS) (HyClone). Cells were grown to 80% confluence in 96-well plates at 37 °C in 5% CO2 and then washed with RPMI (Mediatech) containing 5% FBS. Bacteria were added in 100 µl RPMI at a multiplicity of infection (MOI) of 10. After a 5-minute centrifugation at 250xg, the RAW 264.7 cells and bacteria were incubated for four hours. Percent lactate dehydrogenase (LDH) release, a measure of cytotoxicity, was determined using the Cytotox96 Kit (Promega) according to the manufacturer's protocol.

Complement killing assay

Complement killing assays were performed as previously described [178]. Briefly, blood freshly collected from C57BL/6 mice was pooled, incubated at 4°C for 1 hour and centrifuged at 2000g for 10min. The serum was collected and diluted 1:5 in PBS. Complement-inactive serum was prepared by heating an aliquot of the diluted sample at 56°C for 30 min. From mid-log-phase cultures, approximately 500 CFU of RB50, RB50*ΔrseArseB* and RB50*Δwbm* in 5 µl of PBS were incubated with 45 µl of diluted serum or PBS for 1 hour at 37°C. Bacterial numbers before and after incubation were determined by plating and CFU counts.

Animal experiments

C57BL/6, TNF-$\alpha^{-/-}$, Rag1$^{-/-}$, C3H/HEN and C3H/HEJ mice were obtained from Jackson laboratories (Bar Harbor, Maine, USA). All mice were bred in our *Bordetella*-free, specific pathogen-free breeding rooms at The Pennsylvania State University. For

inoculation, mice were sedated with 5% isoflurane (Abbott laboratory) in oxygen and 50μL of phosphate buffered saline (PBS) containing 1×10^5 or 5×10^5 CFU of the indicated bacteria was pipeted onto the external nares [144]. This method reliably distributes the bacteria throughout the respiratory tract [181]. Survival curves were generated by inoculating C3H/HEJ, TNF-$\alpha^{-/-}$ and Rag1$^{-/-}$ mice with either RB50 or RB50Δ*rseArseB*. Mice suffering from lethal bordetellosis as determined by severe hunched posture, ruffled fur, extremely labored breathing and apathy were euthanized to prevent unnecessary suffering [151]. For quantifying bacterial load, mice were euthanized via CO_2 inhalation and lung, trachea, nasal cavity, spleen, liver and/or kidneys were excised. Tissues were homogenized in PBS, aliquots were serially diluted, plated, incubated at 37°C for 2 to 3 days, and CFUs were determined. All protocols were reviewed by the university IACUC and all animals were handled in accordance with institutional guidelines.

Lung pathology

For analysis of lung pathology, mice were intranasally inoculated and euthanized as described above. The tracheas and lungs were excised and inflated with approximately 2 mL of 10% formaldehyde. The tissues were then sectioned and stained with haemolysin and eosin (H&E) at the Animal Diagnostic Laboratories Facility of The Pennsylvania State University. Sections were examined and scored by a veterinarian with training and experience in rodent pathology (M.J.K.) who was blinded to experimental treatment (25, 37). A score of 0 indicates no noticeable inflammation or lesions; a score of 1 indicates few or scattered foci affecting <10% of the tissue; a score of 2 indicates frequent mild perivascular and/or peribronchial lymphoid aggregates, with overall inflammation affecting no more than 10 to 20% of the tissue; a score of 3 indicates moderate lesions,

typically with abundant perivascular and peribronchial lymphoid infiltrates, with inflammation affecting ~20-30% of the tissue; and a score of 4 indicates extensive pneumonia and marked inflammation affecting >30% of the tissue; a score of 5 indicates extensive lesions with >50% of the tissue affected. If a severity falls between categories, 0.5 was added to the pathology score of the lower category.

Statistical analysis

The average of at least two independent experiments is shown ± the standard error of the geometric mean was determined and expressed as error bars, when appropriate. Two-tailed, unpaired, Student's T-tests were used to determine statistical significance between groups.

Acknowledgements

We would like to thank Mary Kennett for assistance with the pathology scoring, Sandra Armstrong for providing pMP220, Katsuhiko Murakami for providing *E. coli* σ^{70}, Scott Stibitz for providing the plasmids pSS1827 and pSS4245, and Laura Weyrich for help constructing pBBR1BAD.

Chapter 4

Identification of genes in the SigE regulon

Abstract

The *Bordetella bronchiseptica* extracytoplasmic function (ECF) sigma factor, SigE, is required for survival during heat and ethanol stress, for resistance to the detergent SDS and cell wall-active antibiotics, for killing or surviving within phagocytes, and for virulence. Previous work has shown that both high SigE and no SigE activity are detrimental to cytotoxicity, survival during heat stress, and virulence. Once activated by an inducing signal, ECF sigma factors bind to core RNA polymerase and direct transcription to a subset of genes that allow the cell to cope with changes in the environment. The *E. coli* σ^E regulon consists of genes involved in assembly of outer membrane proteins (OMPs) and lipopolysaccharides, as well as small RNAs that target OMPs and other proteins for degradation to help minimize the effects of envelope stress. In this chapter, multiple methods are employed to identify candidate SigE-regulated genes in *B. bronchiseptica*, many of which are predicted to encode cell envelope proteins, proteins involved in envelope biogenesis, combating stress, and virulence. *B. bronchiseptica* with high SigE activity has an altered envelope composition, and the SigE system may also be important for regulating some virulence factors in *B. bronchiseptica*, separate from the BvgAS two-component system believed to regulate expression of most virulence factors in this respiratory pathogen.

*Sara Hester performed the cytotoxicity assays and qRT-PCR, and Tracy Nicholson (USDA-ARS) performed the microarray experiments

Introduction

Sigma factors are the dissociable subunits of core RNA polymerase (RNAP) responsible for conferring promoter specificity. Bacteria have a housekeeping sigma factor (σ^{70}) responsible for the bulk of transcription in the cell, and most bacteria also have an array of alternative sigma factors that redirect transcription to a subset of promoters under specific conditions, such as starvation or cell envelope stress [33]. ECF sigma factors are the largest and most diverse family of alternative sigma factors [39, 40]. Recent work has categorized ECF sigma factors into different groups based on sequence similarity and gene context [40]. The RpoE-like group of sigma factors (ECF02) includes σ^{E} from *Escherichia coli*, AlgU from *Pseudomonas aeruginosa*, and SigE from the respiratory pathogen *Bordetella bronchiseptica*. We recently described a role for *B. bronchiseptica* SigE in cell envelope stress response and virulence (Chapter 2 and Chapter 3).

B. bronchiseptica, *Bordetella pertussis*, and *Bordetella parapertussis* are three of the nine known members of the *Bordetella* genus, which are Gram-negative β-proteobacteria and important respiratory pathogens [135, 139]. *B. pertussis* and *B. parapertussis* are the causative agents of whooping cough in humans, and are thought to have independently evolved from a *B. bronchiseptica*-like ancestor through significant gene loss [135, 136]. *B. bronchiseptica* infects a variety of mammalian hosts, and causes a wide range of disease in these hosts, from asymptomatic carriage to fatal pneumonia, including kennel cough in dogs, atrophic rhinitis in pigs, and snuffles in rabbits [136, 138, 156]. Though not typically associated with human disease, *B. bronchiseptica* has been isolated from immunocompromised patients [137]. It is important, therefore, to understand how these respiratory pathogens sense and adapt to environmental changes in a host.

In previous work, I have shown that the SigE system of *Bordetella bronchiseptica* is important for responding to stress conditions, such as heat shock, ethanol stress, and specific cell envelope stresses (Chapter 2, Chapter 3). Cells lacking *sigE* (RB50Δ*sigE*) are more sensitive, and cells with high SigE activity, lacking the negative regulators *rseA* and *rseB* (RB50Δ*rseAB*), are more resistant to many of these stresses. The SigE system is also involved in important interactions with the host immune system. RB50Δ*sigE* no longer causes lethal, systemic infection in mice lacking an adaptive immune response; RB50Δ*rseAB* is defective in colonizing the lower respiratory tract of immunocompetent mice, and is also no longer able to cause lethal infection in immunocompromised mice (Chapter 2, Chapter 3).

As a sigma factor, SigE presumably directs transcription of many genes that could contribute to the phenotypes described above. Much work has been done in *E. coli* to identify members of the orthologous σ^E regulon. Rhodius, et al. identified nearly 100 genes as σ^E-regulated in *E. coli*, then used this information to predict regulon members in closely related species, such as *Salmonella enterica*. From this, the authors compiled a "core regulon" comprised of genes predominantly involved in the elaboration of cell envelope structures, such as outer membrane proteins (OMPs) and lipopolysaccharides (LPS), and also described an "extended regulon," the unique genes transcribed by RpoE-like sigma factors in different bacterial species [48]. In this section, I present preliminary data identifying candidate members of the *B. bronchiseptica* SigE regulon through gene expression analysis, in vitro transcription, and bioinformatic prediction, and examine the role this regulon plays in both maintaining the cell envelope and in pathogenesis.

Results and Discussion

Identifying candidate SigE regulon members

To identify candidate genes regulated by SigE, we used three methods: 1) comparing gene expression in cells with high SigE activity (RB50Δ*rseAB*) to wild-type RB50, 2) directly determining genes transcribed by SigE in vitro using run-off transcription coupled to microarray analysis (ROMA), and 3) predicting SigE-dependent promoters in the *B. bronchiseptica* genome using a bioinformatic approach. The results from each of these methods are outlined below.

Gene expression analysis

We compared gene expression in RB50Δ*rseAB*, which I previously showed to have high SigE activity (Chapter 3), to that of wild-type RB50. Genes with higher expression in RB50Δ*rseAB* compared to RB50 may be directly or indirectly regulated by SigE, and are candidate regulon members. Both strains were grown at 37 °C to mid-exponential phase, then cells were harvested and total RNA was extracted, made into cDNA, labeled, and used in microarray analysis, as described in the Materials and Methods. Genes with at least two-fold higher expression in RB50Δ*rseAB* compared to RB50 in a minimum of three out of the four replicates (three biological replicates and one technical replicate/dyeswap) were selected for further analysis. The results are shown in Tables 4-1 and 4-2.

Table 4-1: Genes with increased expression in RB50Δ*rseAB* compared to RB50
* denotes genes also identified by ROMA (Table 4-3); [P] denotes genes with predicted SigE-dependent promoters (Table 4-4).

Gene	Name	Product	Avg. Fold Change
BB0098		put. exported protein	20.32
BB4490		outer membrane porin	10.53
BB1931[P]		put. exported protein	8.50
BB0419[P]	*sphB1*	autotransporter subtilisin-like protease	8.30
BB4284[P]		put. membrane protein	8.25
BB3264[P]		put. exported protein	7.88
BB3826	*bfrD*	prob. TonB-dependent receptor	7.75
BB0783		put. membrane protein	7.50
BB3068[P]		put. exported protein	7.44
BB1936	*fhaL*	adhesin	6.89
BB3993	*ompQ*	outer membrane porin protein OmpQ	6.85
BB1291		conserved hyp. protein	6.56
BB4265[P]		put. exported protein	6.42
BB3108[P]		hyp. protein	5.43
BB2245		put. exported protein	5.35
BB4652		hyp. protein	5.29
BB1292		put. exported protein	5.23
BB0975		prob. hydrolase	4.99
BB2033	*bapC*	put. autotransporter	4.94
BB3424[P]		fimbrial protein	4.93
BB3766		conserved hyp. protein	4.89
BB0501	*htpG*	heat shock protein	4.57
BB1932	*bph3*	put. DNA-binding protein (histone)	4.46
BB4017	*osmB*	osmotically inducible lipoprotein B	4.34
BB4835*[P]	*fam*	RNA polymerase sigma-32 factor	4.19
BB1864	*vag8*	autotransporter	4.16
BB1289[P]		put. integral membrane protein	4.14
BB1368[P]		put. membrane protein	4.05
BB1894[P]		heat shock protease	3.96
BB4518*[P]		put. thioredoxin	3.64
BB1610	*bscE*	hyp. protein	3.64
BB1609	*bscF*	put. type III secretion protein	3.63
BB3765		put. membrane protein	3.59
BB1956		hyp. protein	3.57
BB4939		put. exported protein	3.55
BB2667[P]		put. universal stress protein	3.49
BB0452		autotransporter	3.48

Gene	Name	Product	Avg. Fold Change
BB4940[P]	*dsbA*	thiol:disulfide interchange protein	3.46
BB1733[P]	*ppiB*	peptidyl-prolyl cis-trans isomerase B	3.41
BB2885		put. exported protein	3.35
BB1933[P]		put. hydrolase	3.34
BB3097	*mrcA*	penicillin-binding protein 1A	3.30
BB3505		phage-related conserved hyp. protein	3.23
BB4491[P]		put. exported protein	3.18
BB3025[P]	*trxC*	thioredoxin 2	3.16
BB3846*		put. membrane protein	3.13
BB2228*[P]		phage-related conserved hyp. protein	3.12
BB4905		conserved hyp. protein	3.11
BB1617	*bsp22*	put. secreted protein	3.08
BB3802		BolA-like protein	3.06
BB1143		conserved hyp. protein	3.00
BB3493[P]		phage-related conserved hyp. protein	2.98
BB0611		acyl-CoA dehydrogenase	2.98
BB3749*[P]	*mucD*	serine protease	2.96
BB3932[P]		put. zinc protease	2.91
BB3801[P]		put. intracellular septation protein	2.87
BB1629	*bscO*	put. type III secretion protein	2.80
BB4506[P]	*rpoN*	prob. σ^{54}-modulation protein	2.74
BB3496		phage-related conserved hyp. protein	2.71
BB4492		prob. LysR-family transcriptional reg.	2.71
BB1359	*secF*	protein-export membrane protein	2.66
BB2528	*acrB*	acriflavine resistance protein B	2.64
BB3491		phage-related conserved hyp. protein	2.60
BB2233		phage-related conserved hyp. protein	2.58
BB1709[P]		phage-related hyp. protein	2.57
BB1360[P]	*secD*	protein-export membrane protein	2.53
BB1620	*bopD*	put. outer protein D	2.52
BB2213		phage-related put. DNA binding protein	2.49
BB4018		conserved hyp. protein	2.45
BB2232[P]		phage-related conserved hyp. protein	2.44
BB3678[P]		conserved hyp. protein	2.41
BB3241		put. exported protein	2.40
BB0168*		put. penicillin-binding protein	2.39
BB4235		put. TolR-like translocation protein	2.39
BB3739[P]		put. integral membrane protein	2.29
BB2214		put. phage terminase	2.25
BB3419		globin-like protein	2.09

Table 4-2: Genes with decreased expression in RB50Δ*rseAB* compared to RB50

Gene	Name	Product	Avg. Fold Change
BB2983		put. membrane protein	-12.73
BB3750	*rseB*	put. sigma factor regulatory protein	-11.66
BB2982		put. exported protein	-11.51
BB2985		put. exported protein	-8.93
BB3751	*rseA*	put. membrane protein	-8.56
BB2981		conserved hyp. protein	-7.04
BB3795		branched-chain amino acid permease	-6.82
BB2308		put. exported protein	-6.37
BB1658	*fim3*	serotype 3 fimbrial subunit precursor	-5.84
BB2978	*oplaH*	5-oxoprolinase	-5.69
BB2309		put. membrane permease	-5.48
BB2979		conserved hyp. protein	-5.27
BB2907	*livM*	high-affinity branched-chain amino acid permease	-5.08
BB3794		put. branched-chain amino acid transport sys. prot.	-4.78
BB2908	*livH*	high-affinity branched-chain amino acid permease	-4.46
BB2986		put. membrane protein	-4.42
BB2905	*livF*	high-affinity branched-chain amino acid transport, ATP-binding protein	-4.17
BB2987		conserved hyp. protein	-4.10
BB3447	*cysN*	sulfate adenylyltransferase subunit 1	-3.70
BB4954		put. exported protein	-3.66
BB3355		put. transmembrane component of ABC transporter	-3.65
BB3674	*fim2*	serotype 2 fimbrial subunit precursor	-3.55
BB3448	*cysD*	sulfate adenylyltransferase subunit 2	-3.53
BB0377		put. membrane protein	-3.49
BB2906	*livG*	high-affinity branched-chain amino acid transport, ATP-binding protein	-3.49
BB4831		put. bifunctional protein	-3.48
BB2002		put. membrane protein	-3.32
BB4830	*ctaD*	cytochrome c oxidase polypeptide I	-3.23
BB2024	*putA*	bifunctional prolineoxidoreductase	-3.23
BB2909	*livJ*	leu/ile/val-binding protein precursor	-3.13
BB3977		put. exported protein	-3.10
BB4943		metallo-beta-lactamase family protein	-3.06
BB3218	*ectC*	L-ectoine synthase	-3.03
BB1747		put. exported protein	-2.78
BB3217		put. L-proline 4-hydroxylase	-2.73
BB4610	*atpE*	ATP synthase c chain	-2.70

Gene	Name	Product	Avg. Fold Change
BB0209		put. ATP-dependent RNA helicase	-2.69
BB4746		conserved hyp. protein	-2.67
BB1702		phage-related conserved hyp. protein	-2.66
BB3002		put. transport protein	-2.63
BB2980		put. LysR-family transcriptional regulator	-2.60
BB1696		phage-related hyp. protein	-2.60
BB2303	*bipA*	put. outer membrane ligand binding protein	-2.49
BB2307	*greB*	transcription elongation factor	-2.49
BB3764		put. exported hydrolase	-2.41
BB3187	*rimM*	16S rRNA processing protein	-2.40
BB4731		phosphoglycerate dehydrogenase	-2.23
BB1431		conserved hyp. protein	-2.19

Over half of the 77 genes identified as having increased expression in RB50Δ*rseAB* relative to RB50 encode envelope proteins or predicted exported proteins. This suggests that the SigE regulon may play an important role at the cell envelope of *B. bronchiseptica.* Also included among genes with increased expression when SigE activity are genes putatively responsive to heat shock: *fam*, the heat shock sigma factor σ^{32}; an ortholog of the chaperone *htpG*, a protease, *bb1894*; and an ortholog of the *E. coli* periplasmic protease *degP* [197], *mucD*, encoded in the same operon as *sigE* in *B. bronchiseptica.* I have previously shown that cells lacking SigE are more sensitive to heat shock, and that SigE can direct transcription from the promoter region of *fam* both in vitro and in *B. bronchiseptica* (Chapters 2 and 3).

Many genes encoding virulence factors have increased expression in RB50Δ*rseAB*, including at least three predicted autotransporters, an adhesin and a predicted fimbrial protein, along with some components of the type three secretion system (T3SS) apparatus. However, the entire T3SS locus is not differentially expressed when SigE activity is high. A functional T3SS is required for full cytotoxicity [163], and we have previously

demonstrated that *B. bronchiseptica* either lacking *sigE* or with high SigE activity (RB50Δ*rseAB*) have decreased cytotoxicity to macrophages, though both strains exhibit higher cytotoxicity than RB50 lacking a functional T3SS (RB50Δ*bscN*).

To validate some of the expression changes seen in the microarray analysis, we performed quantitative RT-PCR (qRT-PCR) on RNA isolated from either RB50 or RB50Δ*rseAB* to analyze expression of *fam*, as well as various genes encoded in the T3SS locus: *bscL, bscN* (ATPase)*, bsp22* (secreted protein)*, bopN,* and *bopD* (both involved in translocation) [163, 198]. *bopD* and *bsp22* have higher expression in RB50Δ*rseAB* in our microarray analysis, whereas the other genes had less than a two-fold change in expression relative to RB50. In the qRT-PCR analysis, *fam* did have significantly increased expression in RB50Δ*rseAB* compared to RB50, validating its increased expression in our microarray analysis. However, analysis of the various genes in the T3SS locus showed transcript levels in RB50Δ*rseAB* roughly similar to those in RB50, with perhaps the exception of *bscN,* which encodes the ATPase that provides energy to the secretion system (S. Hester, preliminary results) [163]. Further validation of the expression levels of these and other genes identified in the microarray is necessary to determine the role that SigE may play in their regulation.

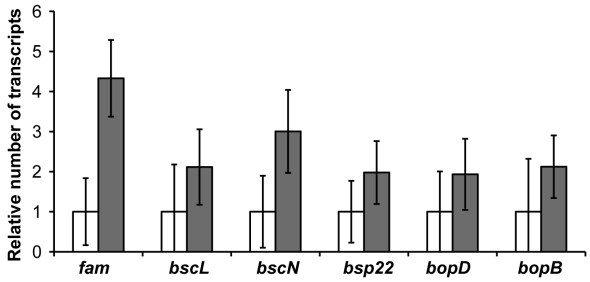


Fig. 4-1. Preliminary qRT-PCR results for *fam* and T3SS locus genes. Expression of *fam* is increased in RB50Δ*rseAB* (grey bars) compared to RB50 (white bars). Average of three independent reactions ± SD is shown.

Run-off transcription coupled to microarray analysis

I previously purified His-SigE and demonstrated that it can bind core RNA polymerase (RNAP) and direct transcription from at least one *E. coli* σ^E-dependent promoter, *rpoH*P3, and one *B. bronchiseptica* SigE-dependent promoter, *fam*, in vitro (Chapter 2). To identify additional candidate regulon members, I used a technique developed by the Helmann group, run-off transcription coupled to microarray analysis, or ROMA [199, 200]. In this technique, in vitro transcription reactions were performed using holoenzyme formed with SigE (ESigE) and sheared *B. bronchiseptica* genomic DNA as a template. Any RNA isolated from these reactions was dependent on ESigE for transcription. cDNA generated from this RNA was then labeled and hybridized to microarray chips, using the cDNA made from reactions with core RNAP alone as a control. Regions of the genome transcribed by ESigE show higher signal than core RNAP alone. I isolated RNA from in vitro transcription reactions where 1.0 μM or 0.1 μM His-SigE (both are saturating concentrations of SigE) was incubated with 0.05 μM *E. coli* core

RNAP to form holoenzyme, and this was used for subsequent microarray analysis. The results are shown in Table 4-3.

Table 4-3: Genes identified as transcribed by SigE in vitro by ROMA. Fold-change denotes the relative amount of that region of the genome in reactions with SigE vs. reactions with core RNAP alone. The genes listed comprise two independent experiments, and average fold-change from three replicates is shown. Genes for which a promoter has also been predicted are indicated. Colored lines at the side indicate genes encoded in a common operon. * Indicated genes with increased expression in RB50Δ*rseAB* (Table 4-1)

Gene	Name	Product	Fold-change above RNAP	Predicted Promoter
BB0061		thiol:disulfide interchange protein	2.51	no
BB0062		put. membrane protein	2.17	no
BB0168*		put. penicillin-binding protein	4.67	no
BB0169		prob. class IV aminotransferase	2.25	no
BB0171	*lipB*	lipoate-protein ligase B	2.11	no
BB0247	*def*	polypeptide deformylase	2.58	yes
BB0270		conserved hyp. protein	2.02	no
BB0355		conserved hyp. protein	2.54	no
BB0614		put. thiolase	2.48	no
BB0618		LysR family reg. protein	2.19	no
BB0620	*eftB*	electron transfer flavoprotein	2.01	no
BB0628		conserved hyp. protein	2.46	no
BB0737	*hpaD*	homoprotocatechuate 2,3-dioxygenase	2.34	no
BB1244	*ttrC*	tetrathionate reductase subunit C	2.04	no
BB1281		two component system sensor kinase	2.28	no
BB1282		two component system response reg.	2.88	yes
BB1283	*cyoA*	put. ubiquinol oxidase polypeptide II	2.43	no
BB1284	*cyoB*	ubiquinol oxidase polypeptide I	2.18	no
BB1286	*cyoD*	cytochrome 0 ubiquinol oxidase	2.01	no
BB1616	*bopN*	put. outer protein N	2.31	no
BB1844	*acnA*	put. aconitate hydratase	3.61	yes
BB2177		prob. fatty acid desaturase	3.03	yes
BB2222		phage-related hyp. protein	2.34	no
BB2228*		phage-related cons. hyp. prot.	2.07	yes
BB2462		put. lipoprotein	2.13	no
BB2677*	*ldcA*	put. muramoyltetrapeptide carboxypeptidase	2.01	no
BB2734		put. exported protein	2.75	no
BB2762		conserved hyp. protein	3.51	no

Gene	Name	Product	Fold-change above RNAP	Predicted Promoter
BB3051		put. exported protein	3.39	yes
BB3170	*hfq*	put. RNA-binding reg. protein	5.65	yes
BB3677	*gltA*	citrate synthase	5.55	no
BB3678		conserved hyp. protein	3.68	yes
BB3748	*lepA*	GTP-binding protein	3.23	no
BB3749*	*mucD*	serine protease	2.76	yes
BB3750	*rseB*	put. sigma factor reg. protein	3.56	no
BB3751	*rseA*	put. membrane protein	4.48	yes
BB3846*		put. membrane protein	2.36	no
BB3880		put. lipoprotein	2.83	no
BB4217	*lgt*	pro-lipoprotein diacylglyceryl transferase	2.06	no
BB4252		prob. acyl-CoA dehydrogenase	2.22	no
BB4516		prob. phosphosugar isomerase	2.58	no
BB4517		put. exported protein	3.22	yes
BB4518*		put. thioredoxin	4.03	yes
BB4520	*argB*	acetylglutamate kinase	2.21	yes
BB4827	*coIII*	put. cytochrome c oxidase	2.09	no
BB4831		put. bifunctional protein	3.77	no
BB4833		put. sodium	2.10	no
BB4835*	*fam*	RNA polymerase sigma-32 factor	8.80	yes
BB4858	*hisA*	histidine biosynthesis protein	2.13	no

Many of the genes identified through ROMA also have bioinformatically predicted SigE-dependent promoters (see below), including *fam*, which I previously identified as being transcribed by ESigE (Chapter 2, Chapter 3). Candidate SigE-regulated genes also identified through ROMA include the gene encoding the sRNA-binding protein *hfq,* some parts of the *sigE* operon itself, including *rseA*, *rseB*, and *mucD*, many genes encoding proteins involved in electron transport, and other cell envelope proteins. There are a few clusters of genes encoded in the same operon; rather than each gene having its own SigE-dependent promoter, it is likely that SigE-dependent transcription began at a single

promoter and continued through the operon. This polycistronic RNA was then used in microarray analysis, resulting in the gene clusters identified by ROMA. It is also likely that some of the genes identified by ROMA are only transcribed by SigE in vitro, but are not SigE-dependent in vivo. Further work characterizing these candidate promoters in vitro and in vivo is needed to confirm which are members of the SigE regulon. Preliminary results from some of this work will be described later in this chapter (See 'Confirming SigE-dependent transcription of candidate regulon members').

Prediction of putative SigE-regulated promoters in B. bronchiseptica

While there is much flexibility in portions of the promoter sequences recognized by RpoE-like sigma factors, there are also some common characteristics. Like many ECF sigma factors, σ^E-dependent promoters often have a characteristic AAC motif in the -35 region, and in *E. coli*, all known σ^E-dependent promoters contain a conserved cytosine in the -10 region. A consensus promoter for RpoE-like sigma factors from a few selected organisms was described by Staron, et al, as DGAACTY-n16-TCHVW (D=A, G, or T; Y=C or T; H=A, C, or T; V=A, G, or C; and W=A or T) [40]. I previously demonstrated that *B. bronchiseptica* SigE can recognize *E. coli* σ^E-dependent promoters, and that a *B. bronchiseptica* SigE-dependent promoter, upstream of *fam*, has similarity to this consensus sequence (Chapter 2).

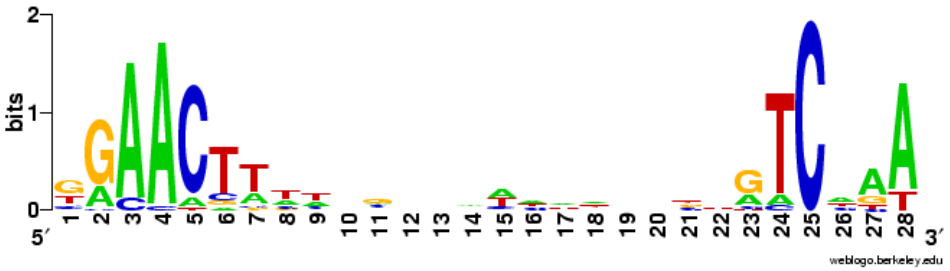


Figure 4-2: Alignment of σ^E-dependent promoters in *E. coli*. The height of the letter represents the information content at each position. For instance, in every σ^E-dependent promoter in *E. coli* there is a cytosine in the -10 region (base 25 on this plot), but the following nucleotide can be either A, T, or C, with fairly similar frequency.

Therefore, I began to search for *B. bronchiseptica* SigE-regulated promoters by aligning known *E. coli* σ^E-dependent promoters, building a position weight matrix (PWM) from this alignment, and using the matrix to scan the RB50 genome for similar sequences. The information content of the alignment is shown in Fig. 4-2. I selected a threshold of 72% identity with the consensus sequence, the median identity score for the promoters included in the PWM, and I also selected only sequences within 1000 bp of the nearest ORF on the same strand to enrich for likely SigE-regulated promoters. I first used the program, coded with the help of B. Tracy Nixon at Penn State University, to identify σ^E-dependent promoters in *E. coli* and validate the program design. The program predicted 522 candidate σ^E-regulated promoters in *E. coli*, similar to the 553 promoters identified in a previous attempt to bioinformatically predict σ^E-regulated promoters in *E. coli* [48]. Our program identified correctly 75% (24 out of 32) of the promoters originally used in the matrix that were above the threshold of identity to the consensus. Because the program successfully identified many known *E. coli* σ^E-dependent promoters, I then used it to scan

the *B. bronchiseptica* genome for similar sequences. From this, I identified 406 candidate SigE-regulated promoters for further analysis (Table 4-4).

Table 4-4: Predicted SigE-regulated promoters in *B. bronchiseptica*. Scores are defined as percent identity to the consensus σ^E-dependent promoter sequence (see Materials and Methods for more details). The distance from the end of the -10 region of the promoter to the translation start site of the nearest open reading frame is indicated by "Dist. to ORF." The -35 and -10 regions of the predicted promoters are indicated in bold. Asterisks denote genes that also have increased expression in RB50Δ*rseAB* compared to RB50 (Table 4-2).

Gene	Name	Product	Score	Dist. to ORF	Promoter
*BB3108		hyp. protein	94.5	113	**GGAACTT**GTGCGTGGCAGCATAG**TCCAA**
BB3756	*fabG*	3-oxoacyl-[acyl-carrier protein] reductase	93.9	149	**GGAACTT**GTGCGTGGCAGCATAG**TCCAA**
BB2611	*dxr*	1-deoxy-D-xylulose 5-phosphate reductoisomerase	93.9	234	**GGAACTTA**TGGCCAGGCGCTGG**TCGAA**
BB4503		put. exported protein	92.9	284	**TGAACTT**GCCCAGCGCGTGGGCG**TCCAA**
*BB3068		put. exported protein	92.6	87	**CGAACTT**CACCGCGCCTCCTGTG**TCTAA**
BB4917		put. exported protein	91.3	461	**GGAACTG**TTCCAGGCGGTGGAAG**TCAAA**
BB1844	*acnA*	put. aconitate hydratase	89.4	139	**GGAACTT**TTGACGCATCGATCAAG**TCTTA**
BB4281		put. membrane-bound lytic murein transglycosylase D	89.4	268	**CGAACTT**GCGCTGGGCGCTGGCGG**TCGAA**
*BB3739		put. integral membrane prot.	88.4	70	**GGAACTG**TCAGGTAGACTTATCG**TCTAA**
*BB4284		put. membrane protein	87.1	61	**TGAACTA**TTTGGCAGACCGCCAG**TCAGA**
BB0499		hyp. protein	86.8	601	**TGAAATT**GGTAGACGCAGGGGAC**TCAAA**
BB1089		put. enoyl-CoA hydratase/isomerase	86.8	74	**TGAAATT**TACTCGCTTGTTGTGA**TCAAA**
BB1133		hyp. protein	86.8	234	**GGCACTT**GCACCATACCGGCAGAG**TCAAA**
BB4348	*glyA*	serine hydroxymethyltransferase	86.5	870	**TGAACTT**ATCGAGTCTGTCATTT**TCCGA**
BB4234		put. TolQ-like transloc. prot.	85.8	121	**GGAACTG**CTTGCCCAGGGCAACA**TCCAA**
BB3977		put. exported protein	85.2	413	**AGAACTT**GTTGGGGGATGGCGC**TCATA**
BB4083		cons. hyp. secreted protein	85.2	77	**TGACCTT**CACTTGCGAGACGCT**TCAAA**
*BB4835	*fam*	RNA polymerase σ^{32} factor	84.9	64	**GGAACTT**TGGCGCGGCGCAGCAG**TCGTA**
BB0879		put. glycosyltranferase	84.9	177	**CGAACTT**CCTGCTCAAGGCGGTGC**TCAAT**
BB1497		put. exported protein	84.9	237	**GGAACCT**GCCGCAGTTCGCCGGCA**TCGAA**
BB0606		acetyl-CoA hydrolase/transferase	84.6	665	**TGAACTT**CGACGCCAAGGCGGATT**TCGTA**
BB4127		put. LysR-family transcript. reg.	84.2	491	**TGAACTT**ATCGCTCAGGGATCTG**GCAAA**
BB2111	*aroC*	chorismate synthase	83.9	848	**GGAACAT**GTTCGAGCCGTCCATGC**TCGAA**
BB4343		conserved hyp. protein	83.6	142	**GGAACTT**GTGCGCAGCGATGCAAG**CCTAA**
BB0729		put. short chain dehydrogenase	83.6	740	**AGAACTT**CCTGGGCGATACCTA**TCACA**
BB4590		conserved hyp. protein	83.6	837	**TGAACTC**CAGGGCCAGCGACGTA**TCGAA**
BB2199		put. phage repressor protein	83.3	550	**CAAACTA**TAGCGTTGCTATAGC**TCAAA**
BB2383		put. HlyD-family secretion prot.	83.3	530	**GGAACTT**GGGCAGCACCTCTTCG**TCGAT**

Gene	Name	Product	Score	Dist. to ORF	Promoter
BB4767		put. enoyl-CoA hydratase	83.3	251	**TGAACGT**CTCCGGCGGCCTGCTT**TCCAA**
BB4402	*pyrC*	dihydroorotase	83.0	700	**TGAACAT**CGCCGCCTCGTCCACG**TCCAA**
BB4378	*bph1*	histone protein	83.0	380	**GCAACTT**TTTTGAAACTCGCGTGA**TCGAA**
BB3520		phage-related hyp. protein	82.3	359	**TGAACTT**GCTTTATGCCAGCGTCG**TCCAT**
BB1566		biotin carboxyl carrier protein	82.3	685	**TGAACAT**GATCACCGGCGACCAGG**TCGAA**
BB1645		put. anti-sigma factor	82.0	666	**TGAACTT**CGAGCTGGACACCGAAA**TCTAC**
BB1396	*chlG*	molybdopterin biosynthesis prot.	82.0	921	**CGAACTT**GGCGGGCACGATGCCGA**TCAAC**
*BB4265		put. exported protein	82.0	82	**GGAACTG**GCGGGCACGCCGCATGG**TCTGA**
BB3005		prob. ATP-binding transp. prot.	81.7	58	**GGAACCT**GTGGGCGCGGCCGCGG**TCTGA**
BB4192	*asmB*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	81.7	223	**TGAACCA**AAGAGCCGTCCTACGGC**TCTAA**
BB2083		integral membrane protein	81.4	158	**AGAAATT**CGCGCGAACCTATAA**TCAAA**
*BB3264		put. exported protein	81.0	105	**CGAACCT**TTCGTTGGACACCTGC**TCCAA**
BB1089		put. enoyl-CoA hydratase/isomerase	81.0	73	**GAAATTT**ACTCGCTTGTTGTGA**TCAAA**
BB0409		conserved hyp. protein	80.7	543	**CGAACTG**CTGGAACACCTGCGGA**TCGAA**
BB0918		put. exported protein	80.7	728	**CGAACTC**GACACCAACAGGCTTA**TCGAA**
BB2050		hyp. protein	80.7	222	**CGAACTG**GCCAGGATTCCCGAGG**TCGAA**
BB2438		prob. 3-hydroxybutyryl-CoA dehydrogenase	80.7	124	**TGATCTT**CACGGCGCCTCACTGA**TCTAA**
BB3161	*rpoD*	RNA polymerase σ^{70}	80.7	841	**GGAACTG**TCGAGCAGAATCATGG**TCTTA**
BB3170	*hfq*	put. RNA-binding reg. prot.	80.7	128	**GGAATTT**CCTGAAAACTCTGGTG**TCATA**
BB0012	*rplJ*	50S ribosomal protein L10	80.7	209	**CGGACTT**GTCCGGATGCTGCTG**TCAAA**
BB0247	*def*	polypeptide deformylase	80.7	77	**GAAACCT**ATTAGAAATCAAGTG**TCTAA**
BB1352		put. aldo/keto reductase	80.7	490	**CGAACTG**GCGTCCCCTAGGGGAT**TCGAA**
BB3258	*radC*	DNA repair protein	80.7	399	**CGAACTC**GACCAGGTCGCCGGGC**TCGAA**
BB4843		put. exported protein	80.7	479	**CGAACTG**CTGGCGGCCTGGCCTT**TCGAA**
BB3990		put. exported protein	80.4	34	**CAAACTT**TTTACCACCTCTGGCTA**TCATA**
BB1588		put. cyclodeaminase	80.4	344	**AGAACTT**CGTGCGCTCCGGCGCCT**TCGGA**
BB1705		phage-related put. exported prot.	80.4	199	**TGAACTG**GGAGCTCTTCGGATG**TCTGA**
*BB0419	*sphB1*	autotransporter subtilisin-like protease	80.1	733	**TGAACTA**CCTGCAGATCGAGCTGA**TCAAG**
*BB1894		heat shock protease	80.1	193	**CGAACAT**GGCTGACGCCAACAACA**TCCAA**
BB2504		put. membrane transport ATPase	80.1	671	**GGAACTT**CATCGAACGGCGCGACG**TCAGC**
*BB3025	*trxC*	thioredoxin 2	80.1	363	**GGAACGA**GTGGGAGCGCTCCGCGG**TCCAA**
*BB4940	*dsbA*	thiol:disulfide interchge. prot.	80.1	54	**GGAACTT**CCCCGCGGCATCCGCAG**TCTGT**
BB0109		put. methyltransferase	79.8	889	**TGCACTT**CGTCGATGGTACGAAGA**TCCAA**
BB0655		put. transcriptional regulator	79.8	812	**CGAACGT**CCCCGACGTCGGGCCG**TCGAA**
BB3629		phage-related cons. hyp. prot.	79.8	461	**TGCACTT**GGTGCGGCAGACCAAG**TCCAA**
BB0906		put. integrase	79.8	30	**GGTACTT**TTCCCCGCAGTGGAG**TCCAA**
BB3577		conserved hyp. protein	79.8	133	**GGAACGT**AACAGGCAAAAGCGCCA**TCTTA**
BB0631		TetR family transcript. reg.	79.4	134	**GGAACAT**ATATTTTTCTTTTAAAT**TCAAT**
BB0672		conserved hyp. protein	79.4	262	**AGAACTT**CATGCCCGTGCATCC**TCAAG**
BB4004		conserved hyp. protein	79.4	906	**AGAACTT**CCAGCGCATGGGGGCGT**TCAAG**
BB0682		conserved hyp. protein	79.4	866	**AGAACTT**CCAGGGCGGCCTGGCCA**TCTAT**
BB1721		put. exported protein	79.4	109	**CAAACTG**ACGTTTTTCTGAATC**TCAAA**
BB4238		put. peptidoglycan-assoc. lipoprotein	79.4	451	**CGAACTT**TACGCCCGACGGCCGT**TCCAT**

Gene	Name	Product	Score	Dist. to ORF	Promoter
BB4960	*topB*	DNA topoisomerase III	79.4	122	**GGAACTG**ACGCCATGCGTACAGC**TCCGA**
BB1941		put. exported protein	79.1	400	**CCAACTT**CACCTACATGATCCAG**TCCAA**
BB4825		put. membrane protein	79.1	305	**GGAACTC**GCGCCGCCGTGCGCTA**TCTCA**
BB1438	*dnaX*	DNA polymerase III sub. Tau	79.1	86	**GGAACGA**AGCAGCCATAGTCGT**TCAGA**
*BB1709		phage-related hyp. protein	79.1	101	**CGAACTT**TCGTGTCAGCGGTGA**TCTAG**
BB2719		put. regulatory protein	79.1	141	**CGAATTA**AAATGGATTCTATGTAT**TCAAA**
BB3790		put. permease	79.1	569	**TGACCTT**CGGCCTGACGCTGCTCA**TCGAA**
*BB3932		put. zinc protease	79.1	58	**TGAACCA**ATCGAAACGTATACAC**TCCAA**
BB4839		conserved hyp. protein	79.1	989	**TGACCTT**GGTGTTGCCGATGCAC**TCGAA**
BB3204		prob. solute-binding prot.	78.8	199	**TGAACTT**CGTGCTGGACGCCGTGC**TCGAC**
BB3595	*phbB*	acetoacetyl-CoA reductase	78.8	651	**GGAACTA**CGTCGTCTCCAATTACC**TGAAA**
BB4355	*argC*	N-acetyl-gamma-glutamyl-phosphate reductase	78.8	113	**GGAACTC**CGCGCCGGTCGTCCGG**TCGGA**
BB0956		conserved hyp. protein	78.8	265	**GGAACTC**GGCGTGTCCGACCACG**TCGGA**
BB3051		put. exported protein	78.8	280	**GGAACTT**CTGTACGTGCCCGAGG**TCATG**
BB3386		put. ribonuclease	78.8	725	**GAAACCA**GGATCCGTATCGCTTGC**TCAAA**
BB0831		put. exported protein	78.5	699	**TGAACTT**CGGCCTGCGGCCAGCCG**TCAGC**
BB4989	*dnaA*	chromosomal repl. init. prot.	78.5	159	**GGAACCT**TCTGACAGACAGAAT**TCGGA**
BB1461	*dhpS*	dihydropteroate synthase	78.5	462	**TGAACCA**GATGACCACCGGCGCT**TCGAA**
BB3006	*apeA*	acyl-CoA thioesterase I	78.5	304	**GGAACGT**CCAGGCCCGCCAGCAG**TCCGA**
BB1884		put. periplasmic transport prot.	78.2	447	**GGCACTC**CGACAGAATGAGCTCTT**TCAAA**
BB2003		cons. integral membrane prot.	78.2	201	**TGATCTT**CGGCATGCCGCAGGAAG**TCCAA**
BB3325		put. ferredoxin	78.2	188	**GGAAGTG**CTGACGCCCGAGCAGA**TCAAA**
BB0523		put. ferredoxin	78.2	70	**TGAAATC**TTCGGAAATATTTTTT**TCAAA**
BB4233		conserved hyp. protein	78.2	975	**GGACCTG**CCGCTGGCGCGCACCA**TCAAA**
BB4274		put. membrane protein	77.8	295	**GAAACTT**GTACCACCCACCAGCA**TCCCA**
BB2346		put. acyl-CoA transferases/ carnitine dehydratase	77.8	832	**GGAAATT**CCGCCACGCATGGAA**TCAAT**
BB4582		put. cytochrome c	77.8	207	**GAAACTG**GCGCGCCCGACAGGAA**TCGAA**
*BB4905		conserved hyp. protein	77.8	86	**TGAACCC**CTCGCCCGCGGGGCAG**TCTAA**
BB3447	*cysN*	sulfate adenylyltransferase	77.5	426	**AGAACAT**GCGGGTGTTCCCGATT**TCCAA**
BB3923		hyp. protein	77.5	545	**TGAACCT**GAACGGCGACTACATC**TCCGA**
BB4467		put. membrane protein	77.5	139	**GGAACTT**GTGGCGTGCCGGCTTG**TCCGT**
BB1295		put. exported protein	77.5	40	**CGAACTC**GCGCTGCAGGCGCGCT**TCTGA**
BB1635	*bscU*	put. type III secretion protein	77.5	512	**GGATCTT**CGAGGCCATCGGCTTCG**TCATA**
BB1657		put. lipoprotein	77.5	199	**TGAACTC**AATGCACGGTTAGCCCG**TCTCA**
*BB1931		put. exported protein	77.5	68	**GAAACCT**TCGCCGGGCGCAGCGG**TCGAA**
BB2933		hyp. protein	77.5	595	**TCAACAT**CTGAGCGCCGCCGCCGC**TCAAA**
BB3013		MerR-family transcript. reg.	77.5	538	**TGACCTT**CGCGATTCTGCAGCAT**TCATA**
BB4462		put. outer membrane protein	77.5	291	**TGAACAA**TGTGGCCGACGTCAAGG**TCGAA**
BB4212	*alsB*	leucine-responsive reg. protein	77.2	210	**TGAACTC**GAAACCGGCCGTACCCA**TCGGA**
BB1779	*maiA*	maleate cis-trans isomerase	77.2	146	**GGAACTT**TTATGCGGGTTTTCCC**TTAGA**
BB3151		put. ABC-transporter permease	77.2	774	**TGAACTT**TGCCCAGGGCGAGTTCG**TCATG**
BB4472		two-component sensor kinase	77.2	134	**TGAACTT**CGACAGCGACACCAACG**TCATC**
BB0070		penicillin-binding protein 1A	76.9	364	**GGAACCC**GCACGCCGCAACGCGCC**TCCAA**
BB0251		hyp. protein	76.9	584	**AGAACAT**CTGTGTGTACTGCGGC**TCGAA**
*BB1360	*secD*	protein-export membrane prot.	76.9	87	**GGAACCA**TCAAGGCCCTGTAAGG**TCACA**
BB1894		heat shock protease	76.9	67	**GGAACCC**GCGCCCCGCCCGCCAG**TCCAA**
BB1955		put. integral membrane protein	76.9	591	**GGAACCC**TGGCACCACCCTTGGTG**TCCAA**
BB3020	*surE*	stationary-phase survival prot.	76.9	767	**GGAAGAT**TGATGCGTATTCTGGTT**TCAAA**
BB4385	*pcp*	put. lipoprotein	76.9	163	**GGAACCC**TGGCCGCGGCGGACGA**TCCAA**
BB0934		put. exported protein	76.9	31	**TAAACTT**GTCCGGATATTCAAGC**GCAAA**

Gene	Name	Product	Score	Dist. to ORF	Promoter
BB1997		put. malate dehydrogenase	76.9	680	**TGAACGT**GGGCGACAAATTGCCT**TCCGA**
BB3647		phage-related cons. hyp. prot.	76.9	884	**AGAACAT**TCCCGCGAAAGACGTCA**TCGAA**
BB3782		put. ferredoxin-NADP reductase	76.9	282	**CGAACTT**CCTCGTGATCGACCCCG**ACGAA**
BB4116		prob. transcriptional regulator	76.9	88	**GGAACTT**AAGTATCGCTGTGAGATA**CTTA**
*BB4518		put. thioredoxin	76.9	117	**GGAACCA**AATCCGGCGAGGGCTCG**TCTGA**
BB1849		hyp. protein	76.6	450	**TGAACTT**CGACCGCATCCCCGAC**TACAA**
BB1924		conserved hyp. protein	76.6	229	**AGAACTT**CCTCATGATGGGAAT**CCTAA**
BB2257		prob. LysR-family transcript. reg.	76.6	152	**CAAATTT**GATATACGTTGCATA**TCAAA**
BB2526		LysR-family transcript. reg.	76.6	87	**ACAACTT**GTTTTCGAAAACATCAT**TCCAA**
BB3632		phage-related cons. hyp. prot.	76.6	85	**CGAACTT**GCAGGAAGCCGCACGTG**CCCAA**
BB0054	*rpsM*	30S ribosomal protein S13	76.6	131	**GGATCTG**CCGCAACTGCAAAGTTA**TCAAA**
BB1574		put. biotinylated protein	76.6	366	**TGAACAT**GGGCGGCGGCAGGAAC**TCCGA**
BB1812		put. nuclease	76.6	182	**TGAACTT**CGAATACATGCCCGA**ACTGA**
BB0475		put. sulfatase	76.2	611	**CGAACTG**TCGAGCGTGCTGGTGG**TCAAT**
BB1533		put. creatinine amidohydrolase	76.2	81	**CGAACTG**CGCAAGCTGTGGTCCG**TCAAT**
BB1642		put. regulator	76.2	100	**CGAACCA**CGTACTCATGGCTGAA**TCCAA**
BB2954	*glnA*	glutamine synthetase	76.2	402	**CGAACTG**CCGTTCGAGTTCATGC**TCAAT**
BB3239	*polA*	DNA polymerase I	76.2	45	**GGAATTT**GTTTGCCGGCGGGTG**TCTCA**
*BB3493		phage-related cons. hyp. prot.	76.2	150	**GGAACTA**TCATCAGCGCAATATTC**CCCAA**
BB3751		put. membrane protein	76.2	297	**GAAACTT**TTAACGAAACCGACAA**CCTAA**
BB3841	*nuoA*	prob. NADH-ubiquinone oxidoreductase chain 3	76.2	180	**GGAACGC**AAGTTCCGCTTCGAT**TCCAA**
BB3862		conserved hyp. protein	76.2	503	**CGAACTC**GATGACCGGCCATTTGC**TCAAT**
BB4704		prob. enoyl-CoA hydratase/isomerase	76.2	984	**TGAACTC**CGATTTGCCCACGCCA**TCAAC**
BB1579		put. membrane protein	76.2	571	**TGAACTG**TCCGAACGACGGGAGTA**CCAAA**
BB1837		put. ECF sigma factor	76.2	48	**GGAACGC**GGGCGCGGCCCGCGCCA**TCCAA**
BB1838		put. exported protein	76.2	139	**GGAACGG**CAGGGGCGGCCGCCTCA**TCCAA**
BB2227		phage-related cons. hyp. protein	76.2	151	**GGAACTA**TCATCAGCGCAATATTCCCCAA
BB2607	*pyrH*	uridylate kinase	76.2	116	**CGAACTG**CCGGCCTATTTCTTG**CTTAT**
*BB3801		put. intracellular septation prot.	76.2	434	**CGAACCA**ATTCGCACGGCGGGAC**TCCAA**
BB4072		conserved hyp. protein	76.2	95	**GGAACAT**CCACTTGCTACCGCACG**TCGTA**
BB4224		put. oxidoreductase	76.2	268	**GGAACCG**CGGGCTCTCGGTCGATA**TCGAA**
BB4688		conserved hyp. protein	76.2	487	**GGAACCC**CTGTCCATCGACGACA**TCGAA**
BB4895	*ptlA*	pertussis toxin transport prot.	76.2	407	**GAAACTT**TCTGCATCACGACCA**TCTAC**
BB0068	*lysA*	diaminopimelate decarboxylase	75.9	233	**AGAAATT**CGGTTTCCGTCATAA**TCCAA**
BB0203		put. integral membrane transport prot.	75.9	176	**GGAAATA**TCGCAATTCATTGAT**TCATA**
BB1136		put. calcium/proton antiporter	75.9	287	**TGAACTT**GCCCATTTCGGCCAGCA**TCCGT**
BB1270		put. ArsR-fam. transcript. reg.	75.9	582	**CGAACTT**CACCACCGGGTTGAGC**CCGAA**
BB2257		put. LysR-fam. transcript. reg.	75.9	153	**TCAAATT**TGATATACGTTGCATA**TCAAA**
BB2448		put. exported protein	75.9	556	**GGAACCG**GTCGGTCAACAATGC**TCAGA**
BB2950		conserved hyp. protein	75.9	143	**TCAAATT**AATTGAATGATGTCA**TCAAA**
BB3158		put. phage-related protein	75.9	662	**GGAAATT**TCTCGGCTCATGGAG**TCGGA**
BB4036		put. conserved membrane prot.	75.9	693	**CGAACTT**GGGCAGCGACGTGGTGC**CCGAA**
BB4182		conserved hyp. protein	75.9	460	**TGAACCT**GACCGTCTTCGTGTCCA**TCAAG**
BB0069	*cyaY*	conserved hyp. protein	75.9	764	**CGAACTT**GTTTTCCTTCAGGCCCG**TCGAG**
*BB1368		put. membrane protein	75.9	127	**GGAACCC**TGGGCGGACGGGCGGG**TCAGA**

Gene	Name	Product	Score	Dist. to ORF	Promoter
*BB1933		put. hydrolase	75.9	51	**AGAAATT**GCCCGGCACGGTGTGCT**TCCAA**
BB4103		conserved hyp. protein	75.9	852	**TGAAATA**GTACGGCAGCGAGATG**TCGAA**
BB4128		put. exported protein	75.9	417	**ATAACTT**GACGCCGAGCAGGTCT**TCCAA**
BB4248		conserved hyp. secreted prot.	75.9	279	**TGAACCT**GGGCAATCCGCGCGCCT**TCAAC**
BB0702		put. GntR-fam. transcript. reg.	75.6	554	**CGAACTT**GCCGCTCAGCACGTA**CCAGA**
BB0936		put. ArsR fam. transcript. reg.	75.6	123	**GGAAGTA**AGCTCTTGCTGATATT**TCAGA**
BB0995		put. dipeptidase	75.6	453	**CGAACTT**CATCGCCGCCACCCTGC**TCAGC**
BB1206		put. exported protein	75.6	39	**CGAACCA**GGAGATTTTTAGAAT**TCGAA**
BB2166		DedA-family integral membrane prot.	75.6	979	**GGAACGT**GTAGTAGCCGGCGAT**TCCCA**
BB0368		branched-chain amino acid transport system permease	75.6	766	**TGAACTT**CGCGCACGGCGAGTTCT**TCACC**
BB0456		prob. transcriptional regulator	75.6	337	**GCAACTT**CCGCCTGGTCGGGCGCA**TCAAG**
BB0600		put. branched-chain amino acid transport permease	75.6	766	**TGAACTT**CGCTCACGGCCAGTTCT**TCACC**
BB1282		two component system response regulator	75.6	472	**GGAACGT**ATCCGCAATCTGGCGG**TCCCA**
BB2045		put. lipoprotein	75.6	551	**TGAACTT**TGCCCGCAGCCAGGCT**TCTGG**
BB2563	*flgF*	flagellar basal-body rod prot.	75.6	572	**TGAACTT**CGACACCGCCGGCAA**TCTGC**
BB2636		put. exported protein	75.6	212	**CGAACCA**TCGGCGCGCGCGGGAA**TCGAA**
BB3009		conserved hyp. protein	75.6	242	**GGAACGC**GTGCCGGCCGACTGGC**TCGAA**
*BB4506	*rpoN*	prob. σ^{54} modulation prot.	75.6	143	**GCAACTT**GTATTGGCCAATAGAG**TCTAT**
BB0526		put. phage-related protein	75.3	554	**GGAAATT**GCAGACGCCATCCGT**TCCTA**
BB0761		branched-chain amino acid ABC transporter	75.3	203	**CGAATTC**GATGAACTTCATTGACT**TCAAA**
BB1251		put. exported protein	75.3	164	**GGAACAG**GCGCAGGTCGGTGAGG**TCGAA**
BB1521		put. diaminopropionate ammonia-lyase	75.3	772	**GGAACAC**GAAGTGGATCGTGCCG**TCGAA**
BB2826	*citE*	put. citrate lyase beta chain	75.3	265	**GGAACTA**CAAGCCGGCGGCTTTCG**TCAGC**
BB0330	*gltJ*	glutamate/aspartate transport system permease	75.3	164	**TGAACGT**GGCAATGAGCGAGGTGC**TCAAG**
*BB1617	*bsp22*	put. secreted protein	75.3	60	**GGAACTT**GCCGCGGGCGCAGGGTT**ACTCA**
*BB4235		put. TolR-like translocation prot.	75.3	290	**TGAACTT**CCTCGCCTCGGCGGGC**TCGGT**
BB4517		put. exported protein	75.3	133	**GGAACTC**ATCACCCCGCACGACG**TCCAT**
BB4993		prob. inner-membrane protein	75.3	233	**TGAACGT**CATGGCGCATGGCGCGG**TCTAT**
BB0108		prob. amino acid permease	75	953	**GGAACCT**TTCGTGATGACAGGAAC**TCCAT**
BB0583		put. ligase	75	751	**TGAACAT**TCATTCGATGTCAATT**TCAAG**
BB0841		put. exported protein	75	34	**GGAACCT**TTAGCCTGGTTTTACA**TCCAT**
BB1515		put. ABC transport membrane prot.	75	650	**GGGACTA**TGGCTTTCGCCAGGTC**TCCAA**
BB2412		conserved hyp. protein	75	72	**CTAACCT**GTATTAAATTAGCACT**TCAAA**
BB2998		put. amidase	75	475	**CGAACTG**TACGTCAGCCTGGCC**TCCGA**
BB3239	*polA*	DNA polymerase I	75	45	**GGAATTT**GTTTGCCGGCGGGTGTC**TCAAG**
*BB3678		conserved hyp. protein	75	53	**GGAACTG**ATGGTCCGTCGTACGG**TCTAG**
BB3877		prob. zinc-bind. dehydrogenase	75	347	**GAAACCG**CATGGCTGTAAGGAG**TCAAA**
BB0901	*prs*	ribose-phosphate pyrophosphokinase	75	960	**CGAACTG**CCCGGCGTGGCCGAG**TCCGA**
BB1494		conserved hyp. protein	75	651	**TGAATTT**GCGCGCGAAGATCTGG**TCCGA**
BB1910	*xseB*	exodeoxyribonuclease VII	75	968	**GGAATTT**TATTTTGCTATCCGAAA**TCAAC**
BB2562	*flgE*	flagellar hook protein	75	31	**TCAACTT**CTGAGAGCTAACGGG**ACAAA**
BB2995	*bvgS*	virulence sensor protein	75	169	**GCAACTG**CTGGCGCAAGGCATG**TCCAA**

Gene	Name	Product	Score	Dist. to ORF	Promoter
BB3873		prob. transcriptional regulator	75	435	**GGAACTC**ACCTGGATGATGGACG**TAAAA**
BB3965		put. exported protein	75	934	**TGAACAT**CCTGAAGTCGTTCGACC**TCAAG**
BB0940	*fdx*	ferredoxin	74.6	254	**TGAACCG**CCACCTGCTGCCCGACG**TCGAA**
BB1067		put. exported protein	74.6	714	**GGAACTC**GGCGGCAAGAGCCCC**TCGAT**
BB3128		put. membrane protein	74.6	653	**TGAACGT**ACATCTCTTCCAGCCCG**GCAAA**
BB4651		put. ferredoxin	74.6	113	**TAAACTG**CATACGGTATACAATGT**TCATA**
BB0360		transcriptional regulator	74.6	97	**TGAAGTT**GCTCATGTTGATTGAAT**TCAAT**
BB0599		put. branched-chain amino acid transport permease	74.6	754	**TGACCTT**TACGCAGATGAACTCGG**TCAAT**
BB1204		LysR-family transcript. reg.	74.6	160	**CAAACTT**GCCGGCCGCCCCTACAA**CCAAA**
BB1348		put. exported protein	74.6	388	**GGAACTG**TGCGCCCGCACGGGCA**TCGAT**
BB2330		prob. membrane efflux protein	74.6	744	**GGAATCA**GTGGTCGCCGTTCGAAA**TCAAA**
BB3663		putattive exported protein	74.6	488	**GGAACTG**AACATTGCCGTTGCCA**TCGAT**
BB4233		conserved hyp. protein	74.6	976	**TGGACCT**GCCGCTGGCGCGCACCA**TCAAA**
BB4239		put. periplasmic protein	74.6	326	**CGAACAA**CCAGCAGCGCATCAAGA**TCGAA**
BB4923		put. membrane protein	74.6	43	**CGAAGTT**CTAGCGCGGGACGAG**TCATA**
BB5011		put. membrane protein	74.6	346	**CGAACCT**GTAGGGCCGCGCGCATT**TCCGA**
BB1150		put. transcriptional regulator	74.3	149	**GGAACTT**TGTCTCTATTTTCGAT**TATGA**
BB1372	*ubiD*	3-octaprenyl-4-hydroxy-benzoate carboxy-lyase	74.3	687	**CGAACTT**GTCCTGGTGCACCCGGG**TCATC**
BB2252		put. membrane protein	74.3	836	**GGAACTT**GAGCGGGAACACCTTGC**TCTCG**
BB2504		put. membrane transp. ATPase	74.3	49	**GTAACTC**CAGGGTGTGAAATGCG**TCCAA**
BB2657		put. membrane protein	74.3	358	**AGAAGTT**CAACGAGTTCGACGTG**TCCAA**
BB3605		put. iron uptake protein	74.3	634	**AGAAGTT**CAACGAGTTCGACGTG**TCCAA**
BB4772	*vanB*	put. vanillate O-demethylase oxidoreductase	74.3	999	**GGAAATG**CCCAGCGAGAAGAAT**TCCAA**
BB4869	*mscL*	large-conductance mechanosensitive channel	74.3	963	**TGAAATT**CTTCGGGGCGTAGTAC**TCGGA**
BB1151		put. branched-chain amino acid-binding protein	74.3	272	**GGAAGTT**TCTCTGTAGCCATGGGT**TCGGA**
BB1172		conserved hyp. protein	74.3	509	**GGAATTC**GTCCCATTGCGCGGTG**TCCAA**
BB1272		conserved hyp. protein	74.3	61	**TGAATTT**CGTCGCCGAGAAAGTGC**TCGGA**
BB1460	*ftsH*	cell division protein	74.3	271	**GAAACCT**TTCTGGTCGCCCGCGA**TCTGA**
BB1980	*tonB*	siderophore-mediated transport protein	74.3	769	**TGAACAA**AACCGAACTCATCGA**TCACA**
BB2177		prob. fatty acid desaturase	74.3	112	**GCAGCTT**AGTGCTTGATTTCCC**TCAAA**
BB2541	*flbB*	flagellar transcriptional activator	74.3	204	**GTAACTT**TACCGGAGATAATCT**GCAAA**
BB0344		conserved hyp. protein	74.0	652	**TGAACGC**GCTGCGGATCGGCTATG**TCGAA**
BB3264		put. exported protein	74.0	104	**GAACCTT**TCGTTGGACACCTGC**TCCAA**
BB3492		phage-related cons. hyp. prot.	74.0	493	**TGTACGT**GCGGTCAACACAGCA**TCAAA**
BB3884	*pssA*	put. CDP-diacylglycerol-serine O-phosphatidyl-transferase prot.	74.0	448	**TGAACTA**CTCGATCTCGAACAACG**CCGAA**
BB0586		put. exported protein	74.0	995	**TGAACTA**CGACGACTTCAAGCACA**TCGAC**
BB0802		conserved hyp. protein	74.0	291	**TGAACTA**CGTCGACGGCGACCCG**GCCAA**
BB1155		prob. permease of branched-chain amino acid transporter	74.0	269	**GCAACCT**GGTCGGCCCGGGCGACA**TCGAA**
BB1554		put. UDP-N-acetylglucosamine 1-carboxyvinyltransferase	74.0	309	**TGAACTA**CACCCTGAGGCAAACCA**GCCAA**
*BB2228		phage-related cons. hyp. prot.	74.0	494	**TGTACGT**GCGGTCAACACAGCA**TCAAA**
BB2467		put. exported protein	74.0	668	**TGAACTA**CGGCTCGATCTCGCACA**TCGAC**

Gene	Name	Product	Score	Dist. to ORF	Promoter
BB2589	*bvgA*	virulence factor transcript. reg.	74.0	695	**GGAACTT**GTTGTGCGAGACGCCG**CCGGA**
BB3318		conserved hyp. protein	74.0	595	**TGAACCT**GTCGTCGCTGCAGTTC**TCGCA**
BB3334		hyp. protein	74.0	947	**GCAACCT**GGACCTGCACGAACAGG**TCGAA**
BB3691		conserved hyp. protein	74.0	68	**TGATCGTG**GTCACGCTGGTTGTCG**TCAAA**
BB3737		conserved hyp. protein	74.0	350	**GCAACGT**CAATAGCGGCCTGTTC**TCCAA**
BB4369		put. GntR-fam. transcript. reg.	74.0	228	**CGAACGT**ATACTATATTCTCAAAT**TCCGA**
BB4520	*argB*	acetylglutamate kinase	74.0	779	**GGAACTT**CAGCTTCTTGGGGTTC**TCGGC**
BB4594		put. binding-protein-dependent transport permease	74.0	751	**TGAACTA**CGCGCCGGCGCGCTTCG**TCGAC**
BB4627	*trpG*	anthranilate synthase comp.	74.0	737	**GGAACTT**CGGCGATTTCCAGGTGG**TCGGC**
BB0161	*birA*	put. biotin protein ligase	73.7	88	**GGAAATG**CGCGGATACACAGAAAT**TCGAA**
BB0308		put. transcriptional regulator	73.7	218	**TGAACAG**ATGCTGCGCAGCTTTG**TCGAA**
BB0481		put. TetR-fam. transcript. reg.	73.7	146	**GGAAATT**CCTGCCTGGCCTCCCTC**TCCCA**
BB1001		put. enoyl-CoA hydratase	73.7	172	**CCAACTA**CAAGGTGCCCAAGCGCA**TCGAA**
BB1458		conserved hyp. protein	73.7	310	**TGAACTC**CACCCCGGCCACCGGC**TCCAT**
BB1490		put. branched-chain amino acid transporter	73.7	49	**GGAAATC**CGGCTGGCGCTACTTCA**TCGAA**
BB1842		conserved hyp. protein	73.7	135	**GGAATTT**TCAGCCCAGCAATAAA**TCCCA**
BB2381		put. membrane protein	73.7	64	**TGAACAC**GGCCTTCATCGGCATCG**TCGAA**
BB2433	*norM*	multidrug resistance protein	73.7	43	**GGAAAAA**AGGTAATATCCCGAT**TCAAA**
BB3029		put. exported protein	73.7	163	**GGAAATC**CTGCGTCTGCTTGGCA**TCGAA**
BB3134	*ilvB*	acetolactate synthase	73.7	371	**TGAACAC**CACGGCCAAGTCGCTGG**TCGAA**
BB3623		phage-related cons. hyp. Prot.	73.7	381	**GGAAATC**ATCCGCCTGACCGAC**TCGAA**
BB4802	*glmS*	glucosamine--fructose-6-phosphate aminotransferase	73.7	82	**CAAAATT**CGTAGAATTAAGAAA**TCTAA**
BB0479		hyp. protein	73.7	185	**TGACATT**GGGGGCGGCTTATACGC**TCAAA**
BB0650		put. exported protein	73.7	71	**TGAACAC**CGGCGTGCGCGCCGATA**TCGAA**
BB1311		put. membrane transport prot.	73.7	226	**GAAACGT**AACCGTGGGAAATGAA**TCTGA**
BB2771	*paaG*	put. enoyl-CoA hydratase	73.7	787	**TGAACAC**ATCCCCGACCGTAAGAT**TCGAA**
BB4092		prob. acyltransferase	73.7	510	**TCAACCA**GGACAGCGATGCCTTCG**TCAAA**
BB4145		phage-related hyp. protein	73.7	505	**GGCACTT**GATGACATCAAGGGAT**TCCTA**
BB0178	*hslU*	ATP-dependent Hsl protease .	73.3	124	**CGAACTG**CCGCCCGAAGCCATCG**TCAAG**
BB0341		prob. GntR-family transcript. reg.	73.3	571	**CGAACTG**GGCGCCGACCACGTCA**TCAAC**
BB0790		put. type II secret. syst. prot.	73.3	755	**CGAACTG**GGCTGCCCGGCCTCGT**TCAAG**
*BB1733	*ppiB*	peptidyl-prolyl cis-trans isomerase B	73.3	330	**TGAACTC**GAACGCGCCCAGGACAG**TCTAC**
BB2374		put. malate/L-lactate dehydrogenase	73.3	961	**AGAACTT**CGAGAACACCATGGCC**GCCAA**
BB2418		put. oxidoreductase subunit	73.3	279	**CGAACTC**GGGGTCGATGCGCCTGC**TCAAG**
BB2438		prob. 3-hydroxybutyryl-CoA dehydrogenase	73.3	104	**TGATCTA**ATTTATCGATCATTCTA**TCCAA**
BB2587	*flaG*	flagellar motor switch protein	73.3	462	**TGAACTA**CCTGCCCAACAAGGAAG**GCGAA**
BB2721		put. transcriptional regulator	73.3	167	**TAATCTT**GATTATTTAAAGTGACT**TCTAA**
BB2724		prob. short-chain dehydrogenase	73.3	613	**GTAACCT**CGATTCCGACATTGA**TCGAA**
BB3007	*ppiD*	peptidyl-prolyl cis-trans isomerase D	73.3	956	**GTAACTG**AACAAGGCGGCTCCGCG**TCAGA**
BB3416		put. polyamine transport ATP-binding protein	73.3	119	**TGATCTA**CCCGCCCGCCGACGTG**TCCAA**
BB3556		put. transcriptional regulator	73.3	70	**TCAACTA**CTCCATCTTGTCCCGCA**TCTGA**
BB3841	*nuoA*	prob. NADH-ubiquinone oxidoreductase chain 3	73.3	256	**AGAACTT**CGCGTTCCTGGAAGAT**GCCAA**

Gene	Name	Product	Score	Dist. to ORF	Promoter
BB4193	*ftsZ*	cell division protein FtsZ	73.3	405	**CGAACTG**GCCGAGGACGTGTTCC**TCAAG**
BB4347		conserved hyp. protein	73.3	195	**TGAACTG**ACCGCCCGCCTGCCCG**TCTAC**
*BB4491		put. exported protein	73.3	304	**TGAAATT**CTTGCGCTTGGCCACT**TCAAC**
BB4607	*atpA*	ATP synthase alpha chain	73.3	182	**CGAACTC**AAATTCGGTCTCAAGC**TCAAG**
BB4913		put. exported protein	73.3	281	**ACAACTC**CAGCGGCAGCCTGTCGC**TCAAA**
BB0060	*cutA*	put. periplasmic divalent cation tolerance protein	73.3	72	**GGAACAT**ACCCGCCGATCGTTA**TCGAT**
BB0876	*wbpO*	polysaccharide biosyn. prot.	73.3	504	**GGAACTT**CCCGTTCGGACGGGTGT**TCCTG**
BB1218		prob. enoyl-CoA dehydratase	73.3	913	**AGAACTT**CAAGGCCGGCGGCCTG**GCCAA**
BB1301		put. exported protein	73.3	552	**CGAACTG**TCCACCCGTTTCGATT**TCAAG**
BB1752		conserved hyp. protein	73.3	106	**AGAACTT**CCGCCAGCAGGTGGTG**GCCAA**
BB1885		put. transcriptional regulator	73.3	634	**CGAACGA**ATAGGTCAGCAGGCTT**TCATA**
BB1939	*glcE*	glycolate oxidase subunit	73.3	169	**CGAACTC**GATGCCTTCCTGGCGG**TCAAG**
BB2246	*phg*	autotransporter	73.3	63	**TAAAATT**CGATTTTGTATTTTGTA**TCGAA**
BB2569	*flaT*	flagellar hook-associated prot.	73.3	276	**TGAACGA**GATGTTCTCGCAGCTCG**TCAAT**
BB2880		put. membrane protein	73.3	114	**TCAACTG**TCACGCTTGTTGTGTT**TCCAA**
BB3224		put. cell surface protein	73.3	314	**CGAACTG**GCCCTGCACGCACGCG**TCAAG**
BB3402		put. dioxygenase	73.3	286	**AGAACTT**GATCAGGCCGAAGATGT**TCGAC**
BB4138		phage-related integrase	73.3	952	**CAAACTC**GATCGCCCGCAACTA**TCGAA**
BB4571		branched-chain amino acid transp. ATP-bind. prot.	73.3	570	**GGAAATC**AGCCGCCTGAAGCTGCA**TCAGA**
BB4758		MarR-family transcript. reg.	73.3	705	**GGACCTT**CTCGGCCGGCGGCGACC**TCAAC**
BB4842		put. outer membrane protein	73.3	47	**CGAACTG**TCACGATTTCCACACAC**TCAAC**
BB4865	*tatC*	put. Sec-independent translocase prot.	73.3	286	**CGAACTC**GACGAACTGCGCAAGT**TCAAG**
BB4945	*ivd*	isovaleryl-CoA dehydrogenase	73.3	173	**TGAAATT**GCTAGGGAAAACCCGAC**TCAAC**
BB0317		LysR-family transcript. reg.	73.0	35	**TGAACTT**TCGCCAGGTCGAGACTT**TCCGG**
BB2310		put. integral membrane protein	73.0	967	**CGAACAT**CATCAGCTCGACGTCGG**TCGGA**
BB2352		prob. LysR-family transcript. reg.	73.0	526	**CGAACTG**GCGCCGCTGGGCCTG**TCGTA**
BB2499		conserved hyp. protein	73.0	964	**AGAACTT**CCAGCTGTTTCCCCA**CCAGA**
BB2539	*flaA*	flagellin	73.0	154	**CGAACTT**CACTTTTTTTGCTTAAG**TCCGT**
BB3251		put. LysR-fam. transcript. reg.	73.0	73	**CCAATTT**TAAGTGCTTGAAACGCA**TCAAA**
BB3420		put. inner membrane transport permease	73.0	511	**GGAACTT**CGCCAGCATATCCTGG**CCTAT**
BB3488		phage-related cons. hyp. prot.	73.0	224	**TGAGCTT**CGTGCTGGATGGTGA**TCAAT**
BB3837	*nuoE*	respiratory-chain NADH dehydrogenase I	73.0	503	**TGGACTT**CGACGTGCCGGTCGGCG**TCAAT**
BB4406	*rubA*	rubredoxin	73.0	184	**AGAACGA**GGGGAGGCGTAACGGAG**TCCAA**
BB0472	*cspA*	cold shock protein	73.0	841	**GGAAAAT**TTTGGGCATGCTTGC**TCCAA**
BB0673	*gst*	glutathione S-transferase	73.0	99	**TGAATTT**GTTTAAAGTGAATATA**TCGTA**
BB1501		put. enoyl-CoA hydratase	73.0	106	**TGATCTT**GCATGTATGATTGCAT**TCAAT**
BB1957		put. lipoprotein	73.0	53	**GGAATTT**TGTTTGTCAAGAACCTG**TCGCA**
BB2177		prob. fatty acid desaturase	73.0	87	**AAAACTA**TTTTGAAACCACGCTG**TCATA**
BB2232		phage-related cons. hyp. prot.	73.0	225	**TGAGCTT**CGTGCTGGATGGTGA**TCAAT**
BB2287	*livM*	branched-chain amino acid transport system permease	73.0	31	**GGAACCC**GCGAGATCGAGAGGC**TCTGA**
BB2552	*cheZ*	chemotaxis protein CheZ	73.0	218	**CGAACCT**CGACGGCCTGGAGATGC**TCAAG**
BB3060		put. glycerol-3-phosphate dehydrogenase	73.0	834	**CGAACAT**GACCCTCAATCCCAG**ACAAA**
BB3308		amino acids ABC transporter	73.0	535	**CGAACCT**GCAGACCATCGAGTCGG**TCAAC**
BB3421		put. periplasmic substrate-binding transport protein	73.0	701	**CGAACTG**CATGCCGCGATCGGGC**TCGTA**

Gene	Name	Product	Score	Dist. to ORF	Promoter
BB4213	*bllY*	put. hemolysin	73.0	754	**GGAACGA**ACCGCCCCTTTTTATT**TCGGA**
BB0409		conserved hyp. protein	72.7	350	**GAAACCA**TGGCGCCCGGGCTGG**TCGAA**
BB0825	*modB*	molybdate-binding periplasmic prot.	72.7	394	**GGACCTG**CTGCTCGACTACGCG**TCCAA**
BB1256		put. dimethylmenaquinone methyltransferase	72.7	613	**TGAGCTA**CGGCGAGAACGGCGAA**TCGAA**
BB1807	*goaG*	4-aminobutyrate aminotransferase	72.7	578	**CGAACTT**CTGGTGCGGGAACTCGG**TAAGA**
BB2329		put. TetR-fam. transcript. reg.	72.7	928	**CGAACTG**CAGCAACAACGTCAGC**GCAAA**
BB2884		conserved hyp. protein	72.7	794	**GGAACTT**GCCCAGCGCATGCGG**CCGTA**
BB3324		put. lipoprotein	72.7	309	**AGAACGT**ATACCGCCTGCAGATCA**TCAAT**
BB3420		put. inner mem. transp. permease	72.7	511	**GGAACTT**CGCCAGCATATCCTG**GCCTA**
*BB3749	*mucD*	serine protease	72.7	651	**AGAACTA**CCACATCCGGGCCGAA**TCGCA**
BB3889		put. membrane transport prot.	72.7	754	**TGGACTA**CGACTCTCCCGTTCACG**TCGAA**
BB4343		conserved hyp. protein	72.7	142	**GGAACTT**GTGCGCAGCGATGCAA**GCCTA**
BB4817	*glmU*	UDP-N-acetylglucosamine synthesis bifunctional protein	72.7	562	**TGAACCT**GTCGATGCTGTCGAGC**TCGAT**
BB0079		conserved hyp. protein	72.7	172	**GGACCTA**TGGCCTGACTGCCAG**TCTGA**
BB0520		put. membrane protein	72.7	195	**ACAACTT**GCAGTTGTACGCAAA**TCTGA**
BB1065		put. NAD dependent epimerase/dehydratase	72.7	480	**GAAACCA**GCGCGCGGCTGCTGC**TCGAA**
BB1543		conserved hyp. protein	72.7	822	**TGCACTT**TCTTTGCCGCGGACGGC**TCGGA**
BB1703		phage-related hyp. protein	72.7	291	**TCAACTT**CATCCCGACGGCAGAG**TCCCA**
BB2185		integrase	72.7	900	**GAAACCA**TGGCGATCGGACAGC**TCGAA**
BB4226		LysR-family transcript. reg.	72.7	240	**GGATCTT**GACCGTACCGCCGGGT**TCGGA**
BB4234		put. TolQ-like translocation prot.	72.7	58	**CGAACTG**CCGGACGATATTCGC**GCAAA**
BB4524	*ldh*	put. L-lactate dehydrogenase	72.7	663	**CGAACTC**GGCCCCATGACCTCGGC**GCAAA**
BB4841		glutamate synthase [NADPH]	72.7	180	**GAAACCA**TTCTGCGCCGCCTGG**TCGAA**
BB4990	*rimA*	50S ribosomal protein L34	72.7	419	**AGAATTC**TTTCATGTCGCTCTTGT**TCAAA**
BB0500		put. integrase	72.4	208	**GAAATTG**GTAGACGCAGGGGAC**TCAAA**
*BB1289		put. integral membrane protein	72.4	589	**GGAACTG**GTCGGACCGGACTTCT**TCCAC**
BB1891		conserved hyp. int. mem. prot.	72.4	180	**TGAACTT**GACCACGATGGCCTTG**CCGGA**
BB2054		put. GntR-family reg. protein	72.4	353	**TGAACGT**AATCATAACCCATTT**TGAAA**
BB2296	*recJ*	ss-DNA-specific exonuclease	72.4	142	**GGAACGC**CTGGTGGTCTCGCCCCG**TCTGA**
BB2443		conserved hyp. protein	72.4	292	**CGAACGT**CGGCCTGCAGTTCGACA**TCTAT**
BB2743		conserved hyp. protein	72.4	450	**GGAACTG**GCCAACTACCTGGCC**TCCAC**
BB3062		put. oxidoreductase	72.4	166	**TGAACCA**GGCGGTCGACGGTTT**TCTCA**
BB3362		put. GntR-fam. transcript. reg.	72.4	861	**AGAACTC**GACCAACCTTCGCTG**TCCGA**
BB3731		put. LysR-fam. transcript. reg.	72.4	506	**GCAACTT**CGTCGGCCCGACCATCT**TCGAT**
BB3924	*icd*	isocitrate dehydrogenase	72.4	217	**GGAAGGA**TTGAACATCGGCGCT**TCAAA**
BB0301	*thiF*	adenylyltransferase	72.4	784	**GGAACTC**GGCGCCAAGGGTGCG**CCCAA**
BB0849		hyp. protein	72.4	105	**GAAATTC**TGCAGCATCCTCGGGCG**TCAAA**
BB1693		portal protein	72.4	143	**TGAACTT**GTCGATGGCGCCGGA**CCGGA**
BB2912	*ugpB*	glycerol-3-phosphate-binding periplasmic protein	72.4	239	**TGAACGC**GATGGAGAGATCGTTTT**TCATA**
BB2917		conserved hyp. protein	72.4	989	**TGAACTT**GGCATCCAGGCCCTGC**TCGGC**
BB3555		put. fumarylacetoacetate-family hydrolase	72.4	812	**CGAACCT**CGGCCTGCCCGTCTC**GCAAA**
BB3650		phage-related hyp. protein	72.4	313	**GGAACTG**ATCATCTGCGATGACA**TCCAG**
BB3665	*fabG*	3-oxoacyl-[acyl-carrier protein] reductase	72.4	69	**GGAACTA**GGGCAATGCATTTAA**CCACA**

Gene	Name	Product	Score	Dist. to ORF	Promoter
BB3821		put. exported protein	72.4	110	**TAAACAG**AATAGATTATACGAA**TCAAA**
BB4223		put. gamma-glutamyltranspeptidase	72.4	487	**TGAACTA**CAACACGTTCGACATTT**TCATG**
BB4240		put. inner membrane permease	72.4	460	**GGAACTG**ATTTCGCGCCAGCAG**CCCAA**
BB4258		put. Mg^{2+} transporter	72.4	128	**GAAACAA**GATAGTCCATCGGCGTC**TCCAA**
BB4362		put. exported protein	72.4	577	**AGAACGA**GCGCGGCATTCCCGCC**TCGAA**
BB0362	*scoA*	put. succinyl-CoA:3-ketoacid-coenzyme A transferase	72.1	678	**CGAACTG**ATGGCGCAGCCCCAGTT**TCCCA**
BB0402		put. permease	72.1	65	**CAAACGT**TTCATAATCGACTTCGA**TCAGA**
BB0422		put. acetyltransferase	72.1	141	**CGAACTG**CACGCGGCGTGCCTG**TCCCA**
BB1221		hyp. protein	72.1	222	**TGAACAT**CGATGGAGTTTTTCA**TGAAA**
BB2173	*ribF*	riboflavin biosynthesis protein	72.1	46	**AAAACTT**TCCAAAAGCGGGTCAAC**CCAAA**
BB2252		put. membrane protein	72.1	740	**GGAAGTC**CTCGGCCTTGCCGCCT**TCGAA**
BB2423		put. membrane protein	72.1	889	**GGACCTG**CGCTTCAATACCGACG**TCGAA**
BB3228		put. membrane protein	72.1	870	**GGACCTG**TGGCTGCACAGCAATG**TCGAA**
BB3329		prob. cytochrome C oxidase	72.1	98	**GGAACGA**AGGCCGCTTGACGCAAA**TCAAG**
BB4336		put. GntR-fam. transcript. reg.	72.1	126	**GGAACGT**CCTGCTGCATGGGTT**TCATT**
BB4514		put. tetrapyrrole methylase	72.1	924	**GGACCTG**CCGCTCGAAGATCTCG**TCGAA**
BB4681		hyp. protein	72.1	703	**GGAACGA**TTTCATCGCACTCGA**TCAAC**
BB4802	*glmS*	glucosamine--fructose-6-phosphateaminotransferase	72.1	832	**TGAACGT**CGACGGCGGCGCGATC**TCGAT**
BB0110		put. adhesin	72.1	139	**CGAAATTT**CCTCTCTGCCATTTCC**TCCGA**
BB1358		put. membrane protein	72.1	638	**CGAACTT**CTCGATCTCGGGAAA**GCTGA**
BB2042	*ihfA*	integration host factor	72.1	907	**GGAAGTG**GTCAATTACAGCTTCG**TCGAA**
BB2159		put. exported protein	72.1	830	**GGAACGA**CGCCGGCTATTTCGGCG**TCAAG**
BB2247		put. membrane protein	72.1	104	**TGAACAT**CGGCTACGCGTACCGCT**TCTAG**
BB2881		put. bacterioferritin comigratory protein	72.1	836	**TGAACAT**CACGCGCGTCACCCACG**TCTAC**
BB3267		put. exported protein	72.1	866	**GAAACGA**CGAATGCGTGTCGCCC**TCGAA**
BB3392		conserved hyp. membrane protein	72.1	59	**AGAACAA**GTAATGGAAAAATCGC**TCGAA**
*BB3424		fimbrial protein	72.1	54	**TCAACTT**TTAATATTCTTGTAA**TCGCA**
BB3573		put. exported protein	72.1	951	**GGACCTG**GCCGTCACGGTACGTA**TCGAA**
BB3811	*risS*	sensor kinase protein	72.1	97	**GCAAATT**GATCGAACCCAACCCG**TCCAA**
BB4006	*ruvB*	holliday junction DNA helicase	72.1	649	**GGAAGTC**GAGTTCGCCGACCTCG**TCGAA**

Among the candidate promoters identified include sequences upstream of genes encoding many envelope and exported proteins, some transcriptional regulators, and many proteins of unknown function. Also included among the candidate promoters is a promoter upstream of *fam*, validating that our program can identify known SigE-regulated promoters in *B. bronchiseptica* (Chapter 2). Some of the candidate SigE-dependent promoters are upstream of genes orthologous to genes with σ^E-dependent promoters in *E. coli*, such as

fam (σ^{32}, described above), *rseA*, *mucD* (*E. coli degP*), *ftsz*, and *bb4517* (*E. coli yraP*) [48]. This suggests that, as with other species where σ^{E}-dependent genes have been identified or predicted, there is some overlap in the regulon between different bacteria. Some of the genes identified as having putative SigE-regulated promoters were also identified by ROMA, including *hfq* (sRNA-binding protein), *bb3678* (conserved hypothetical protein), *bb4518* (putative thioredoxin), *bb2177* (fatty acid desaturase) and *mucD* (serine protease) (Table 4-3). Because this program identifies sequences with similarity to the position weight matrix, it is likely that some of the genes identified as having SigE-regulated promoters are false positives, and are not dependent on SigE for transcription in vivo; it is also likely that some SigE-regulated promoters were missed because their similarity to the consensus promoter is lower than the threshold used in this analysis. For example, the promoter upstream of *E. coli rseA*, which has been experimentally verified, would score just below the median score of the promoters used to build the PWM, and therefore would not have been identified in this analysis. As indicated in the previous section, further work, some of which is described later in this chapter, is necessary to confirm identified genes as members of the SigE regulon.

Combining methods to predict the SigE regulon

Over 30 of the candidate SigE-regulated promoters are upstream of genes that also have increased expression in RB50Δ*rseAB*, indicated by asterisks in Table 4-4. Only five genes are in common to all three of the methods used in this study (Fig. 4-3), including *bb2228* (phage-related conserved hypothetical protein), *bb3678* (conserved hypothetical protein), *mucD* (serine protease), *bb4518* (putative thioredoxin), and *fam* (σ^{32}). There is

some overlap between the three methods, but a significant number of genes are identified by one method only. The Helmann group previously determined that no one method is sufficient to accurately determine the regulon of a sigma factor, and rather a combination of methods not only provides multiple pieces of evidence for a particular regulon member, but also identifies regulon members missed by other methods [201].

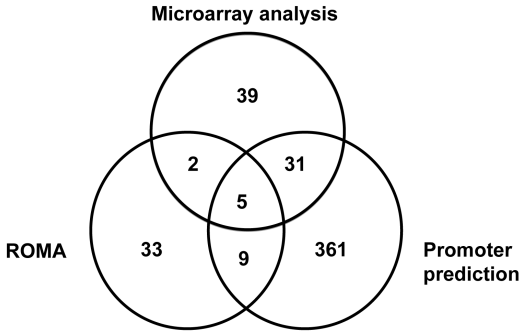


Figure 4-3: Combining all three methods to determine the SigE regulon. Genes identified by microarray analysis, ROMA, and bioinformatic prediction of SigE-regulated promoters are compared.

In order to build a short list of candidate regulon members, I selected the strongest candidates from each method: genes with at least two-fold higher expression in RB50Δ*rseAB* compared to RB50, those with at least a 2.5-fold difference between RNA transcribed by ESigE compared to core RNAP in the ROMA analysis, and those with the strongest-scoring promoters (at least 80% identity to consensus). It is likely that the SigE system only indirectly regulates some of the candidate members identified by microarray analysis (Table 4-2), that some of the SigE-dependent promoters identified by ROMA and predicted bioinformatically are simply good binding sites for the sigma factor and not regulated in vivo, and that there are still some SigE-regulated genes missed by all three

methods. Therefore, this is only a list of potential regulon members yet to be confirmed as SigE-regulated in *B. bronchiseptica* (Table 4-5).

Table 4-5: Putative SigE regulon. Candidate regulon members compiled from all three methods of identification, grouped by putative function. The gene, product, and method by which each gene was considered a candidate are listed. A= microarray, P= promoter prediction, R= ROMA, M=Mass Spectrometry

Predicted functions	Gene		Product	Method
Transport, Export, and Secretion				
Autotransporters	BB0419	*sphB1*	autotransporter subtilisin-like protease	A,P
	BB0452		autotransporter	A
	BB1864	*vag8*	autotransporter	A
	BB2033	*bapC*	put. autotransporter	A
Type III Secretion	BB1609	*bscF*	put. type III secretion protein	A
	BB1610	*bscE*	hyp. protein	A
	BB1617	*bsp22*	put. secreted protein	A,P
	BB1620	*bopD*	put. outer protein D	A
	BB1629	*bscO*	put. type III secretion protein	A
	BB1645	*btrW*	put. anti-sigma factor	P
Secretion	BB1359	*secF*	protein-export membrane protein	A
	BB1360	*secD*	protein-export membrane protein	A
	BB2383		put. HlyD-family secretion protein	P
	BB3739		put. integral membrane protein	A,P
Transport	BB2504		put. membrane transport ATPase	P
	BB2528	*acrB*	acriflavine resistance protein B	A
	BB3005		prob. ATP-binding ABC transporter prot.	P
	BB3826	*bfrD*	prob. TonB-dependent receptor	A
	BB4235		put. TolR-like translocation protein	A
Porin	BB3993	*ompQ*	outer membrane porin protein OmpQ	A
	BB4490		outer membrane porin	A,M
Export	BB0098		put. exported protein	A
	BB0918		put. exported protein	P
	BB1292		put. exported protein	A
	BB1931		put. exported protein	A
	BB2245		put. exported protein	A
	BB2734		put. exported protein	R
	BB2885		put. exported protein	A
	BB3051		put. exported protein	R
	BB3068		put. exported protein	A,P
	BB3241		put. exported protein	A
	BB3264		put. exported protein	A,P
	BB3990		put. exported protein	P
	BB4265		put. exported protein	A,P
	BB4491		put. exported protein	A

Predicted functions	Gene		Product	
Export (cont.)	BB4517		put. exported protein	R
	BB4939		put. exported protein	A
Other envelope proteins				
Adhesion	BB1936	*fhaL*	adhesin	A
	BB3424		fimbrial protein	A
Lipoproteins	BB3880		put. lipoprotein	R
LPS biosynthesis	BB4516		prob. phosphosugar isomerase	R
	BB4192	*lpxC*	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	P
Protein folding	BB0061		thiol:disulfide interchange protein	R
	BB1733	*ppiB*	peptidyl-prolyl cis-trans isomerase B	A
	BB4940	*dsbA*	thiol:disulfide interchange protein	A,P
Other redox	BB3025	*trxC*	thioredoxin 2	A,P
	BB4518		put. thioredoxin	A,R,P
	BB3766		cons. hyp. protein	A
Peptidoglycan	BB0168		put. penicillin-binding protein	A,R
	BB3097	*mrcA*	penicillin-binding protein 1A	A
	BB3801		put. intracellular septation protein	A
	BB3802		BolA-like protein	A
Other membrane proteins	BB0783		put. membrane protein	A
	BB1289		put. integral membrane protein	A
	BB1368		put. membrane protein	A
	BB3765		put. membrane protein	A
	BB3846		put. membrane protein	A
	BB4017	*osmB*	osmotically inducible lipoprotein B	A
	BB4284		put. membrane protein	A,P
	BB4831		put. bifunctional protein	R
Other proteins				
Hydrolases	BB0975		prob. hydrolase	A
	BB1933		put. hydrolase	A
Proteases	BB1894		heat shock protease	A,P
	BB3749	*mucD*	serine protease	A,R,P
	BB3932		put. zinc protease	A
Stress Response	BB0501	*htpG*	heat shock protein	A
	BB2667		put. universal stress protein	A
	BB3750	*rseB*	put. sigma factor reg. protein	R
	BB3751		put. membrane protein	R,P
	BB4506	*rpoN*	prob. sigma(54) modulation protein	A
	BB4835	*fam*	RNA polymerase sigma-32 factor	A,R,P
Transcription/ Translation	BB1282		two component system response reg.	R
	BB1837		putative ECF sigma factor	P
	BB1932	bph3	put. DNA-binding protein (histone)	A
	BB3161	*rpoD*	RNA polymerase sigma factor 70	P
	BB3170	*hfq*	put. RNA-binding reg. protein	R,P
	BB3748	*lepA*	GTP-binding protein	R
	BB4127		put. LysR-family transcriptional reg.	P
	BB4492		prob. LysR-family transcriptional reg.	A

Predicted functions	Gene		Product	
Other metabolism	BB0247	*def*	polypeptide deformylase	R,P
	BB0611		acyl-CoA dehydrogenase	A
	BB1844	*acnA*	put. aconitate hydratase	R,P
	BB2177		prob. fatty acid desaturase	R,P
	BB2438		prob. 3-hydroxybutyryl-CoA dehydrogenase	P
	BB2762		conserved hyp. protein	R
	BB3677	*gltA*	citrate synthase	R
	BB3756	*fabG*	3-oxoacyl-[acyl-carrier protein] reductase	P
	BB4348	*glyA*	serine hydroxymethyltransferase	P
	BB4402	*pyrC*	dihydroorotase	P
Unknown function	BB0355		conserved hyp. protein	R
	BB0409		conserved hyp. protein	P
	BB0499		hyp. protein	P
	BB1143		conserved hyp. protein	A
	BB1291		conserved hyp. protein	A
	BB1956		hyp. protein	A
	BB2050		hyp. protein	P
	BB3108		hyp. protein	A,P
	BB3419		globin-like protein	A
	BB3678		conserved hyp. protein	A,R,P
	BB4018		conserved hyp. protein	A
	BB4343		conserved hyp. protein	P
	BB4652		hyp. protein	A
	BB4905		conserved hyp. protein	A
Phage-related proteins	BB1709		phage-related hyp. protein	A
	BB2199		put. phage repressor protein	P
	BB2213		phage-related put. DNA binding protein	A
	BB2214		put. phage terminase	A
	BB2228		phage-related conserved hyp. protein	A,R,P
	BB2232		phage-related conserved hyp. protein	A
	BB2233		phage-related conserved hyp. protein	A
	BB3491		phage-related conserved hyp. protein	A
	BB3493		phage-related conserved hyp. protein	A
	BB3496		phage-related conserved hyp. protein	A
	BB3505		phage-related conserved hyp. protein	A
	BB3520		phage-related hyp. protein	P

Confirming SigE-dependent transcription of candidate regulon members

To confirm direct SigE regulation, I PCR-amplified the promoter regions upstream of some of these candidate regulon members. These were used as templates for in vitro transcription assays to confirm SigE-dependent transcription of these promoters. Figure 4-4 shows the results of these assays. I have confirmed SigE-dependent promoters upstream

of *fam* (Chapter 2), *hfq, rseA, bb3108* (hypothetical protein)*, bb1282* (two component response regulator), *bb1837* (ECF sigma factor), *bb4281* (putative membrane-bound lytic glycosylase)*, bb0098* (putative exported protein)*,* and *bb3842* (putative outer membrane porin). I have also tested the promoter regions of 27 other genes (Table 4-6). Preliminary results do not indicate SigE-dependent transcription from the promoter regions upstream of these genes. However, ongoing work will confirm the presence or absence of SigE-dependent promoters for these and the genes encoding other putative regulon members.

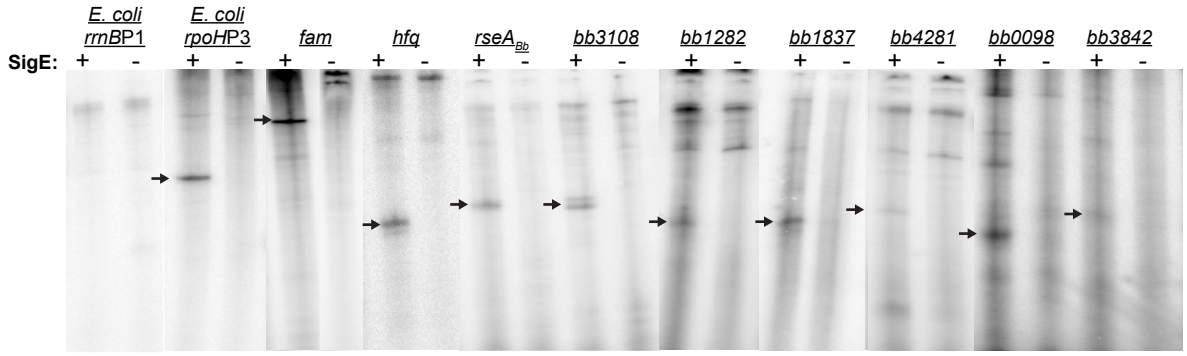


Figure 4-4: SigE-dependent promoters in *B. bronchiseptica*. In vitro transcription reactions with either ESigE (+) or core RNAP alone (-), and using the promoter region of the indicated gene as a template. The σ^{70}-dependent promoter *rrnB*P1 was used as a negative control for SigE-dependent transcription (far left lanes)., and the previously described SigE-transcribed promoters *E. coli rpoH*P3 and *B. bronchiseptica fam* were used as positive controls.

Table 4-6: Promoters tested for in vitro transcription. This table lists the gene associated with the promoter region tested, the reason it was chosen for testing, and whether or not it has been confirmed by in vitro transcription. A=microarray analysis, R=ROMA, P=promoter prediction, S=part of *sigE* locus, M=mass spectrometry. + indicates SigE-dependent transcription, - indicates no observed SigE-dependent transcription, +/- indicates preliminary results not yet reproduced.

Gene		Product	Reason tested	SigE-transcribed
BB0098		Putative exported protein	A	+
BB1282		two component system response regulator	R	+
BB1837		putative ECF-family sigma factor	P	+
BB3108		hypothetical protein	A,P	+
BB3170	*hfq*	putative RNA-binding regulatory protein	R,P	+
BB3751	*rseA*	putative membrane protein	S,P	+
BB3842		outer membrane porin protein precursor	M	+
BB4281		putative membrane-bound lytic mureintransglycosylase D	P	+
BB4835	*fam*	RNA polymerase sigma-32 factor	A,R,P	+
BB0070	*pbp1a*	penicillin-binding protein 1A	A,P	-
BB0419	*sphB1*	autotransporter subtilisin-like protease	A,P	-
BB1089		putative enoyl-CoA hydratase	A,P	-
BB1286		cytochrome 0 ubiquinol oxidase	P	-
BB1366	*prn*	pertactin precursor	A,P	-
BB1617	*bsp22*	putative secreted protein	A,P	-
BB1635	*bscU*	putative type III secretion protein	P	-
BB1894		heat shock protease	P	-
BB1975	*rnr*	putative ribonuclease R	P	-
BB2177		probable fatty acid desaturase	P	-
BB2256	*lon*	ATP-dependent protease La	R	-
BB2994	*bvgA*	virulence factors transcription regulator	P	-
BB3068		putative exported protein	R	+/-
BB3172		putative GTP-binding protein	P	-
BB3678		conserved hypothetical protein	A,P	-
BB3739		putative integral membrane protein	P	-
BB3749	*mucD*	serine protease	S,R,P	+/-
BB3752	*sigE*	RNA polymerase sigma factor	S,P	-
BB3754	*fabF*	3-oxoacyl-[acyl-carrier-protein] synthase II	S,P	-
BB3756	*fabG*	3-oxoacyl-[acyl-carrier protein] reductase	P	-
BB3932		putative zinc protease	S	-
BB4193	*ftsZ*	cell division protein FtsZ	P	-
BB4284		putative membrane protein	S	-
BB4490		outer membrane porin	A,M	-
BB4518		putative thioredoxin	R,P	-
BB4894	*ptxC*	pertussis toxin subunit 3	P	-
BB4940	*dsbA*	thiol:disulfide interchange protein	S	-

Role of the SigE regulon at the cell envelope and in pathogenesis

Cells with constitutively active SigE have an altered cell envelope protein composition

Expression of multiple genes encoding cell envelope and envelope-related proteins increases when SigE activity is high. To determine whether this has any impact on the composition of the cell envelope, I separated the outer membrane, inner membrane, and cytoplasm/periplasm of RB50, RB50Δ*sigE*, and RB50Δ*rseAB* via the differential solubility of the outer membrane in sarcosyl. The outer membrane fraction of RB50 and RB50Δ*sigE* are indistinguishable (data not shown). However, the outer membrane protein composition of RB50Δ*rseAB* differs from RB50 by at least one protein band around a molecular weight of 35kDa (Fig. 4-5A). I sent this band for analysis by mass spectrometry, and it was identified as BB4490 and/or BB3842, both of which are outer membrane porins. While some porins of *E. coli* and other organisms have been studied extensively, the role of porins in *B. bronchiseptica* has not been examined in any detail. Outer membrane porins (OMPs) are typically trimeric β-barrel proteins that span the outer membrane, allowing hydrophilic molecules to cross the very hydrophobic phospholipid bilayer [16].

BB3842 is an ortholog of the *B. pertussis* "40K porin", which is constitutively and abundantly expressed in both the Bvg^+ and Bvg^- phases in *B. pertussis*, and is thought to have a small channel and be selective for anions [202, 203]. Outside of this and some structural data [204], little is known about what role this porin actually plays in the cell. BB4490, on the other hand, is not conserved in *B. pertussis*, but has orthologs in *B. avium* and *B. parapertussis*. I have also identified potential orthologs in the genomes of other bacteria, such as *Burkholderia* spp., *Shewanella* spp., and *Nitrosomonas europaea*, but again, none has been examined experimentally.

Because BB3842 is orthologous to the most abundant porin identified in *B. pertussis*, it is possible that BB3842 actually corresponds to the large predominant band above the smaller, differentially expressed band sent for analysis. If a small amount of this more abundant protein were present in the sample, it would be difficult to separate the two. Consistent with this interpretation, BB3842 is predicted to be a slightly larger protein (41.3 kDa) than BB4490 (39.6 kDa), which could correspond to the larger, more abundant protein, and the smaller, less abundant protein, respectively. However, there is significant sequence similarity between the two porins (73.5%); therefore, it cannot be ruled out that both BB3842 and BB4490 are present in the protein band differentially expressed between RB50 and RB50Δ*rseAB*.

bb4490 has increased expression in cells with high SigE activity (10.53-fold, Table 4-1), but *bb3842* does not have significantly changed expression (-1.2-fold, Table 4-1) in RB50Δ*rseAB*. Neither gene was predicted to have a SigE-regulated promoter from my bioinformatic analysis. However, even in *E. coli*, biologically relevant σ^E-dependent promoters fall below the threshold of many bioinformatic analyses. To test whether the genes encoding either of these porins is regulated directly by *B. bronchiseptica* SigE, 400 basepairs upstream of the translation start site was PCR-amplified and used in in vitro transcription reactions. Preliminary results indicate a possible SigE-regulated promoter upstream of *bb3842*, but not *bb4490* (Fig. 4-4 and data not shown). Ongoing work will confirm the presence or absence of a SigE-regulated promoter upstream of either of these genes.

In other bacterial systems, downregulation of certain OMPs corresponds to a gain of resistance to some antibiotics, particularly β-lactams, which can use porins to cross the

envelope into the cell. This is different from our results in *B. bronchiseptica*, where under conditions when SigE activity is high, cells are more resistant to β-lactam antibiotics, despite upregulation of at least one porin [16, 205]. There is some evidence that replacing certain OMPs with other, less permissive, porins, is one mechanism that contributes to antibiotic resistance [16]. Whether BB4490 or BB3842 may fall into this class has yet to be determined. An increase in porin expression when SigE activity is high is also different from the *E. coli* or *Salmonella* σ^E systems. In *E. coli*, constitutive activation of σ^E leads to a decreased accumulation of OMPs in the outer membrane because σ^E transcribes sRNAs such as MicA and RybB, which target OMP mRNAs for degradation [77]. Identification at least one porin with increased expression in a strain with high SigE activity suggests different specific roles for the SigE system in *B. bronchiseptica* compared to *E. coli* or *S. enterica* serovar Typhimurium σ^E in regulation of OMPs.

Different antibodies are produced during RB50Δ*rseAB* infection compared to RB50 infection

To determine whether differential expression of envelope proteins may affect interactions of *B. bronchiseptica* with the host immune system, outer membrane proteins from RB50, RB50Δ*rseAB*, or RB50Δ*wbm*, which lacks the O-antigen component of LPS, were probed with naïve serum, serum collected from C57BL/6 mice inoculated with RB50, or serum collected from C57BL/6 mice inoculated with RB50Δ*rseAB*. LPS is a particularly antigenic component of the cell envelope [14]; in strains lacking O-antigen, a large smear recognized by immune serum is missing, aiding visualization of antigenic proteins (Fig. 4-5B). No proteins were recognized by naïve serum (data not shown). Different proteins

were recognized by immune serum from infection with either RB50 or RB50Δ*rseAB*, as indicated by the arrows in Fig. 4-5B. Interestingly, the proteins against which antibodies were produced during infection with each bacterial strain are still present in the outer membranes of all *B. bronchiseptica* strains tested. This suggests that when SigE is constitutively active, it may not be differential expression of antigenic proteins, but differences in availability of proteins to the host immune system for antigen presentation that promotes a different set of antibodies generated against RB50 and RB50Δ*rseAB*. To ensure that differences seen between serum from mice infected with either RB50 or RB50Δ*rseAB* are not due to mouse-to-mouse variation, I performed western blots on outer membrane protein samples using serum from four additional mice infected with RB50, and will perform western blots using serum from multiple mice infected with RB50Δ*rseAB* when it is available. There were few discernible differences between samples probed with serum from different mice infected with RB50 (data not shown), indicating that the differences observed are not due to variations between mice. This provides further evidence that there are differences in the envelope protein composition of cells when SigE activity is high.

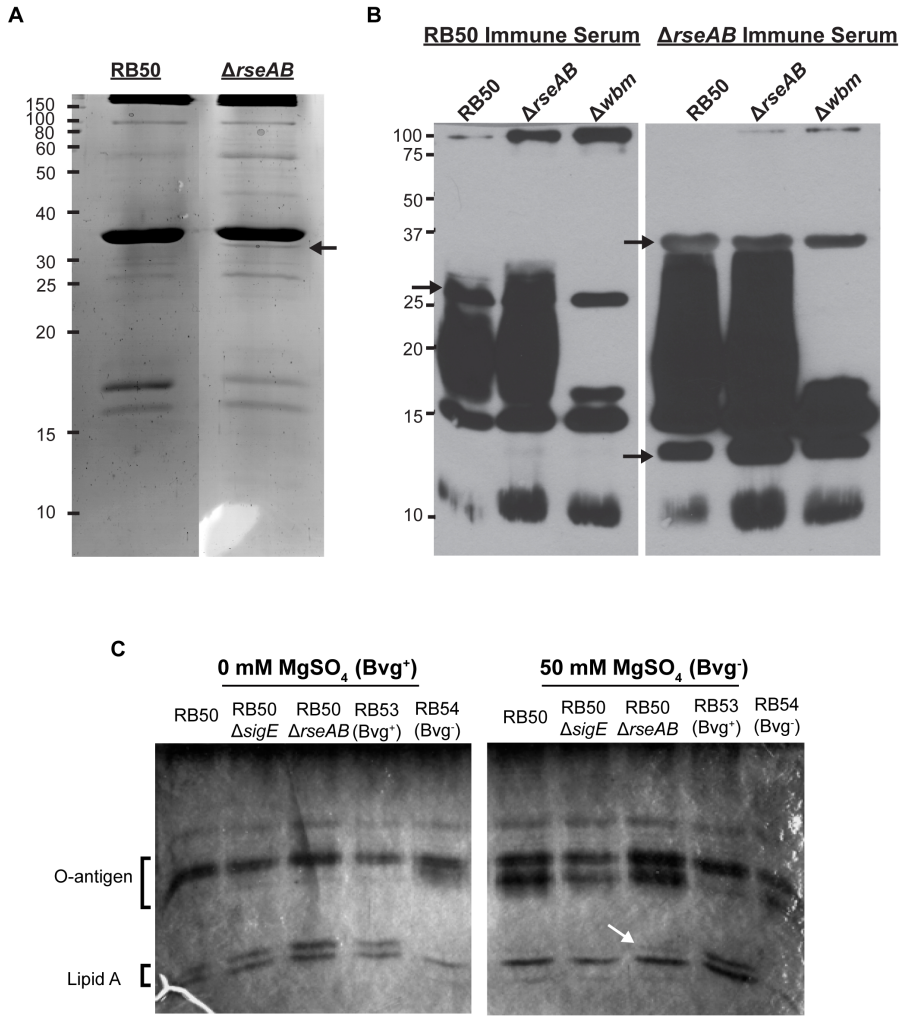


Figure 4-5: RB50Δ*rseAB* has an altered cell envelope composition. (A) At least one protein is more highly expressed in RB50Δ*rseAB, right,* compared to RB50, *left*. A differentially expressed protein indicated with an arrow. (B) Different antibodies are produced in by C57BL/6 mice in response to infection with RB50Δ*rseAB, right*, than to infection with RB50, *left*. Consistent differences between serum recognition of outer membrane proteins are indicated with arrows. (C) LPS from samples grown under Bvg^+ conditions, *left*, and under Bvg^- conditions, *right*. A white arrow indicates the band of interest.

Cells with high SigE activity have minor differences in lipopolysaccharides

In *E. coli* and other bacteria, the core σ^E regulon includes genes important for proper synthesis and assembly of lipopolysaccharides (LPS). I have also identified a couple of genes with roles in LPS biosynthesis as putative SigE regulon members in *B.*

bronchiseptica (Table 4-5). *B. bronchiseptica* LPS consists of a lipid A core, complex branched core oligosaccharides, and a complex trisaccharide, as well as O-antigen [12]. In *B. bronchiseptica*, various components of LPS, such as O-antigen and modifications to lipid A, including PagP-mediated palmitoylation of lipid A, are expressed in a BvgAS-dependent manner [206]. The BvgAS two-component phosphorelay system regulates most of the virulence factors in the bordetellae. When BvgAS is activated, the cells are considered Bvg^+, and express most known virulence factors. When growing at 25 °C, or in the presence of $MgSO_4$ or nicotinamide, the expression of many virulence factors is repressed, such as the type three secretion system (T3SS) and adenylate cyclase toxin (ACT), and *B. bronchiseptica* instead expresses flagella and O-antigen [27].

To determine whether *B. bronchiseptica* SigE also influences the composition of the LPS, I isolated LPS from RB50, RB50Δ*sigE*, RB50Δ*rseAB*, a Bvg^+-locked strain (RB53) and a Bvg^--locked strain (RB54) by overnight proteinase K digestion. Under Bvg^+-conditions, RB50, RB50Δ*sigE*, and RB50Δ*rseAB* have similar LPS profiles, including decreased expression of O-antigen and introduction of a PagP-mediated palmitoylation of lipid A, visible as the upper band of a doublet at the bottom of a silver-stained gel (Fig. 4-5C, [206]). When the cells are modulated to Bvg^- with the addition of 50mM $MgSO_4$, expression of O-antigen increases, and PagP-mediated modification to lipid A is no longer present in RB50 and RB50Δ*sigE*, but is still somewhat visible in RB50Δ*rseAB* (Fig. 4-5C). This suggests that constitutive activation of the SigE system either directly affects this modification, or indirectly affects it through perturbations to the envelope, which can activate PagP post-transcriptionally (A. Preston, personal communication).

Possible roles of SigE regulon members in pathogenesis

Cytotoxicity

B. bronchiseptica is cytotoxic to macrohpages, and this cytotoxicity is dependent on a functional T3SS and adenylate cyclase toxin [163, 193]. As previously described, both a strain lacking *sigE* (light grey bars) and a strain lacking *rseA* and *rseB* (dark grey bars) are less cytotoxic to macrophages than wild-type RB50 (white bars) when pre-incubated in Stainer-Scholte media (Fig. 2-7, Fig. 3-8, Fig. 4-6, *left*). When *B. bronchiseptica* is incubated with bovine serum prior to incubation with macrophages, its cytotoxicity toward macrophages increases, suggesting that *B. bronchiseptica* recognizes components of the bloodstream and upregulates factors important for cytotoxicity (S. Hester, unpublished results). Ongoing work is investigating what factor or factors are responsible for this increase in cytotoxicity. Cells lacking *sigE* still increase cytotoxicity to nearly wild-type levels in response to serum (Fig. 4-6, *right*). RB50Δ*rseAB*, however, is not stimulated by serum to the same extent, indicating that constitutive activation of SigE interferes with the ability of RB50 to properly sense and respond to these signals.

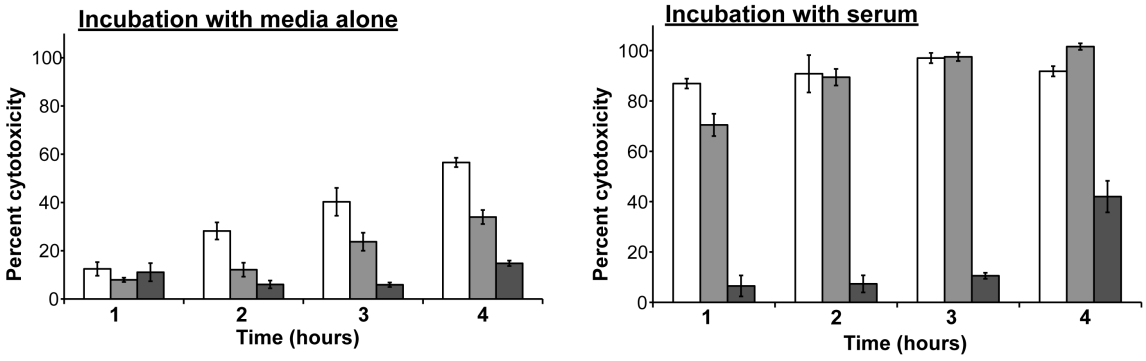


Figure 4-6: RB50Δ*rseAB* does not increase cytotoxicity in response to serum to the same extent as RB50. RB50 (white bars), RB50Δ*sigE* (light grey bars), or RB50Δ*rseAB* (dark grey bars) were incubated in media, *left*, or in the presence of 100% fetal bovine

serum, *right*, for 1 hour, then added to RAW 264.7 murine macrophages. Percent cytotoxicity of each strain as measured by LDH release, was measured after 1, 2, 3, or 4 hours of incubation with macrophages.

To determine what part(s) of the SigE regulon might be involved in this phenomenon, I compared the genes differentially expressed in RB50Δ*rseAB* (Tables 4-1 and 4-2) with genes that are regulated positively or negatively by incubation with bovine serum (S. Hester and T. Nicholson, unpublished results) (Table 4-7). Few genes were significantly, differentially regulated under these conditions. Both subunits of sulfate adenylyltransferase, which is involved in purine and sulfur metabolism, as well as a membrane component of an ABC transport system and a putative exported hydrolase all had higher expression in cells incubated with serum, but lower expression in RB50Δ*rseAB*. The role these genes might play in cytotoxicity is unclear. It is possible that changes in gene expression of some important factor, below the two-fold threshold set by our analysis, are sufficient to prevent RB50Δ*rseAB* from robustly increasing cytotoxicity in response to serum. It is also possible that through changes in the cell envelope composition, RB50Δ*rseAB* is no longer able to correctly assemble, present, or secrete some unknown factor important for sensing and responding to serum, or that some other post-transcriptional regulation is involved in response to serum that could not be identified from gene expression analysis.

Table 4-7: Genes differentially expressed in serum compared to RB50Δ*rseAB*

Gene	Name	Product	Fold change RB50Δ*rseAB* vs. RB50
Higher expression in serum, lower expression in RB50Δ*rseAB*			
BB2907	*livM*	high-affinity branched-chain amino acid transport system, permease	-5.08
BB3447	*cysN*	sulfate adenylyltransferase subunit 1	-3.7
BB3448	*cysD*	sulfate adenylyltransferase subunit 2	-3.53
BB3764		put. exported hydrolase	-2.41
Lower expression in serum, higher expression in RB50Δ*rseAB*			
BB0098		put. exported protein	20.32
BB4652		hyp. protein	5.29

Interactions with BvgAS

We previously demonstrated that cells with constitutive SigE activity are defective in colonization and persistence in the lower respiratory tract (Chapter 3). To determine if differential regulation of some virulence factors may account for this defect, I compared the genes that are differentially regulated in RB50Δ*rseAB* with genes identified as maximally expressed in the Bvg^+ or Bvg^- phase. As mentioned above, the BvgAS two-component system regulates most known virulence factors in the bordetellae, which tend to be maximally expressed in the Bvg^+ phase and repressed in the Bvg^- phase [27]. In particular, I was interested in genes that are known to be expressed in the Bvg^+ phase, but have decreased expression in cells with high SigE activity, even under typical Bvg^+ conditions (37 °C, Stainer-Scholte media). Gene expression data from Table 4-1 and 4-2 (increased and decreased expression in RB50Δ*rseAB* compared to RB50) were compared to results from Cummings, et al. [141], which identified genes with increased expression in a Bvg^+ phase-locked strain (RB53) compared to a Bvg^- phase-locked strain (RB54), and to results from Nicholson, et al. [182], which identified genes with increased expression in

RB50 grown under Bvg^+ conditions compared to a Bvg^- phase-locked strain (RB54). I identified six genes with decreased expression in Δ*rseAB* that were identified as Bvg-activated genes in both published studies (Table 4-8). Interestingly, half of the genes identified in this analysis were adhesins: *fim2*, *fim3*, and *bipA*.

Fimbriae and other adhesins are required for colonization and persistence in the lower respiratory tract, particularly the trachea [147]. Perhaps the misregulation of these genes could contribute to the defects in colonization and persistence observed in C57BL/6 mice (Chapter 3). This is among the first evidence that a transcription factor other than BvgAS contributes to regulation of major virulence factors, and suggests that SigE, in part, helps to fine-tune a subset of virulence factor expression under particular conditions or specific times during infection. Future work will be needed to confirm differential regulation of these virulence factors by the SigE system (discussed in Chapter 6).

Table 4-8: Genes differentially expressed in Bvg^+ compared to RB50Δ*rseAB*

Gene	Name	Product	Fold change RB50Δ*rseAB* vs. RB50
Higher expression in Bvg^+, lower expression in RB50Δ*rseAB*			
BB1658	*fim3*	serotype 3 fimbrial subunit	-5.84
BB1696		phage-related hyp. protein	-2.60
BB1702		phage-related conserved hyp. protein	-2.66
BB3674	*fim2*	serotype 2 fimbrial subunit	-3.55
BB3764		put. exported hydrolase	-2.41
BB2303	*bipA*	put. outer membrane ligand binding prot.	-2.49
Lower expression in Bvg^+, higher expression in RB50Δ*rseAB*			
BB0975		prob. hydrolase	4.99
BB1733	*ppiB*	peptidyl-prolyl cis-trans isomerase B	3.41
BB1894		heat shock protease	3.96
BB2667		put. universal stress protein	3.49
BB3108		hyp. protein	5.43
BB3264		put. exported protein	7.88
BB3678		conserved hyp. protein	2.41
BB3801		put. intracellular septation protein	2.87
BB4491		put. exported protein	3.18
BB4506		prob. sigma(54) modulation protein	2.74

Gene	Name	Product	Fold change RB50Δ*rseAB* vs. RB50
BB4652		hyp. protein	5.29
BB4835	*rpoH*	RNA polymerase sigma-32 factor	4.19
BB4905		conserved hyp. protein	3.11

Conservation of SigE-regulated promoters in the classical bordetellae

The *sigE* operon is nearly 100% conserved in *B. bronchiseptica* and the closely related *B. pertussis* and *B. parapertussis*. However, even in very closely related bacteria, σ^E systems can be adapted for a particular species' requirements, based on different environmental niches and/or presence of redundant available response systems for a particular stress condition (Chapter 1). To begin to explore whether parts of the SigE regulon might be conserved in the human pathogens *B. pertussis* and *B. parapertussis*, I used the promoter prediction program described earlier to predict SigE-regulated promoters in *B. pertussis* and *B. parapertussis.* A comparison of the promoters predicted for each species is shown in Fig. 4-7. Over 10% of the promoters predicted in the three species are common to all three, indicating some possible conservation of function for the SigE regulon. Included in the SigE-regulated promoters predicted for all three of the classical bordetellae are those upstream of *hfq, bb4281* (putative membrane-bound transglycosylase)*,* and *mucD,* which I have already confirmed as SigE-regulated in *B. bronchiseptica*.

13 of the 77 genes identified by microarray analysis, 4 of those identified by ROMA, and one identified by both methods (*mucD*) also have predicted promoters in all three genomes. Three genes also identified by other methods (microarray, ROMA) are

common to both *B. bronchiseptica* and *B. pertussis*, and four genes are common to both *B. bronchiseptica* and *B. parapertussis*, suggesting some limited overlap between genes identified as candidate SigE regulon members by multiple methods. Most of the promoters predicted for each species are unique to that species. This observation might indicate that while some regulon members may be conserved between the classical bordetellae, each has adapted the SigE system to transcribe a subset of genes, based on the set of conditions encountered by each and the array of response systems expressed by each in order to respond. Further work is necessary to experimentally confirm SigE regulon members in each of the bordetellae.

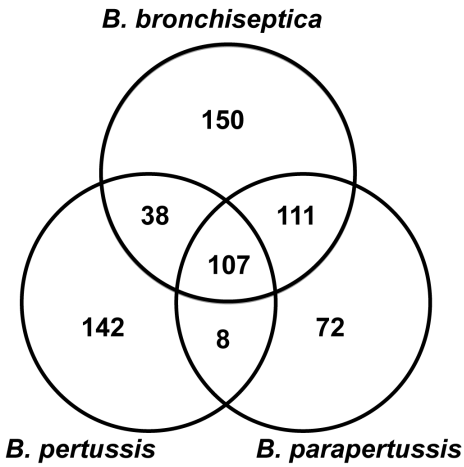


Figure 4-7: Comparison of candidate SigE-regulated promoters in the bordetellae

In this work, I have identified candidate members of the SigE regulon in *B. bronchiseptica* through a variety of methods, and demonstrated possible roles for this regulon in expression of genes at the cell envelope and important in stress response and pathogenesis. It will be of interest to further determine what specific roles members of this

regulon play in the infection process, and the extent to which the regulon is conserved among the classical bordetellae.

Materials and Methods

Bacterial strains and growth

A list of strains and plasmids can be found in Table 4-9. *B. bronchiseptica* strains are derivatives of RB50 [167]. *B. bronchiseptica* was maintained on Bordet-Gengou (BG) agar (Difco) containing 10% defibrinated sheep's blood (Hema Resources) and 20 μg/ml streptomycin. In liquid culture, *B. bronchiseptica* was grown at 37 °C in Stainer-Scholte broth [168] with aeration. *E. coli* strain XQZ001 was maintained on LB agar containing 30 μg/mL kanamycin, and grown in liquid culture in LB broth with shaking at either 30 °C or 25 °C, as indicated below.

Table 4-9: Strains and plasmids

	Strain name	Genotype	Source, Reference
B. bronchiseptica	RB50	RB50	[167]
	RB50Δ*sigE*	RB50Δ*sigE*	Chapter 2
	RB50Δ*rseAB*	RB50Δ*rseAB*	Chapter 3
	RB53		[167]
	RB54		[167]
	RB50Δ*wbm*		[172]
	WD3	RB50Δ*bscN*	[163]
	AVS	RB50Δ*cyaA*Δ*bscN*	[149, 173]
E. coli	XQZ001	BL21(DE3) *slyD*::kan pLysS pXQZ001	Chapter 2
Plasmid	pXQZ001	*sigE* in the T7 expression vector pET15b, Kan^R	Chapter 2
	pSEB015	isolated *rpoH*P3 promoter in pRLG770, Ap^R	[92]
	pRLG770	*rrnB*P1 promoter in pRLG770, Ap^R	[207]
	pSEB051	*fam* promoter in pRLG770, Ap^R	This work

RNA isolation, preparation of labeled cDNA, and microarray analysis.

Total RNA was extracted with Trizol (Invitrogen), treated with RNase-free DNaseI (Invitrogen) and purified using RNeasy columns (Qiagen) according to the manufacturer's instructions. A 2-color hybridization format was used and dye-swap experiments were performed. For each reaction, 5mg of cDNA from RB50 and the *rseAB* deletion mutant was fluorescently labeled and directly compared as previously described [182, 208]. The two differentially labeled reactions to be compared were combined and hybridized to a *B. bronchiseptica* strain RB50 specific long-oligonucleotide microarray [182]. Slides were then scanned using a GenePix 4000B microarray scanner and analyzed with GenePix Pro software (Axon Instruments). Spots were assessed visually to identify those of low quality and arrays were normalized so that the median of ratio across each array was equal to 1.0. Spots of low quality were identified and were filtered out prior to analysis. Ratio data from the two biological replicates were compiled and normalized based on the total Cy3% intensity and Cy5% intensity to eliminate slide-to-slide variation.

qRT PCR

Total RNA (1 μg) from three independent cultures of RB50 or RB50Δ*rseAB* was reverse transcribed using 300 ng of random oligonucleotide hexamers and SuperScript III reverse transcriptase (Invitrogen) according to the manufacturer's protocol. The resulting cDNA was diluted 1:100, and 1 μl of this dilution was used in qPCR mixtures containing 300 nM primers and 2× SYBR green PCR master mix (Applied Biosystems) using an Applied Biosystems 7300 real-time PCR detection system (Applied Biosystems). Primer sequences are listed in Table 4-10. To confirm the lack of DNA contamination, reactions without reverse transcriptase were performed. Dissociation curve analysis was performed

for verification of product homogeneity. Threshold fluorescence was established within the geometric phase of exponential amplification, and the C_T value for each reaction was determined. The C_T values from all three biological replicates for each strain were compiled, and expression of *recA* was used as an internal control for data normalization. Change in transcript level was determined using the relative quantitative method ($\Delta\Delta C_T$) [209, 210].

Promoter Prediction

Known *E. coli* σ^E-dependent promoters as determined by Rhodius, et al. [48], were aligned to generate a position weight matrix (PWM). This matrix was used to score every sequence of a particular length over the entire *B. bronchiseptica* RB50 genome, based on methods developed by [211, 212]. The length of the sequence to be scored is determined by the spacer length between the -10 and -35 regions of the promoter; most known σ^{70}-family-dependent promoter spacers are between 15-18 nucleotides long, so each of these was independently scored. The score of a particular sequence is defined as:

$$Score = \left(\frac{\left(\sum basescores - B \right)}{M - B} \right) * 100$$

, where the basescore is the number of times that base is in that same position in the PWM, M= maximum score, which is sum of all the most prevalent bases in each position of the PWM (corresponds to consensus sequence), and B=baseline score, which is described as $Baseline = \sum_{A}^{T} (pB * fB)$, where pB=probability of the base occurring randomly in the genome, as determined by the genomic GC content, and fB= frequency of the base in that position of the PWM. Starting with the first base in the genome, each window was scored for its identity to the consensus sequence. Sequences above a threshold score of 72 (the median score of the *E. coli*

promoters in the PWM), and within 1000 bp of the nearest ORF on the same strand were considered candidate SigE-dependent promoters.

Protein purification

His-SigE was purified from strain XQZ001 as previously described [92, 213]. Briefly, cells were grown at 25 °C to an OD_{600} of 0.5, at which point IPTG was added to a final concentration of 1 mM to induce protein production. Following 1.5 hours of induction, cells were harvested by centrifugation and resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 2.5 mM β-mercaptoethanol, 1 mM PMSF). Resuspended cells were then lysed by sonication, and the lysate cleared by centrifugation. The supernatant containing soluble His-SigE was loaded onto a Ni-NTA column (Qiagen). Bound proteins were eluted with a step-wise gradient of 20, 60, 100, and 200 mM imidazole in column buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2.5 β-mercaptoethanol). Fractions containing SigE were pooled and dialyzed into 20 mM Tris-HCl pH 8.0, 50 mM NaCl, and 2.5 mM β-mercaptoethanol.

In vitro transcription and ROMA

In vitro transcription reactions on PCR-generated templates consisting of the sequence surrounding a putative SigE-dependent promoter were performed as previously described (Chapter 2). Briefly, 100 nM *E. coli* core RNA polymerase (Epicentre) was incubated with 400 nM His-SigE or His-σ^E in transcription buffer (40 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT, 0.1 mg/ml BSA) for 10 min at 30 °C to form holoenzyme. Multi-round transcription reactions were initiated by addition of holoenzyme at a final concentration of 40 nM sigma factor and 10 nM core RNA polymerase, to prewarmed (30 ºC) transcription mix containing 5 nM template DNA, 5%

glycerol, 200 mM ATP, 200 mM CTP, 200 mM GTP, 10 mM UTP, and 2.5 mCi [α-^{32}P]-UTP in transcription buffer. After 10 min at 30 °C, reactions were stopped by the addition of stop solution (80% formamide, 20 mM EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue). Samples were electrophoresed on 6% polyacrylamide gels containing 7.5 M urea, and RNA was visualized by phosphorimaging. Templates were generated by amplification of the promoter regions using the oligonucleotides described in Table 4-10.

For ROMA experiments, His-SigE and *E. coli* core RNA polymerase (Epicentre) were mixed and incubated at 30^{o}C for 15 minutes. RB50 genomic DNA that had been sheared by vortexing was added to the mixture (to a final concentration of 500ng/reaction) and incubated for 10 minutes at 30^{o}C to form open complexes. Samples were diluted to a final concentration of 50 nM core RNA polymerase, and either 0, 0.1, or 1.0 μM SigE, and added to a mix of NTPs (800 μM each) and transcription buffer (50 mM NaCl, 40 mM Tris pH 8.0, 10 mM $MgCl_2$, 100 μg/mL BSA, 5% glycerol, 1 mM DTT) to start transcription. Reactions proceeded for 30 minutes at 30^{o}C, and were stopped with 200 μl of stop solution (2.5 M ammonimum acetate, 10 mM EDTA, 0.1 mg/mL glycogen). The reactions were then extracted twice with phenol, once with chloroform, and precipitated with ethanol and sodium acetate. After DNase treatment, RNA was extracted twice with phenol, once with chloroform, and precipitated with ethanol and sodium acetate. The resulting RNA was resuspended in nuclease-free water (Ambion), quantified, and microarray analysis was performed as described above.

Table 4-10: Primer sequences, 5'-3'

Primer name	Sequence
qfamF	TGGAGGCCGCAAACATG
qfamR	GGCGTTGCCGGAAAGG
qbscNF	AGCCAGCCGGGTCATGA
qbscNR	ATACGTCCGGCCAGGTACTTG
qbsp22F	CGGCACGGGCGTCAT
qbsp22R	GGTGTAGGCACTTTCGAGTTCCT
qbscLF	CGACTACTTCGCGGGTATCG
qbscLR	GCGGACCGCGCTCAT
qbopDF	CGGCTCGGTGAAGACATCTAC
qbopDR	GCCTCCCGCATCTGTTGA
qbopBF	GCTCAATTCGACGAGGCCTAT
qbopBR	TGTGCGTACTCGCCATATCG
1089F	GGCGGCCTCGAGTAGGGACAGGCACACCGGA
1089R	GCCGCCAAGCTTTATAGACACCGCGCGCATC
1282F	GGCGGCGAATTCCGAACTGGCCGCGCTGCGCGA
1282R	GGCGGCAAGCTTGATCGACCCGCAGGCTGGCG
1286F	GCCGGCGAATTCCCGCACATCGACTCGCTCGAC
1286R	GCCGCGAAGCTTGTCGTGGTGACCGTGATGGT
1837F	GCGGCGGAATTCCTGGTGCGCAACACCGAG
1837R	CCGCCGAAGCTTACGCGCTGTCCTCGCCAGC
1894F	GGCGGCGAATTCAGGGCGAAAAAGCCCAGC
1894R	GGCGGCAAGCTTGCTTGGACTGGCGGGCGGGG
2177F	GACGACGAATTCCGAGGGCGAGAAGGGCG
2177R	GCGACGCCCGGGCTTGGGTCAGACCGCCG
3068F	GCGCGCGAATTCCCACCGGCCTCGCCCTG
3068R	GCCGCCAAGCTTGCCATCCATGCGCGTTCA
3108F	GCGACGGAATTCCGCGCCGCAGATACCGGT
3108R	GCGTGCAAGCTTGTACTTCCTGGTCACGATG
3172F	GCGCGCGAATTCCTGACCGCGCCGGCCTTG
3172R	GGCGCGAAGCTTGAAATCGGCCACCAGCG
3678F	GGCGGCGAATTCAGAACGACGTGGACGACAGC
3678R	GGCGGCAAGCTTTGAGGTCGTTCTCCAGCAGGC
3739F	GGCGGCGAACTTCCCGATACCCCCTTCCAGCC
3739R	GGCGGCAAGCTTGCGCAGCACCCTTGTTCCGGA
3751promF	GCCGCCGAATTCCGCGCATGATCCGCGATCCG
3751promR	GCCGCCAAGCTTCACGCAGGACGATGGCGGT
3932F	GGCGGCGAATTCCAGCGCACGGAAGGCGGAGT
3932R	GGCGGCAAGCTTCGGCGCCGGCCTGGAAAGC
4281F	GCGTCGGAATTCTGCCCGGCCACACGGC
4281R	GCTGCAAAGCTTAGGATTGGTCTCGCGCTC
4284F	GGCGGCGAATTCCCTTGCAGCCGCCCAGCGCG
4284R	GGCGGCAAGCTTCTTGCGCGCCGGCGCCCTG
4490promF	CCGCAGGAATTCCATGGTTTCCGTTACGCTG
4490promR	GCGGCGAAGCTTCATAGCAGTTCCCTACGG
4518F	GGCGGCGAATTCCATCATGACGGCGGTGGGCA
4518R	GGCGGCAAGCTTTCGGGCATCATTCGGTGTCTC

Primer name	Sequence
acnAF	GGCGGCGAATTCAAGCGGGTGGCGCAGCAGGT
acnAR	GGCGGCAAGCTTTGCGGCATGATGGACTCCGT
0098F	GACCGCGAATTCGTGCTGCTGATCGCGGC
0098R	CGGCGGAAGCTTGCGGTAGCAAGGGTCTTC
0619F	GCAGACGAATTCCCTATGCTGACGATGCCG
0619R	GGCGGCAAGCTTGCAGAAAGTGCGCTCACG
3842F	GACCGCGAATTCCCTTGGCGATGGCGCGGAT
3842R	GCCGCCAAGCTTCATTTAAGAAATCTCCGT
4083F	GACGGCGAATTCTACGCACGTCGTCGCGCT
4083R	GCGACGAAGCTTGGCGGCCAGGCGTTGCA
bscUF	GCCGCCGAATTCTCTGCCGATGTTCAACCGTC
bscUr	GACGGCAAGCTTCAATCTGGGATTCATGGT
bsp22F	GGCGGCGAATTCGTGCGATGCGCTGGCCTTGC
bsp22R	GGCGGCAAGCTTGAGTCTGGATATCCATGCTC
bvgA F	GCCGGCGGAATTCGATGATGAGGTCGGCGCCCT
bvgA R	GGGCGGAAGCTTTGGTTCCTCAGGGGCAGACA
dsbAF	GGCGGCCTCGAGATCGGGTGCTCGAAGACGCC
dsbAR	GGCGGCAAGCTTCCTGGGTGGCGGGCGCGAAC
fabFprom_HR	GGCGGCAAGCTTCTTTAGCTTACTGCTTACCGTG
fabFprom_XF	GACACGCTCGAGGAAGCGCGCGGCGCGGCCCAT
fabG F	GGGCGGGAATTCGCGCGCGCTGTTGCTGC
fabG R	GGCGGCAAGCTTGTGATGGCGATGCCCGTCAG
famF	GGGCGGGAATTCTGCCGTTCGTGGATGTCCAG
famR	GGGCGGAAGCTTGGGCCAACGAACTACTGGGT
fhaLpromF	GCAGCAGAATTCGGTGCCCTTCCAGCGA
fhaLpromR	GCAGCAAAGCTTGGCTGGCGAACATAG
ftsZF	GGCGTGGAATTCGACGAGGCGGTCGAAGTG
ftsZR	GCCGCCAAGCTTTGCGCGCTTCCTGCAGCA
hfqF	GGCGGCGAATTCTTTCGACGTGCTGCCTTCGG
hfqR	GGCGGCAAGCTTTTATTGCTCATTGGCCAGGC
htpGpromF	GACCGCGAATTCAACGATCGTAGCCATGAG
htpGpromR	ACGGACAAGCTTTTCGTGGTGGTCTGGCTC
lonF	GCCGCGGAATTCCGAGGCCAATGCCGTCGAGG
lonR	GGCGGCAAGCTTGCAGGTCAATCGGGTCGGAA
mucDpromF	GGCGCCGAATTCGTCCTCGCGCCTGGTGCACCAC
mucDpromR	GGCGCGAAGCTTCACGTCGGCCGCACAGGCGA
pbp1AF	GCAGCTCTCGAGTCGCGGCTGGTGCGCCG
pbp1AR	GGCGCTAAGCTTGATCTTTCTGGTGGGGAT
prnF	GGCGGCCTGCGAGGGCCGCTCCAAGGCAAAAAA
prnR	GGCGGCAAGCTTGCAAACGTACGAAGAACGCC
ptxC F	GGGCGGGAATTCGCATGTTCCTGGGTCCC
ptxC R	GGGCGGAAGCTTGTTCCCGCGTCGGACAGGCAC
rnrF	GGCGGCGAATTCCTTGCGTATCCGGATTTTCT
rnrR	GGCGGCAAGCTTAGTTCGGCCGGCGACAACGG
sigEprom_HR	GGCGGCAAGCTTATTGCAGTTGAACGAGCGCC
sigEprom_XF	GGCAGCCTCGAGTCGGCGACCATGCCTCCAAG
sphB1F	GGCGGCGAATTCCGCAGTAACTGCAACGCTTC
spHB1R	GGCGGCAAGCTTTGGCCATGCTGCGCCTGCT

Separation of cell envelope proteins and western blotting

RB50, RB50Δ*sigE*, RB50Δ*rseArseB,* and RB50Δ*wbm* (lacking O-antigen) were grown at 37 °C to mid-exponential phase. 5 or 10 mL of cells were harvested at 9000 rpm, and the pellets were stored at -20 °C. Cells were resuspended in 100mM Tris-HCl pH 8.0 and 20% sucrose, incubated on ice, and pelleted at 9000 rpm for 10 minutes. The pellet was resuspended in 100 mM Tris-HCl pH 8.0, 20% sucrose, 10 mM EDTA, lysozyme was added to 100 μg/mL, $MgSO_4$ was added to 20 mM, and DNaseI and RNaseA were both added to 5 μg/mL. 1 mL of cold water was added to this mixture, and the cells were disrupted with 5 or more freeze-thaw cycles in a dry ice-ethanol bath. Membranes were then pelleted at 15000 rpm for 25 minutes, washed in 20 mM NaPO4 pH 7.0, and the outer membrane was selectively precipitated in 20 mM NaPO4 pH 7.0, 0.5% sarcosyl. Outer membrane membranes were washed twice more in the same solution, and then resuspended in 75-100 μL Laemmli sample buffer. Protein was separated on a 12% SDS PAGE gel.

For western blotting, protein was transferred to Hybond-P (GE Healthcare), and after blocking in 10% milk in TBST for at least one hour, blots were incubated overnight with serum collected from either naïve mice, mice infected with either RB50, or mice infected with RB50Δ*rseAB*. After washing, blots were incubated with goat anti-Mouse IgG-HRP conjugate (Bio-Rad), and developed with ECL Plus (GE Healthcare).

Preparation of lipopolysaccharides

Mid-exponential phase cultures of RB50, RB50Δ*sigE*, RB50Δ*rseAB*, RB53, or RB54 grown in the absence or presence of 50 mM $MgSO_4$ to modulate the cells to Bvg^- were harvested, and resuspended in PBS, 62.5 mM Tris-HCl pH 6.8, 2% SDS, and 10%

glycerol, and boiled for 5 minutes. The lysate was then added to 62.5 mM Tris-HCl, 0.1% SDS, 10% glycerol, and 10-25 mg/mL proteinase K, then incubated at 55 °C overnight. Samples were then run on tris-tricine gels, and silver-stained to visualize LPS (Bio-Rad).

Cytotoxicity assays

Cytotoxicity assays were performed as previously described [149]. Briefly, RAW 264.7 murine macrophage cells were cultured in Dulbecco modified Eagle medium (DMEM) (HyClone) with 10% fetal bovine serum (FBS) (HyClone). Cells were grown to 80% confluence in 96-well plates at 37 °C in 5% CO_2 and washed with RPMI (Mediatech) containing 5% FBS. Bacteria that had been pre-incubated in either Stainer-Scholte medium or in the presence of 100% FBS for one hour were centrifuged and resuspended in DMEM containing 10% FBS, penicillin, and streptomycin. The bacteria were then added to the macrophages at a multiplicity of infection (MOI) of 10. After a 5-minute centrifugation at 250xg, the RAW 264.7 cells and bacteria were incubated for 1, 2, 3, or 4 hours. Percent lactate dehydrogenase (LDH) release, a measure of cytotoxicity, was determined using the Cytotox96 Kit (Promega) according to the manufacturer's protocol.

Acknowledgements

I would like to thank Xuqing Zhang for some of the initial RNA isolation and serum collection, Laura Goodfield for additional serum collection, and B. Tracy Nixon for assistance writing the promoter prediction program.

Chapter 5

Conclusions and Future Directions

Bordetella bronchiseptica encounters many stressful conditions during infection of a host respiratory tract, and combating these stresses effectively is essential for survival. Despite this, very little is known about how *B. bronchiseptica* senses its environment and responds to stress. In this dissertation, I presented evidence that the extracytoplasmic function sigma factor SigE contributes to the ability of *B. bronchiseptica* to combat stress conditions, colonize the respiratory tract of a host, and cause disease. I demonstrated that *B. bronchiseptica sigE* encodes an RpoE-like sigma factor, and that the downstream genes *rseA* and *rseB* encode negative regulators of its activity. The *B. bronchiseptica* SigE system is required for response to heat shock, ethanol stress, and treatment with cell-wall active antibiotics. *B. bronchiseptica* with constitutively active SigE do not efficiently colonize or persist in the lower respiratory tract; however, both high and no SigE activity decrease the ability of *B. bronchiseptica* to cause systemic, lethal infection in mice lacking an adaptive immune response, and alter the interaction of *B. bronchiseptica* with phagocytic cells, suggesting that maintaining proper regulation of SigE activity is important for many aspects of pathogenesis. Gene expression analysis and bioinformatic prediction of SigE-regulated promoters has generated a candidate list of SigE regulon members, many of which have putative roles at the cell envelope, consistent with the stress-related and virulence phenotypes we have observed. This chapter outlines some of

the major conclusions of my dissertation research, and proposes ways to answer some remaining questions about the role of the SigE system in *Bordetella*.

Components and regulation of the SigE system

SigE, RseA, and RseB

Most RpoE-like sigma factor systems consist of a sigma factor, an anti-sigma factor, and a periplasmic protein that contributes to inhibition of sigma factor activity [53, 57, 59]. I have demonstrated that *B. bronchiseptica sigE* encodes a functional RpoE-like sigma factor, and that the downstream genes *rseA* and *rseB* encode negative regulators of SigE activity. Overexpression of *sigE* and *rseA* together in an *E. coli* strain with an *E. coli* σ^E-dependent reporter (*rpoH*P3::*lacZ*) resulted in about a three-fold decrease in activity below overexpression of *sigE* alone. Overexpression of *sigE*, *rseA*, and *rseB* together decreased reporter activity an additional two-fold, returning it to basal levels, suggesting that both proteins are important for this regulation (Fig. 3-2B). This is more similar to the *P. aeruginosa* AlgU system, where deletion of *mucA* or *mucB* individually increases AlgU activity to about the same extent, suggesting that both contribute fairly equally to inhibition [58, 59]. In the *E. coli* σ^E system, however, deletion of *rseA* increases σ^E activity 25-fold, and deletion of *rseB* increases this activation by only about an additional two-fold [57], indicating that *E. coli* RseA is the primary protein responsible for inhibition of σ^E activity. This suggests that although the negative regulators are conserved between species, there are differences between how these proteins interact with their respective sigma factors to inhibit their activity, or differences in the pathway that releases the sigma factor from its anti-sigma factor. A strain of *B. bronchiseptica* lacking both *rseA* and *rseB*

exhibits increased SigE activity (Fig. 3-2B), but generating single deletions of *rseA* or *rseB* could help define the contribution each makes to regulation of SigE.

I have shown that the cytoplasmic domain of *B. bronchiseptica* RseA ($RseA_{Bb}$-cyto) directly inhibits $SigE_{Bb}$ in vitro (Fig. 3-3). While *E. coli* RseA-cyto inhibits both *E. coli* σ^E and *B. bronchiseptica* SigE in vitro, $RseA_{Bb}$-cyto does not inhibit *E.coli* σ^E-dependent transcription in vitro or in *E. coli* (Fig. 3-3 and data not shown). This further demonstrates that while the sigma factors and many of the residues of the anti-sigma factor that interact with σ^E from both the *E. coli* and *B. bronchiseptica* systems are highly conserved, there are differences in how $RseA_{Bb}$ interacts with $SigE_{Bb}$ versus how $RseA_{Ec}$ interacts with σ^E_{Ec} or $SigE_{Bb}$. With the establishment of a $SigE_{Bb}$-dependent reporter in *B. bronchiseptica* (Fig. 3-2), we will be able to test whether this interaction also occurs in vivo by cloning $RseA_{Ec}$ into an inducible plasmid and moving this into a *B. bronchiseptica* strain with the *fam*P::*lacZ* reporter. If overexpression of $RseA_{Ec}$ decreases *B. bronchiseptica* SigE-dependent reporter activity, then it would be interesting to determine which portions of these anti-sigma factors are responsible for the different interactions they have with their cognate sigma factors. Modeling the structure of *B. bronchiseptica* RseA-cyto and SigE and comparing this to the published *E. coli* σ^E:RseA structure [189] suggests that these proteins share similar secondary structure. However, alignment of the primary sequence has shown both conservation and distinct differences in regions of RseA known to interact with σ^E in *E. coli*. These regions may be good initial targets for examining residues important for the differences between these sigma-anti-sigma factor interactions.

Regulated proteolysis pathway

Many sigma factors, including *E. coli* σ^E, *P. aeruginosa* AlgU, and *B. subtilis* σ^W, are released from their anti-sigma factors by regulated proteolysis [44, 109, 214] (see Appendix C for more detail). An inducing signal activates the first protease, which cleaves the anti-sigma factor, creating a substrate for a second protease, which cleaves a second time, and after a final cytoplasmic protease degrades the remaining portion of the anti-sigma factor, the sigma factor is free to bind core RNAP and direct transcription of its regulon.

I identified a potential ortholog of the second protease in this regulated proteolysis pathway, RseP: BB2612. The protease with the most sequence similarity to DegS in *B. bronchiseptica* is BB4867, also annotated as DegQ. Further study is needed to determine if these proteins function as proteases that degrade RseA in *B. bronchiseptica*. In *E. coli*, *degS* and *rseP* are essential, and their essential function is to provide σ^E activity [215, 216]. In other bacteria, however, such as *P. aeruginosa* and *S. enterica* serovar Typhimurium, neither *degS* nor the gene encoding σ^E is essential [217, 218]. *sigE* is not essential in *B. bronchiseptica*; therefore, I would hypothesize that the genes encoding the putative homologs of *degS* and *rseP* would not be essential in *B. bronchiseptica.* Determining whether SigE activity decreases when the genes encoding one or both of these proteases is deleted would provide evidence for whether these proteases are involved in degrading RseA and releasing SigE to bind core RNAP.

In *E. coli,* the –YXF motif at the C-terminal end of outer membrane porins (OMPs) binds to the PDZ domain of the protease DegS to initiate the proteolytic cascade that results in release of σ^E from its anti-sigma factor RseA [73, 219]. I identified a common

motif in the proteins annotated as porins in *B. bronchiseptica,* also ending with phenylalanine. In fact, three of the six proteins annotated as putative porins end in –HRF (Table 5-1). In *Pseudomonas aeruginosa,* the DegS ortholog AlgW is not activated by the same peptide that activates the *E. coli* protease [109], demonstrating that the activation of the proteases that degrade RseA is species-specific.

Table 5-1: C-terminal three amino acids from annotated porins in *B. bronchiseptica*

Gene	Product	C-terminal sequence
BB3842	outer membrane porin protein precursor	HRF
BB3993	outer membrane porin protein OmpQ	QRF
BB4490	outer membrane porin	HRF
BB4551	outer membrane porin protein precursor	HIF
BB4566	outer membrane porin protein precursor	HRF
BB4842	putative outer membrane protein	YRF

Now that I have constructed a *B. bronchiseptica* SigE-dependent reporter, we can determine whether SigE is also activated by the accumulation of OMPs, similarly to its counterparts in *P. aeruginosa* or *E. coli.* Since there is a common C-terminal motif at the end of annotated porin genes in *B. bronchiseptica*, overexpression of an unrelated periplasmic protein, such as cytochrome c (CytC) with the three C-terminal peptides from an OMP (-HRF) would provide evidence for whether SigE is activated by accumulation of misfolded OMPs, or if something else triggers release of SigE from RseA in *B. bronchiseptica.*

I determined that SigE activity does not increase when the *E. coli* C-terminal peptide –YXF fused to cytochrome c (CytC-YYF) is overexpressed in an *E. coli* strain with a plasmid-encoded copy of *B. bronchiseptica sigE* and *rseA*, but no endogenous *E. coli rpoE.* This suggests that activating *E. coli* DegS does not release *B. bronchiseptica*

SigE from inhibition by RseA, and that there are differences between the proteolytic pathways that releases σ^E or SigE, possibly in the substrates (residues or portions of RseA) recognized by the various proteases of each pathway.

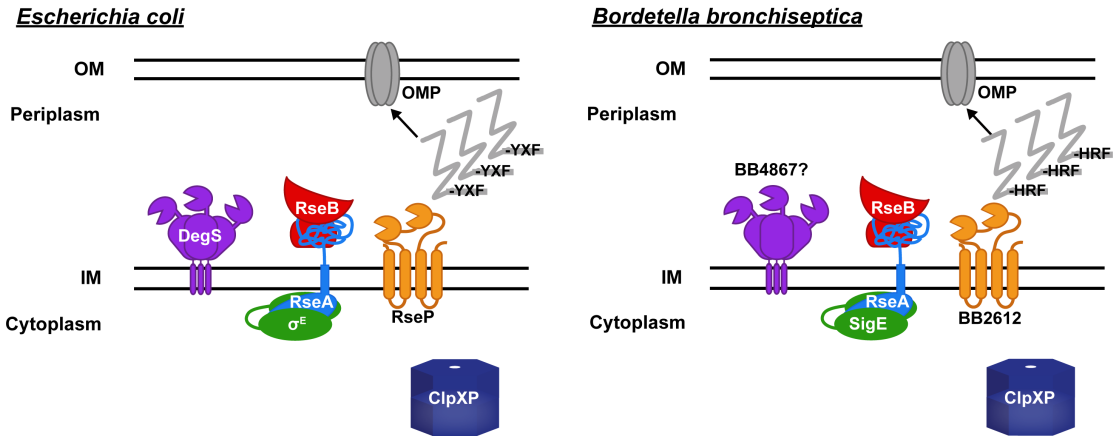


Figure 5-1: The *E. coli* σ^E system and *B. bronchiseptica* SigE system. A model for the regulation of the *B. bronchiseptica* SigE system (*right*) based on the *E. coli* σ^E system (*left*). This work has shown that the basic components are fairly well conserved (σ^E, RseA, RseB), and possible homologs of the other components of the system have been identified in the *B. bronchiseptica* genome, as described above.

ppGpp

We have shown that in *E. coli*, σ^E is also regulated independently of the RseA-dependent regulation described above by the alarmone ppGpp and the protein DksA. ppGpp and DksA together positively regulate σ^E upon entry into stationary phase, through both direct and indirect mechanisms ([92], Appendix B). *B. bronchiseptica* encodes *relA*, *spoT*, and *dksA,* suggesting that it has the capacity to produce the global regulator ppGpp, but this has never been investigated in the bordetellae. In a strain of *B. bronchiseptica* with the SigE-dependent *fam*P::*lacZ* reporter, I have determined that SigE activity does not

increase significantly upon entry into stationary phase, indicating that perhaps ppGpp and DksA do not regulate SigE activity in a growth-phase dependent manner (data not shown). However, there is no direct evidence that ppGpp, if produced in *B. bronchiseptica*, even accumulates upon entry into stationary phase. Ongoing work in the laboratory of one of our collaborators at the USDA-ARS, Tracy Nicholson, includes generating a strain that no longer produces ppGpp, lacking *relA* and *spoT*, to determine whether ppGpp plays an important role in growth phase regulation and in virulence for *B. bronchiseptica*, as it does in many other bacteria [220, 221].

Transcriptional regulation of the *sigE* locus

In *E. coli* and many other bacteria, the gene encoding σ^E, *rpoE*, is the first gene in an operon consisting of *rpoE, rseA, rseB* and *rseC* (Fig. 5-2). This operon has both a σ^E-dependent promoter and a σ^{70}-dependent promoter [69, 70]. However, in *B. bronchiseptica*, *sigE* is not predicted to be the first gene in the operon. Instead, there are two genes predicted to be involved in fatty acid biosynthesis directly upstream of *sigE*, *bb3753* and *fabF*, and each of these have overlapping start and stop codons. Similar to other systems, downstream of *sigE* are encoded two negative regulators, *rseA* and *rseB* (formerly *bb3751* and *mucB*). Although *B. bronchiseptica* does not encode an RseC-like protein, there is a putative serine protease downstream of *rseB*, annotated *mucD* because of homology to the *P. aeruginosa* protein.

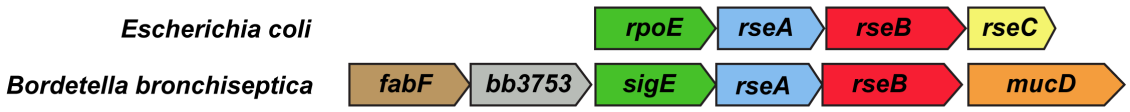


Figure 5-2: The *B. bronchiseptica sigE* operon has different architecture from the *E. coli rpoE* operon.

Preliminary results from PCR on cDNA isolated from RB50 using primers across the junctions between these genes suggest that some of these genes can be co-transcribed, including *bb3753/sigE*, *sigE/rseA*, and *rseB/mucD* (data not shown). Additionally, I have tested for SigE-dependent transcription from the regions upstream of both *sigE* and *fabF*, and thus far have no evidence that SigE transcribes these regions in vitro. In *E. coli* and other systems, autoregulation provides a method of quickly upregulating expression of this response system to combat stress [69, 70]. There is some evidence that in *Xylella fastidiosa*, the gene encoding σ^E is not autoregulated [130], so perhaps in some systems, such as *B. bronchiseptica,* other sigma factors and/or transcription factors that can compensate for the additional σ^E activity provided by autoregulation in other systems. *E. coli rseA* also has an internal σ^E-dependent promoter, so that the system can be quickly shut off when the stress conditions have ceased [48, 81]. I have identified a SigE-regulated promoter upstream of *B. bronchiseptica rseA* (Fig. 4-4). It would be of interest to map the promoter for *sigE* and *rseA*, along with possibly other genes in the predicted operon using 5'RACE or primer extension, to determine whether this system is autoregulated. If it is not, it would also be interesting to determine how transcription of this locus is regulated.

I have shown that the basic components of the SigE system are fairly well-conserved between *B. bronchiseptica* and other described RpoE-like sigma factor systems, such as *E. coli* σ^E and *P. aeruginosa* AlgU. Differences in some of the details of the regulatory paradigm, including the contributions of the negative regulators, the inducing signals for the potential regulated proteolysis pathway that releases the sigma factor, and the architecture of the *B. bronchiseptica sigE* locus suggest that within the basic system, each species has evolved its own method of fine-tuning the regulation of this sigma factor system to best respond to stress conditions.

The role of the SigE system in cell envelope stress response

E. coli σ^E was first discovered for its role transcribing σ^{32} under extremely high temperature [67]. It was subsequently shown to be important for cell envelope maintenance and stress response in both *E. coli* and *S. enterica* serovar Typhimurium, for oxidative stress in *S. enterica* serovar Typhimurium and *P. aeruginosa*, and other subsets of stress conditions in many other bacteria [5, 6, 49-51, 123, 124, 127] (see Chapter 1 for more detail). I demonstrated that *B. bronchiseptica* SigE directs transcription of the heat shock sigma factor σ^{32}, encoded by the *fam* gene (Fig. 2-2). Cells lacking *sigE* are more sensitive to heat shock and ethanol stress. Simply providing *B. bronchiseptica* with increased SigE activity, however, does not confer greater resistance to either treatment. I found that constitutive SigE activity from either a plasmid or deletion of *rseA* and *rseB* does not increase resistance to heat shock (Chapter 2 and Fig. 3-4A). In fact, RB50Δ*rseAB* shows decreased thermotolerance, or the ability to adapt to elevated, sublethal temperatures and

better survive an otherwise lethal temperature (Fig. 3-4A). This suggests that proper regulation of SigE is important for responding to temperature stress in *B. bronchiseptica.*

Cells lacking *sigE* are also more sensitive to treatment with detergent or β-lactam antibiotics, which specifically target the cell envelope. Unlike with heat shock, however, constitutive activation of SigE (RB50Δ*rseAB*) confers resistance to these stresses, and plasmid-encoded *rseA* and *rseB* restored sensitivity to nearly wild-type levels (Fig. 3-4B). RB50, RB50Δ*sigE* and RB50Δ*rseAB* showed no difference in sensitivity to treatment with other antibiotics, detergents, or to osmotic and oxidative stress, suggesting that the differences observed are not due to changes in overall membrane permeability. This indicates a specific role for the SigE system in cell envelope stress response.

The activity of a *B. bronchiseptica* SigE-dependent reporter (*fam*P::*lacZ)* does not increase in response to heat shock, ethanol stress, or treatment with ampicillin or mecillinam (data not shown). However, the *E. coli* σ^E-dependent reporter *rpoH*P3::*lacZ* does not increase much under stress conditions [69], and in the closely related *Burkholderia pseudomallei*, where σ^E is also required for responding to heat shock, activity of a σ^E-reporter increased less than two-fold in response to heat shock [124]. Increases in *E. coli* σ^E activity have been observed when OMPs are overexpressed, or in mutants lacking periplasmic folding factors [69, 73]. To determine specific intracellular triggers of SigE activity in *B. bronchisptica*, similar genetic studies could be performed.

The role of the *B. bronchiseptica* SigE system in virulence

RpoE-like sigma factors play diverse roles in the virulence of many bacterial pathogens. *S. enterica* serovar Typhimurium and *V. cholerae* mutants lacking σ^E are attenuated for virulence [97, 98]. *P. aeruginosa* mutants lacking σ^E (AlgU), however, show increased virulence, and in the porcine respiratory pathogen *Actinobacillus pleuropneumoniae,* mutants with high SigE activity are attenuated, while a strain lacking σ^E is not [60, 64]. In collaboration with the Harvill lab at Penn State, we found that the *B. bronchiseptica* SigE system, particularly regulation of SigE, plays an important role in virulence.

Colonization and persistence in the lower respiratory tract

While RB50Δ*sigE* colonizes and persists in the respiratory tract of wild-type C57BL/6 mice similarly to RB50, RB50Δ*rseAB* is defective in colonizing the trachea and lungs of these mice. This defect in colonization appears to increase later during infection, which led us to hypothesize that RB50Δ*rseAB* could have altered interactions with the adaptive immune response. Despite differences in the type of antibodies produced against RB50Δ*rseAB* and RB50 during infection (Fig. 4-4), the antibodies produced against each strain are equally capable of clearing RB50Δ*rseAB* (Fig. 3-6), suggesting that although different proteins in each strain may be available to the immune system to make antibodies against, this does not affect the overall perceived strength of the adaptive immune response. However, it is still conceivable that RB50Δ*rseAB* is generally more sensitive to antibody-mediated clearance than RB50, so that differences in the ability of serum from RB50 or RB50Δ*rseAB* infection would be masked. Further work, including determining

whether immune serum from mice inoculated with each either RB50 or RB50Δ*rseAB* also affects survival of RB50 similarly, is necessary to differentiate these possibilities. Additionally, the defect in colonization of the lower respiratory tract is not alleviated in mice lacking a functional adaptive immune response, but is partially abrogated in mice lacking a key component of the innate immune response, TLR4 (Fig. 3-7), suggesting that high SigE activity alters the interaction of *B. bronchiseptica* with some component of the innate immune response, discussed later in this section.

In comparing genes regulated by the BvgAS two-component system that regulates expression of most known *B. bronchiseptica* virulence factors with genes having decreased expression in RB50Δ*rseAB*, I discovered that at least three adhesin genes have significantly decreased expression in RB50Δ*rseAB*: *fim2*, *fim3*, and *bipA* (Table 4-7). One of the only known phenotypes associated with deletion of the genes encoding fimbrial adhesins is decreased tracheal colonization [147]. One hypothesis is that decreased expression of fimbrial genes in this mutant contributes to the colonization defect. To examine this, future work will determine the protein level of the fimbrial genes in RB50Δ*rseAB* compared to RB50 to confirm decreased expression of these adhesins. Ongoing work will also examine the ability of RB50, RB50Δ*sigE*, and RB50Δ*rseAB* to adhere to host cells (S. Hester, ongoing work). Alternatively, differential expression of another factor or altered composition of the envelope through constitutive activation of SigE could prevent proper expression or placement of a factor or factors important for colonization and persistence. Ongoing work determining the SigE regulon may provide some insight into the role of SigE in regulating colonization and persistence in the lower respiratory tract.

Interactions with the host immune system

The *B. bronchiseptica* strain RB50 asymptomatically colonizes the respiratory tract of immunocompetent mice, as described above [222]. Other strains of *B. bronchiseptica* cause a range of disease in mice and other hosts, ranging from asymptomatic colonization to systemic lethal infection [223, 224]. In mouse models lacking key components of either innate or adaptive immunity, even RB50, which typically causes asymptomatic colonization, can escape the respiratory tract, colonize systemic organs, and cause lethal disease [138, 143, 144]. In this way, immunocompromised mouse models allow us to investigate the role of bacterial factors both in specific interactions with the immune system as well as in causing systemic, lethal infection. While RB50Δ*sigE* can still cause systemic, lethal infection in mice lacking key components of the innate immune response, it no longer colonizes systemic organs nor causes death in $Rag1^{-/-}$ mice, which lack the ability to make B and T cells. Cells with constitutively active SigE are unable to cause systemic, lethal infection in mice lacking TLR4 or TNF-α, as well as in cells lacking an adaptive immune response ($Rag1^{-/-}$).

Whether the inability to colonize systemic organs is because the bacteria are unable to escape the respiratory tract to colonize systemic organs, or because once the bacteria escape, they cannot survive in the bloodstream, has yet to be determined. To differentiate between these two possibilities, RB50, RB50Δ*sigE*, and RB50Δ*rseAB* will be directly injected into the bloodstream of $Rag1^{-/-}$ mice (L. Goodfield, ongoing experiments). If RB50 and RB50Δ*sigE* or RB50Δ*rseAB* colonize and survive similarly, the defect is likely in escape from the respiratory tract. However, if RB50Δ*sigE* and/or RB50Δ*rseAB* survives in the bloodstream less well than RB50, then it is likely that the SigE system is important

in combating some stress condition unique to this environment not found in the respiratory tract. This would suggest that these RB50Δ*sigE* is able to escape the respiratory tract when numbers are high, but cannot survive once it escapes.

We do know that RB50Δ*sigE* and RB50Δ*rseAB* are not more sensitive to complement, a bactericidal component of the bloodstream. However, both cells lacking *sigE* and those with high SigE activity (RB50Δ*rseAB*) have altered interactions with macrophages, which are also circulating in the bloodstream. *B. bronchiseptica* is cytotoxic to macrophages, and this cytotoxicity is dependent on a functional type three secretion system (T3SS), adenylate cyclase toxin (ACT), and possibly the type six secretion system (T6SS) ([163, 193] and L. Weyrich, unpublished results). Strains lacking either *sigE* or the negative regulators *rseA* and *rseB* are less cytotoxic to macrophages, suggesting that proper regulation of the SigE system is also required for full cytotoxicity (Fig. 2-7, Fig. 3-8). Plasmid-encoded *sigE* complements this defect (Fig. 2-7), and ongoing experiments will determine whether plasmid-encoded *rseA* and *rseB* also complement this phenotype. Incubation in serum or CO_2 has been shown to stimulate cytotoxicity (S. Hester, unpublished results). While cells lacking *sigE* still respond to incubation with CO_2 or serum, cells lacking *rseA* and *rseB* do not (Fig. 4-5 and S. Hester, preliminary data), indicating that high SigE activity interferes with the ability of *B. bronchiseptica* to increase cytotoxicity in response to serum or CO_2. Why cytotoxicity increases in response to these stimuli is still unclear; comparison of genes differentially expressed in RB50Δ*rseAB* vs. RB50 with those differentially expressed after incubation with serum (S. Hester and T. Nicholson, unpublished results), revealed few obvious answers.

As a transcription factor, it is possible that SigE controls genes important in assembly of the secretion system complexes involved in cytotoxicity or the components of the secretion systems themselves. Gene expression analysis reveals possible regulation of parts of the T3SS apparatus, but not the locus in its entirety. Additionally, we have seen an increase in expression of some portions of the T3SS in RB50Δ*rseAB*, but decreased instead of increased cytotoxicity. An alternative hypothesis is that changes in the cell envelope composition through constitutive activation of SigE may indirectly affect proper assembly of this large envelope-spanning apparatus. A more detailed analysis of protein levels to determine whether this secretion system is being properly expressed at the envelope is necessary to determine whether this or some other factor can provide insight into the role that the SigE system plays in cytotoxicity to macrophages.

B. bronchiseptica also has the capacity to be phagocytosed and may sometimes live intracellularly [225]. Cells lacking *sigE* attach better to polymorphonuclear leukocytes (PMNs) than wild-type RB50. RB50Δ*sigE* is also phagocytosed better and killed better by PMNs than RB50. In *Salmonella enterica* serovar Typhimurium, *Haemophilus influenzae* and *Burkholderia* spp., σ^E is also required for survival in phagocytes [50, 65, 125]. In the case of *S. enterica* serovar Typhimurium, σ^E is also required for survival during oxidative stress, which is an important part of intracellular survival [65]. However, *B. bronchiseptica* SigE does not seem to be required for survival in the presence of either paraquat or hydrogen peroxide. The reason for the decreased survival of RB50Δ*sigE* in phagocytes is unclear. Expression of some adhesins, including *fim2*, *fim3*, and *bipA* are decreased in RB50Δ*rseAB* compared to RB50, and expression of all but *fim2* is modestly increased in RB50Δ*sigE* compared to RB50. This could, in part, help explain the increased attachment

of RB50Δ*sigE*. Ongoing work will determine whether high SigE activity also affects the ability of *B. bronchiseptica* to survive intracellularly, possibly decreasing attachment (L. Bendor, ongoing work).

In *P. aeruginosa,* cells with constitutively active AlgU (σ^E) activity display increased biofilm formation, but decreased expression of many virulence factors, promoting the chronic infection observed in many cystic fibrosis patients with *P. aeruginosa* infection [114]. Preliminary data from a collaborator has suggested that transcription of *sigE* is increased in biofilms, and that a strain lacking *sigE* has altered biofilm architecture during early biofilm formation, but these differences are no longer apparent later in biofilm formation (Raj Deora, personal communication). These results are consistent with our predicted role for SigE at the cell envelope, but the specific role of SigE in biofilm formation has yet to be determined.

A mutant of *A. pleuropneumoniae* lacking *rseA* displays increased attachment but attenuated virulence [134]. Our results show that, like this, *B. bronchiseptica* with constitutively active SigE also have decreased virulence. Further work, however, is necessary to determine exactly what SigE regulon members might confer these virulence defects when SigE activity is high. It is also possible that this regulation might be more complex; changing expression of multiple envelope-associated proteins might mask or uncover particular factors that alter the interaction of *B. bronchiseptica* with its host.

The SigE regulon

To date, I have confirmed nine SigE-dependent promoters in *B. bronchiseptica* that are upstream of: *fam, hfq, rseA, bb3108* (hypothetical protein)*, bb1282* (two component

response regulator), *bb1837* (ECF sigma factor), *bb4281* (putative membrane-bound lytic glycosylase), *bb0098* (putative exported protein), and *bb3842* (putative outer membrane porin) (Fig. 4-4). The promoters of these genes contain motifs similar to those transcribed by other RpoE-like sigma factors. While unlikely because of high sequence similarity between the various subunits of core RNAP in *E. coli* and *B. bronchiseptica,* it is possible that forming holoenzyme with SigE and *E. coli* core RNAP during in vitro transcription assays has biased the interaction of *B. bronchiseptica* SigE toward interacting with promoters more similar to those recognized by *E. coli* σ^E. To control for this, I attempted to purify core RNAP natively from *B. bronchiseptica* with the help of Katsuhiko Murakami at Penn State. However, the resulting protein was not competent for transcription under any of the conditions tested. Further work, such as promoter mapping through primer extension, is necessary to determine the sequences of promoters recognized by *B. bronchiseptica* SigE in vitro and in vivo.

Some of the candidate SigE-dependent promoters are upstream of genes orthologous to genes with σ^E-dependent promoters in *E. coli*, such as *fam* (σ^{32}, described above), *rseA*, *mucD* (*E. coli degP*), *ftsZ*, and *bb4517* (*E. coli yraP*). However, these are the only genes out of the 23 genes in the "core σ^E regulon" that have been identified by any of our methods [48]. Six genes of this core regulon do not have known orthologs in *B. bronchiseptica*, leaving 12 that are present in the *B. bronchiseptica* genome, but not predicted to be regulated bo SigE. This suggests that our methods have not yet successfully identified these regulon members in *B. bronchiseptica*, or that while a few regulon members are conserved, *B. bronchiseptica* has adapted its SigE system to transcribe a divergent set of genes compared to other characterized σ^E systems.

The expression of many genes changes in response to σ^E activity, both positively and negatively. Presumably, some of this regulation is direct and some is indirect. Many of these genes do not have obvious SigE-regulated promoters. SigE is predicted to transcribe the genes encoding a handful of other transcription factors, such as *fam* and *bb1837*; their altered activity could be responsible for other transcriptional changes. Additionally, SigE is predicted to transcribe many envelope proteins, and altered expression of these may lead to downstream changes in the transcription of other factors. Confirmation of additional SigE regulon members through in vitro transcription, as described in Chapter 4, as well as possibly through in vivo reporter assays will hopefully provide insight into the various stress and virulence-related phenotypes we have observed for mutants in the *B. bronchiseptica* SigE system.

SigE-regulated sRNAs

SigE could also affect post-transcriptional regulation of other genes through transcription of small, non-coding RNAs (sRNAs). In *E. coli* and *S. enterica* serovar Typhimurium, σ^E transcribes sRNAs that target outer membrane proteins (OMPs) for degradation [80, 94, 226]. This prevents additional burden to an already-stressed system activated by the accumulation of misfolded OMPs. In contrast, however, I have determined that expression of at least one porin is increased when *B. bronchiseptica* SigE activity is high, suggesting that putative SigE-regulated sRNAs may have different functions from those transcribed by σ^E in *E. coli*. It was recently discovered that the sRNAs that target OMPs for degradation in *E. coli* also target other genes, indicating a broader regulatory role for sRNAs than originally hypothesized [80].

In collaboration with Laura Weyrich, Jihye Park, and David Place, we identified many candidate sRNAs in *B. bronchiseptica* that we are in the process of confirming (Table A-1, Table A-2, Fig. A-2). I compared these putative sRNAs to predicted SigE-regulated promoters (Chapter 4), and identified a few candidate sRNAs that are predicted to have SigE-regulated promoters (Table A-5). I previously confirmed the SigE-dependent promoter upstream of one of these candidate sRNAs, bbsRNA058, by in vitro transcription, and assumed this was a promoter upstream of the gene *bb3108*, which encodes a hypothetical protein (Fig. 4-4). This sRNA is predicted to be encoded on the same strand as *bb3108* and ends only 7 nucleotides from the 5' end of the coding region, suggesting that this may be the 5' UTR of *bb3108.* Primer extension analysis will differentiate between these possibilities.

The RNA for this original analysis was isolated under non-stress conditions, when, at least in other bacterial species, such as *E. coli*, σ^E activity is low. To enrich for putative SigE-regulated sRNAs, we have submitted RB50Δ*rseAB* for whole-transcriptome sequencing (RNA-seq) (L. Weyrich and Jihye Park, ongoing work). We can then amplify the promoter region of candidate sRNA genes and look for SigE-dependent transcription, and identify potential targets for these sRNAs through bioinformatic and experimental analysis.

Other ECF sigma factors in *B. bronchiseptica*

Bacteria can encode anywhere from 0 up to 70 or more ECF sigma factors [40]. *B. bronchiseptica* is predicted to encode twelve ECF sigma factors. Five are most closely related to the FecI sigma factor of *E. coli*, which is known to be important for iron sensing

and uptake. Indeed, two of these factors have been shown to be required for iron sensing in bordetellae [227-229]. Of the remaining seven ECF sigma factors, one, BtrS, regulates many of the genes involved in the type three secretion system (TTSS) [230], one is SigE, and five have yet to be studied in detail: BB1302, BB1837, BB2661, BB3268, and BB3300 (Fig. 5-3).

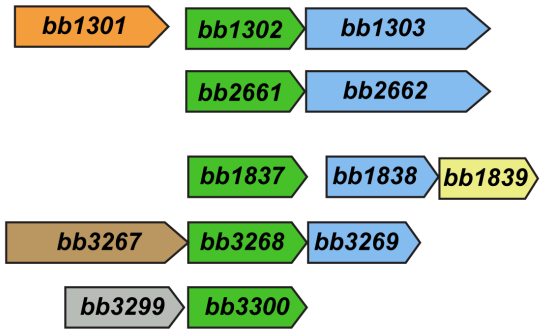


Figure 5-3. Other predicted ECF sigma factors in *B. bronchiseptica*: BB1302, BB2661, BB1837, BB3268, and BB330. The putative sigma factor is indicated in green, and the putative anti-sigma factor in blue.

BB1302 and BB2661 belong to the same group of ECF sigma factors, as organized by Staron, et al. [40]. This is a large and diverse group, with varied gene context arrangement. Some members of this group are encoded downstream of a lipoprotein, some with cytochrome c oxidases and metallophosphoesterases, and some with putative catalases and cytochromes. Only a couple members of this group have been studied at all, including PrtI from *Pseudomonas fluorescens* [231]. PrtI transcribes at least one protease, *aprX*. Mutants lacking *prtI* or the downstream gene, *prtR*, have decreased protease activity, suggesting that instead of acting as anti-sigma factor, PrtR is an activator of PrtI activity [231]. I have cloned the gene encoding BB1302, renamed *prtI-1,* into the expression vector pET15b, which adds an N-terminal 6XHis tag to recombinant proteins. After purification,

PrtI-1 was used for in vitro transcription assays with either the promoter region of PrtI-1 or the upstream gene BB1301 as a template since many ECF sigma factors transcribe their own gene. Under all the conditions tested, I have not demonstrated PrtI-1-dependent transcription of the *bb1301* or *prtI-1* promoter regions in vitro. Further work is needed to determine what promoters are transcribed by PrtI-1, and if any other factors are required for PrtI-1-dependent transcription.

BB1837 is followed by a putative exported protein, and a putative cytochrome c. No members of this group have been studied experimentally. Interestingly, we have confirmed the presence of a SigE-regulated promoter upstream of the gene encoding BB1837, but no other information about this sigma factor is available.

The group of ECF sigma factors that includes BB3300 (group ECF41) was recently studied in the Mascher lab (Wecke, et al., unpublished). This group of ECF sigma factors lacks an obvious co-transcribed anti-sigma factor, but instead has a C-terminal extension that essentially functions as a fused anti-sigma factor domain. Work to determine the regulon of this sigma factor suggested that, in *Rhodobacter sphaeroides* and *Bacillus licheniformis*, ECF41 only directs transcription of the immediately adjacent gene, encoding a carboxymuconolactone dehydrogenase (CMD). In other bacteria, group ECF41 sigma factors are also found adjacent to oxidoreductases and epimerases. In *B. bronchiseptica*, the ECF41 sigma factor BB3300 is also predicted to be adjacent to a CMD protein. In analyzing the promoter sequences of BB3300 and the CMD protein, BB3301, I identified a sequence similar to the promoter identified by Wecke, et al., suggesting that perhaps BB3300 also transcribes BB3301 in *B. bronchiseptica*. Future experiments would include purifying this sigma factor and testing for BB3300-dependent transcription of this

promoter, as well as determining whether, as in *B. licheniformis* and *R. sphaeroides*, this is the only gene transcribed by an ECF41-group sigma factor in *B. bronchiseptica.* While BB3300 is conserved in *B. pertussis* and *B. parapertussis*, suggesting this sigma factor may play an important and conserved role in the classical bordetellae, the orientation of the CMD protein and sigma factor is reversed in the human pathogens compared with *B. bronchiseptica*. It would be of interest to delete *bb3300* from the *B. bronchiseptica* and/or *B. pertussis* chromsome, and test for possible contributions to stress response and virulence for these pathogens.

Because the promoters transcribed by some ECF sigma factors share some similarities, I have cloned the genes encoding BB1302 (*prtI-1*) and BB1837 into pTrc99a, under the control of an IPTG-inducible promoter, and assessed their ability to activate transcription from the *E. coli* σ^E-dependent promoter *rpoH*P3. Neither increased activity from the *rpoH*P3::*lacZ* reporter. The other three sigma factors (BB3300, BB2661, and BB3268) have yet to be tested for activity. Additionally, we can purify the other sigma factors and perform ROMA to identify other candidate promoters they might regulate, and make deletion mutants of each in *B. bronchiseptica* to and gain some insight as to their function.

Encoding multiple ECF sigma factors, along with a multitude of other transcription factors, provides *B. bronchiseptica* the opportunity to have multiple layers of regulation, some of which are likely to have redundant and overlapping regulons. This may partially explain why *sigE* is not essential in *B. bronchiseptica*, and why it is dispensable for many aspects of virulence, though similar genes are essential for survival and virulence in other bacteria.

Final Comments

In this work, I demonstrated that SigE plays a critical role in cell envelope stress response, and that consistent with this finding, candidate members of the SigE regulon have predicted roles at the cell envelope. The cell envelope is a vital compartment of the bacterial cell; therefore, being able to quickly sense and respond to stresses that attack the envelope is essential for cell survival. During infection, bacteria express many of their virulence factors, from adhesins to secretion systems, at the cell envelope, making this compartment also particularly important for virulence. We have demonstrated a possible role for SigE in regulation of some virulence factors, such as adhesins and the type three secretion system, either directly or indirectly.

In addition to changes in temperature, pH, nutrient availability, and other host-associated environmental changes, many host immune responses are designed to recognize and attack various components of the cell envelope, such as cationic antimicrobial peptides that attack the membrane. While the SigE system does not seem to be involved in responding to antimicrobial peptides (Fig. 3-4) or complement (Chapter 2, Chapter 3, data not shown), I demonstrated that the SigE system plays a role in resistance to other stresses that affect the envelope, such as heat shock or treatment with cell envelope-active antibiotics like β-lactams and detergents. In addition, bacteria can modulate the composition of their envelope, such as the LPS, to evade immune attack. I found that there are some differences in envelope composition when SigE activity is high, including LPS modifications, the composition of the outer membrane, and the proteins available for antigen presentation, indicating that the SigE system plays an instrumental part in regulating expression of envelope proteins during infection.

The SigE system is nearly 100% identical in *B. bronchiseptica* and the very closely related human respiratory pathogens *Bordetella pertussis* and *Bordetella parapertussis*. Because these human-adapted bacteria are thought to have diverged independently from a *B. bronchiseptica*-like progenitor through significant gene loss, conservation in all three of the classical bordetellae implies a possibly important and conserved role for this system in *Bordetella* species. However, while the components of similar sigma factor systems even very closely related can be highly conserved, their function can vary dramatically. Preliminary data predicting SigE-regulated promoters in the bordetellae suggests some conserved and some unique roles for this system in each species (Fig. 4-7). Identifying the role that this system plays in responding to stress response and in virulence of *Bordetella* species will be important in understanding how to better combat these pathogens during infection.

Appendix A

sRNA Discovery in *Bordetella bronchiseptica*

Abstract

In addition to regulation of gene expression by transcription factors, bacteria can rapidly alter gene expression through post-transcriptional regulation by small RNA molecules (sRNAs). Bacterial sRNAs are implicated in regulation of outer membrane protein levels, quorum sensing, virulence, and everything in between; the breadth of sRNA-mediated regulation is only beginning to be understood. Outside of the BvgAS two-component system and the extracytoplasmic function sigma factor SigE described earlier in this work, little is known about gene regulation in the mammalian respiratory pathogen *Bordetella bronchiseptica.* Recent work by Hot, et al., has identified some sRNAs encoded by *Bordetella pertussis*. In this section, I describe ongoing, preliminary work identifying, confirming, and characterizing candidate small RNAs in *B. bronchiseptica*.

*L. Weyrich prepared RNA samples and performed the SOLiD sequencing in collaboration with the PSU Nucleic Acid Facility; J. Park and L. Weyrich performed the subsequent analysis of the sequencing output; L. Weyrich predicted targets for the candidate sequenced sRNAs; S. Barchinger performed the cross-species analysis, compared the sequenced and predicted sRNAs, analyzed the location of sRNAs relative to the nearest ORF, predicted SigE-regulated sRNAs and their targets, and performed the northern blots with D. Place

Introduction

Small, non-coding RNAs are increasingly being identified as important regulatory molecules in all domains of life, ranging from siRNAs and miRNAs in eukaryotes, to sRNAs of bacteria [232-235]. sRNAs are a rapid way for bacteria to alter protein expression, as these short molecules are not as energetically expensive to make as larger protein regulators, and can often be quickly turned off and on in response to changing conditions [236]. Regulatory RNAs were identified originally in bacteria for their role in maintaining plasmid copy number and targeting transposase elements [236, 237]. Since these initial discoveries, hundreds of sRNAs with varied functions have been identified in diverse bacterial species.

Regulatory RNAs may act on their targets in a variety of ways. Some cis-acting regulatory RNAs reside in the 5'-untranslated region (5'UTR) of an mRNA and use changes in secondary structure to either promote or prevent translation, depending on conditions such as temperature or metabolite concentration [234, 236, 238]. Yet other cis-acting sRNAs encode extensive complementarity to the gene on the opposite strand and bind that mRNA to prevent its expression. This mechanism is commonly found in transposable elements and plasmids, and one has been implicated in virulence in *M. tuberculosis* and *S. enterica* serovar Typhimurium [234, 239].

Trans-acting sRNAs are often encoded in supposed intergenic regions between protein-coding genes, and target non-neighboring mRNAs or proteins for regulation. Some directly target proteins for activation or sequestering, such as CsrB/CsrC-targeting of the RNA-binding protein CsrA [240]. Other sRNAs promote translation of their targets by sequestering mRNA structures that prevent ribosome binding, as in DsrA/RprA regulation

of σ^S [241, 242]. The majority of known trans-acting sRNAs target mRNAs for degradation [235, 236]. Among the first described examples of this mechanism are the sRNAs MicA, MicF, and RybB in *Escherichia coli* and *S. enterica* serovar Typhimurium, which target the mRNAs for the outer membrane proteins (OMPs) OmpA, OmpF, and OmpC for RNaseE-dependent degradation, respectively [77]. While it was originally thought that sRNAs only targeted one or two mRNAs, recent work on these OMP-targeting sRNAs in *E. coli* has suggested that each sRNA has an extensive potential array of targets, expanding the role that these small regulatory molecules play in the cellular response to changing conditions [80].

Since the initial discovery of sRNAs in bacteria like *Salmonella* and *E. coli*, many methods have been developed to predict and identify sRNAs of diverse functions in many bacteria. In *Vibrio cholerae*, the Qrr sRNAs are required to carry out quorum sensing behaviors [243, 244]. *E. coli* RyhB downregulates non-essential proteins that require iron when iron is scarce, so that the available iron can be used by essential proteins [245]. SgrS targets phosphotransferase transporters for sugars to prevent accumulation of toxic sugar phosphates in the cell [245-248], and *S. enterica* serovar Typhimurium MgrR is involved in lipopolysaccharide biosynthesis and resistance to antimicrobial peptides [249, 250].

Many sRNAs require the assistance of the sRNA-binding protein Hfq to interact with their target mRNAs. Hfq forms a multimeric complex that binds sRNAs and facilitates interaction with their target mRNAs [249, 251, 252]. In *Salmonella* spp., *Pseudomonas aeruginosa*, and *Vibrio cholerae*, a mutant lacking *hfq* is attenuated for virulence, indicating an important role for some sRNAs in pathogenesis for many bacteria [100, 249, 253, 254]. In *Salmonella*, Hfq is required for type three secretion system (T3SS)

expression [254], whereas in enterohemorrhagic *E. coli*, Hfq represses T3SS expression [255]. The roles sRNAs play in regulation of everything from metabolism to biofilm formation to virulence are extensive and species-specific [234, 237, 242, 245, 249]. Even evolutionarily related sRNAs can target similar genes [80], and a single sRNA can target many mRNAs for regulation [80, 256].

In this section, I describe ongoing work identifying sRNAs in *Bordetella* species. The bordetellae are closely related respiratory pathogens [135, 139, 140]. *Bordetella bronchiseptica* can cause a wide range of disease, from asymptomatic carriage to fatal pneumonia in a variety of mammalian hosts, including immunocompromised humans [136-138]. *Bordetella pertussis* and *Bordetella parapertussis*, the etiological agents of whooping cough in humans, are thought to have independently evolved from a *B. bronchiseptica*-like progenitor [135, 139, 140]. Recent work by Hot, et al., identified around 20 putative sRNAs in the respiratory pathogen *Bordetella pertussis* through in silico analysis, and confirmed transcription from these loci via northern blotting [257]. This, however, is the first description of identification of sRNAs in *B. bronchiseptica,* and in any of the bordetellae using deep sequencing methods.

We have identified 181 candidate sRNAs in *B. bronchiseptica* using SOLiD sequencing, and have to date confirmed transcription of six by northern blotting. Analysis of these sRNAs, as well as those in the other bordetellae, suggests that some sRNAs may be conserved at a sequence level, while many are unique to each species. We have predicted targets for the candidate sRNAs, and are in the process of validating these predictions. I have also identified a few candidate *B. bronchiseptica* sRNAs that may be regulated by the extracytoplasmic function sigma factor SigE, and ongoing work will

identify additional candidate SigE-regulated sRNAs through whole-transcriptome analysis of a strain with high SigE activity.

Results and Discussion

Prediction of sRNAs bioinformatically

To bioinformatically predict candidate sRNAs in *B. bronchiseptica*, we submitted the RB50 genome to SIPHT [258], which integrates information about conservation of intergenic regions, prediction of intrinsic terminators, conservation of secondary RNA structure, potential transcription factor binding sites, and genomic synteny to identify candidate small regulatory RNAs. Through this, we identified 468 putative sRNAs, the sequences of which are included in Table A-1.

Table A-1: Candidate sRNAs predicted by SIPHT. The name, length, and sequence of sRNAs predicted by SIPHT are shown.

Name	Length	Sequence
>BbSIPHT001	192	ACCCGGCCGCGGGGCTGGCCGGCCCGCCCCGGCGGGCGCGCGCCGATCGC GCGCCAAGCCGCGCCATTGCTGGACTTGCCGCTCGGCGCACCGGCTCGCG AAAACGCGGACGCGAGCCCGGCAGGAACACCGGGGAAAACCTCGTTGAC ACTCAAAAACACTATCAGTATGCTGAGTAAAACGTCAGCATACT
>BbSIPHT002	50	CTTACAAGCGCGGGCTGCGGTACAGTCAGGCCGTGGCGGCCGCGCCGCGC T
>BbSIPHT003	187	CGGGCGCGCCGCGCGCCCGCAATCCCCGCCATGCGCGGCGTCGAAGGGGG TGGTGTCAATTCTTGTTGCGTTTCAACTCATCTTGTCGTGTGTTGCGGCAAG CCCCGGATAGAATTGCGTCCGCGCAATATTCGGAGGGCTTGCCATGAGGT ATTTCCTGATCCTGGTGGGGGCCGCTGTGGTGGCCT
>BbSIPHT004	170	GCCCGTGGCAAGACGTGCCTGTCCCCGCCGGGGGAGCGACATCTGCGGCC AGGGAGCCTGTCCCTGCCGGGGAGCGACACCTGCGGCCAGGGTGCCTGTC TCCGCCGGGGGAGCGACACCTGCGGCCAGGGTGCCTGTCCCCGCGGGGAC AGGCACCCGTCAAGCGACGGT
>BbSIPHT005	93	GCTGTCTCCTGCCTGTCCGTTGCGGCGGCGCGCCCGCCCGGGCGCGCCGCC ATTTTTTCTGGCTGTTGGTACTAGTGTTCCGTATCGCGGAACT
>BbSIPHT006	126	GCTGTCTCCTTGTTGTGCCTGGGATGTGAGGATGGAGACGACTATATCGAC GGCCGAAATCGGGCGGCAACCGACTATTTCGCAGCGGTGGCTCAGTTTTT CTCAGCCAGGTTTCTTCTGGTTGAGT
>BbSIPHT007	97	ACGGGCAATCTCCGAAACGGCGGACGGTCTCCCAGCGTTCTGTATTTGGG ACGCTGAGGGGCCGATCCCGCATCTAGGCGGAAGCCGGGCGCCGGCCT
>BbSIPHT008	59	GCATTGTCTCCGGATGCGGGGCGGCGGCAGGCGCGCGCCCGGGTCGTTCA CTGGGCGAAT

Name	Length	Sequence
>BbSIPHT009	293	CCGCCAACATGCGCCGAAACCTACGCCTGGCTTACAGGCTTGCAATTCCG GTGGACTTTGCCGACAATGTCACCTGATTGCCCCGGTTCCGACCCGAGCCG GGGTTTTGTTTTGGTCGACGCTTGGCCGCCGGATGCGGCAGGCCGATCAA AGAGGAGACAGCAAAGGGGAGCCTCGGTCGGGTTCGACCTTGTCTCGTCT TTTGTTCGCCTGTCTTCCGCAGCGGCCCGCCTGCTTCATGGCGAGGCCATA CACCAAGCCGAGACCTTCACCGAAACGCTCCGTCGGGGGCGT
>BbSIPHT010	73	CGATGGCGCGGCCGATCCCTGGCCGCGCCGCGGACAAGACCCTGCGCCGA TTGAACGCAGCCTGAGCGCTGCAT
>BbSIPHT011	71	GTCGACGTCCTTGAAGAGAGCGGGCGGTCGTGGCGACCGCCTCCTTACAT GACCCGGCGACTTGCGCCGGCT
>BbSIPHT012	61	AGCTTTCCCGCGCATTGCGGCCAGACCGAAAATCCCATGCAGATCACTGA TTTGCATGGGAT
>BbSIPHT013	99	GTGCCCGGCGGCGCCTTGCAGCCCCGCCGGATGGCCGCGCGGCGCAGCAT GCCTGCGCGATTCTTGCCCTGCATGGGCAAGGCCTTGATCGGCAAGGCGT
>BbSIPHT014	506	GATTTTGTGTCTGCAGCTTGGAATTAATCGGAAATCTGTCAGTACGCGGAG CAATGTATATCGTCTCCGAATATTGCAATTATTTGCGTTGTTTTCTATTCGT GATACCAGGGGAATCGCCATGAATGCTGCAAAACTCATGACAATCATTTC ATGGCAAATCAAATTTTTCCTGTTTTTGATCCATTAACGAGAAATTCTGCC ATCTAAGTGCCTAAAAAAGCGAGTTGCTACCAATTCGAAGAATATCCGTT GCACAAAATATGTCTCAATAATACGAATACCGGGGTAAACCCTGATGGCA ATGAAAAGTTAATAGATAAGACTAGATAAACATTTCAAAACAAAAGATTG CGGATTTTGAAGTTGTTATCCATAAAACAGCGCGGGTATTCGAGGCGGCA CATGAAATCGACTGGCGAGACGCGAGCGGTGCATCGGGCAAGGGGGCGG CCGTGAATCCGGGCGACATCGTGCCGCTTGCGCTGCAGCCCGGCGTGGTT GCGT
>BbSIPHT015	31	ACGGGGTTCCTTCCAGATGGGATGGAAGGGAT
>BbSIPHT016	59	GCGTTTTTGTTTATTGCGAAAAGCGCTAAAAGCTAGTATTGGCGCGGCTCA GGGCGCTGT
>BbSIPHT017	115	GATCTGCGGGGTGAGGGGTTGGAGGGACGCTTATCGCTGCGCTGCGCGGG GATTGTAGACCACCGGACGCGTGGCCGCGGCCGGCGGCCCGGCCCGCAGG CCGTAAAATACGGCAT
>BbSIPHT018	63	CCGTAGCGCCCGGCGCCGGTGCCGGGCGCGTCCATTGTGCGAAAGGGCCG CGATTTTTGCGGCT
>BbSIPHT019	199	TGTCGGGCCACGTCTCGGAGGCCTGTTGCGCCGCCGCCGGGCAGGACAGC GCCAGCGCCGCCGCCAGCGCAAGGGTTCGGGTGAGTGTCATGGCAGATTT CCAATGGATGAGGCGAAGAGGGTGTGCGTGTATTGGCGCGGCCCGGCGTT CAAGCGGCCAGGCGGTTGTCGTCCGGCGCGCGTCCGTCGACGGGCGTGCT
>BbSIPHT020	179	GTTATGCAATTTTATTTTGCTGTTGTCCATATGATCGTTGATTCGGCTGCCT GCAGGCAATAAATCGCTGCCTCATTGTGTCGCATCCGGGCTATGCTGTCAG CTGTCATTTCCGTTCATTCAGCCCTGTTCATTCATCGTGCCATATCTGTCGG GTCTCATGCGGCGCAGCGGCCGCCT
>BbSIPHT021	34	TTTCCTCCTTGGGGATTGGCAGGCAGCGGCTTGCT
>BbSIPHT022	75	GGACCGAGCCGGACGGACGCGAAAAAGAAAACAGCCGCCATTGCTGGCG GCTGTAAATGCCTGAATACTCAGGCCT
>BbSIPHT023	87	CTCTGCTCCTTGGGGGTTCGATCGTCGTCGTTGCGGCGACCTGTGCATGCA GTATGCTTGGCACCCTGGCGGAGCACTTTGCCGGCAT
>BbSIPHT024	39	CAATGCTCCTTCCGTATGCGCATGGGGCCGGGCCCCGACT
>BbSIPHT025	130	ATTTCTAGGGAAAACCCCGAGCAACTGGCGGAAAACCAGGGTTTCCAGCG CCCTTGCCCTCCATTACCATTCGGCCTTCAAAAATACGCAAGCGCCGATTT CAGAAAACCAAACGGCACCACCAAGTGCCT
>BbSIPHT026	45	AGGTGGGCCTCCAGTGGTACGGCGTCGTGCCGTGTAGCCATGGCGT
>BbSIPHT027	209	CCGCGCCTGTCCCGTCCTGTCGCCGGAGGTGCCTGTCCCCGCCGGGACAG GCACCCGGGCCCAGGAAAACCCCCACGCTGCGCCCGCTGTGCGGGCGTGC CTGTCCCCGCCGGGACAGGCACACTGTCGTACGCCATCGACGACAGGCCG AAAAAAAGCCCCGCGGCCTGCGGCCGGCGGGGCGTGTACGGCGCCGGGA CTGCCGGCGCT
>BbSIPHT028	82	TGATCCCGGGCGCTAATCTAAAATAGCACTCATCGTATCAATACCGCCCGC CAACATCCATGGCGGGCGCATGTTTCGCCACT

Name	Length	Sequence
>BbSIPHT029	164	AGCTGAGCATCGCCCAGCGCTTCTAGGAATGCCTGCGCGCGGCTTGCGGC GTGCTCCTCGCGCAACATCAGTGAAAACGTGAAGACCGCGCGCAGCGCAC CGCCCCCTTTTACAGAGCCGCGCGCAGTTGCGGCCAACTTGCAGCGGTCTG GCGCGTACGGCGCC
>BbSIPHT030	97	ACGCCGCCGCTGACATGGCGGCGTCATGTTTGCACCGCAGAATGAAAAGG ATTTCATTTGCGGCCAGAATTGAGCTTCCGCGCAACAAACAAGAAATG
>BbSIPHT031	129	ATGGGCATTTCCCTGCCCCGCGGTGATACAGTTCTGGCCCCGAACCGGTGC TTTCGGCAAGTACGCTTGCGCGGAAGGTAAACGGGAAACAGGAAGGCCCT GCCCAGCCTGTGCTGCCCCCGCAACGGTC
>BbSIPHT032	111	ACGACAGCGCCGCCTGCCGGACTGCCGACGCGACGACAGCACGCACTCTA AAGAGCGGCGACAGACAAAGCAATGGCGCGCGGGGCCTTGTGCAATTGC GCAACAGTTACAC
>BbSIPHT033	300	AGCAATACCCGCCGGATATTACACGTTGTCATACGTCCTGTTGATGGGGTT TTATCAAGATGCCGGCCTTGCGCCGCCGAGGCCGCGTCGATGCTGCGCCG CAAAAATCCGTCTCTCCGCCGGGCGGGCGCGCCGTGAATGGCGCGGGAAC GACAGGAGCCCGCCTGCGCATGACGAGAAGCCTGCCCCCGCATGGCGAGG GTGCCTGTCCCCGCCGGGACAGGCACCCCGTCATACGCCTGCCAAGACAG TGCCTGTCCCCACGGGGACAGGCACCCTGCCGGCGCCTGCTGGAGGACGA
>BbSIPHT034	114	ACGGCAGGCGCGCCGGCCCGCACCGTGCAAGCGCGTGCGCGGGAATCTTC GGCAAGCAAGGGCAGGCGCGGCCGGAACGCGCACAGCGGTCCGGCCGGC ACGGGACGCGGCAGGC
>BbSIPHT035	149	ACGGCCTGCCGGCCATCCATGGATCACGCATGCGGATCGCCGGCCGCCTT GCGCGGCCTCGGCCCGACCCAGGCCATGCTCAGCCACTATGATGCGCCAT CGGCGCGCGCCGCGCATGCGCCGGCGCAAAGCCGGCGGCATGCGTGCCGC
>BbSIPHT036	41	AGGCCGTCAGTGACTACTATCACTAACGGCCTTACAAATCAT
>BbSIPHT037	44	ATGGACGGGTAATTGACCGTCCATTTCCAATATATGGAGTGACAA
>BbSIPHT038	54	AGACACCTCTCAGGGCGCCGCCCCGCGCGCAGGCGGTATCCGGCCCGAAA TGGGA
>BbSIPHT039	496	ACGCACCAATATGGTGCAAAAAACCGTTTTCTGAGGCTTACATCGCCGATT TTCGCCAATTGTCGGTGAATTCGCCATCGAAATTGCCAAACGTGCACTTGT TTTTTTAAAAAGCGGTGAATCCACCAGCGCCCCACACGTGTGCATGGACG CCCGGCCGGCAAGGGGTTTTTTACGCCGCGTTTACGCCCAGCCTCGGAAC ATTTACCCCGTTTTTATCCGCCGATCCGTAACCTGCCAGCGTCCGCCGGTG CTCCGTGCACAACGCACCACCGGCCCAACCCCGCGGCCCCGCGCCCAAGG AGAACGCTACATGGATCTGGTTTACCTGGCCGTCTTTGCCGCCATGGCGGC CCTGACCGGCACGCAACTGCGCCTGTGCGCCCGACTGTTCGCGAGAGGCC GCCCATGACCGCGCTTTACTGGCTCGGCGGCCTGACGGCCGCGCTGCTCTT CGCCTACCTGGCCTACGCGCTGTTCAATCCGGAGAAATTCTG
>BbSIPHT040	399	AGCACAGGGGCACCCATGCCCGAGAAGGCCGGCCCAGGGAGGCAGTGGA TAGCGCTGGAACAGCCGGTTGTTCACTTTCATCGATCCGGGAGAGACCAA TGAGAAAGATCCAACTTTTCGCCACGGCGTGGCTGCTTCTGGCACTCGCCC TGCCGTTTCCGGCCTCGGCCCACGGCGGTGGCTGCCGGAAAGACTCGCCG CCTGGAAAATGCTGCCACATGGACAGAAAGGCAGGCACGATGCATTGCCA CTGAGCCTGGATTGGCAGATGCCCAGGCCATTGATAGCCTGGGCATCTGC CTGGACGACAAAAAGCCCATAGGCGCTGACGCATATGGGCTTTTTTGCTG GCTTTCTCTTGCGGGTCTTTCCCGCACCAGGGCCTGCCTGGACAAATCCT
>BbSIPHT041	73	ATTGCAGCGCTAGCCGCTACAATGAGGGGGCCTCGCATAGCGAGGCCCTT TTTTCTTGCCGCGCCCTCCCGGCG
>BbSIPHT042	220	ACCCGGCGCATTTGCGACTCACGCGCCAAATGGCGTCAATCACAGCAATC ACGCGCCAAAGCGTGCACAACGTGCATTCAAAAGCCGCTGCGCGCAACGC CGTGCTGCACGCGCACTGCCCTGCAAATGCACTGCGCGCCATGAAAACCG TTGATTTTACCAGATAAACCGGAAATTTGCCTGATACGTCGGGAGCTTACG GATAAGCGGGACACATCCGA
>BbSIPHT043	111	ACCAGACGCAAAGGAAGCATCTACTTTGGCATAACACACGACGCGCGAGC GCCCACGGCGCCGGCCAGGCCCTAGACTGCCACGCAATTCGCAGTCCGGC GCTTCGCTCGTT
>BbSIPHT044	146	AGGCCGGTTCGCCGGCACGACGCGCTGCGCTGGTGTCGCAGCCCGGCCCG CCTGCGGGCCGTGGCGCGTCCTTTCTGGCCCTTGGGCCCAAAGACGCGCC GGTACGCCATGCCACCTGGCGGGCATGCCGGGCAACGCGCGGGGTTG

Name	Length	Sequence
>BbSIPHT045	281	ACATCAACGCTGGCATCGATATGCCCTCGAACCGTTCTGGATACCGGCGG GATGCCGGGCTACCCGGACGGCGGCGCGTGCCCGATGGCCCGTGCCGTCA TCCGCGCCCGATTAGTGGAACAGGCAACGCTCGCGTTCCATGTATCAGAA CCGCGGGCGGCAATCAGCAAGAATCAAGACTCTCGGGCCGCGCAGGCGCC GGAAAAGAAAACAGCCGCCATTGCTGGCGGCTGCGGAAGGACTTGCACCT GCGGCAGAATCCAGGACACCACGCGAGCCTGA
>BbSIPHT046	39	AGGCTCACACCACATATCGAGCCGCCATACAGGAGGAGTC
>BbSIPHT047	154	AGGCGTAACGCCGCCCCGCCCGGCCGCCGCCCCGGCGGCCGCGGCATTGC CGCGCATCCGCGCCGACCGGCGCGGCGCCACGGCATCGTTTCCGCGCCTG GCCGGCACGGGATTTCCAGGCGGCCATCGGCGCGCCCGCCGGCCGCCGCA CGCGC
>BbSIPHT048	32	ACCCGCAATCGGCGGGACAACGGGAGACAGGAA
>BbSIPHT049	81	ACGGCGCGGCGAACGCCTTGGCATGCGTTGCTAGACTTTTCCGTGGGCGG CGATCCGCCGCCGGTTTTCCTAGGAGGTGTTC
>BbSIPHT050	94	AGGGCGCGGCGGCGCCGCCTGTCAAAGTGGTGGAGGAAAGGGCGGATGG ACCGATTTTATGCGCTTTGCCGCGGCGACGGGGCCGGACGCCCCGG
>BbSIPHT051	70	ACTCGGGAAAATCCTCCCAAATATAAGTAAAATTTGTAATAATTTGCGAG TAACGCGTCGCGCTGCGCGGA
>BbSIPHT052	56	ATTCCTCGCTGGAACAAGGACTTGTCCGGGGCGGGCAATCCGGGAGCGGG CGGCGGG
>BbSIPHT053	159	ACTATCCACAGAAATTGTGGATAATTCTGTGCAAAACTTGCCTCGAACGC AGTCCCGCATTGCGTTTGCGCTAGGCTGCTCGCAATGTATGCAGCATGCGT GACACTATTTTCAGTGTAAAAACAAAGACTTGCACGGAATTTGCTGGCGTT GCGGGCAG
>BbSIPHT054	64	ATCCCGCATCGGCGGCGGGCTCCGCACGGCGGCGGCCCGTCCATTTTCCC GGAAGCAGCAAAATC
>BbSIPHT055	48	ACCCGCGGGGCGGGTGTAGACGTGGCGGACTGGGCCCGTTTGGACGGGT
>BbSIPHT056	76	ACGCCGCGGCAGACGGCGCCCGGGCACAGGAGGCGCGACGCAATCCCAC CCGGACTGCGCGCCGTGGCCACCATTGA
>BbSIPHT057	70	ACAGGCCGGACGGCGCGGCAGCCCGCCGCGCCGTATGGCGGGGCGGGGC ATAGGCGGCTACGCTTGTATTA
>BbSIPHT058	73	ACGGGCAGATATTGCGCCCAGCCTCGCAACAACGCATCAATCGCGCGCAT GCCCCTCACTAGAATTTTGCCTGT
>BbSIPHT059	72	AGTCGTCTCGTTCGACAACTGACGCGGCGGCGCCAGCCGCGCTCAATCAT TCAATGATGGAAGACGGTCACTC
>BbSIPHT060	39	ATCCCCCGGTGCGGGGCTTGTACTGCGAAATCTTCTAACG
>BbSIPHT061	55	AGCGGCAATAATACCGCGAATCCCGTAGGGATTTTCCTGACAAATCTCGTT GGTTA
>BbSIPHT062	116	ATCGTCGAAATTTTCCACCGTTCCGACGCCTGCATGGCGGAGGCAAAAAA AACGGCGCCCGTCCCTCTGTACAGGGCGGGCGCCGGCACTTACGGCTGCG GCGAGGGGGGCGCCGGG
>BbSIPHT063	139	CCGTTCGGCCCCGGAAAATGTTACAGCGTGGTGGATAATTCTGCGCCGCC GGCGGTTGGCGCGGCCATGCGTTATCGTCGATAATGTCGGTCATCACAGG GCTTCACCCAGTATGCGCCGGTCACCGACCGGCGTGTTTC
>BbSIPHT064	111	CACACAGGTGCCTGTCCCCGCACGGGGGACAGGCACCTGCCGAATCCTGG GTGCCTGTCCTTCCCGTGGGGACAGGCTCCCGCCAACGGGGAGGAGCCCT CCCCCCCGCCGC
>BbSIPHT065	60	GAAAACCCTTGACCCGGGTTTCCCCTGACCCCCCGAGGGCCGCGCTACGC GGCCCTTTGCG
>BbSIPHT066	90	CCAGGCGCGATAGCCGTGCGGACCGGGCGAGTTCGCCGGGTCCCGCGCAT CGCCCTCCAAGGCCGCCCGGCATGCCGGCGCGGCCTTTGTT
>BbSIPHT067	72	TAGAACGGTGTGCCTGTCCCAGTGGGGACAGGCACACGCTGCCGCAGGTG TCCATCCCACGTGGTGCGGCCCC
>BbSIPHT068	101	TAAAAACACCGGAAGGCGGTGGTTCATGGACTGTATGTTATCTACTAGTA GATAGATAAACAAGCCATGGCCCCATTGTTTTTTCGTATCCCGCCGATAGC C
>BbSIPHT069	75	ATGACAGTGTGCCTGTCCCCGCGGGGACAGGCACACTTGCCGGTGCCAGT CCCCCTTTGGGGACAGGCACGCGGCG

Name	Length	Sequence
>BbSIPHT070	117	ATTCAACGTGTGCCTGTCCCCCTGCGGGGACAGGCACACGTTGAATCCTGC TTCGGGGGCAGGCACCCTGAGAAGTCAGCGGGTGCCAGTCCCCCTTCGGG GACAGGCACCCGGGGGG
>BbSIPHT071	92	TTTCTCGGGTGCCTGTCCCCGCCGGGGACAGGCACCTGGACACCCAGATTC AACGTGTGCCTGTCCCCCTGCGGGGACAGGCACACGTTGAAT
>BbSIPHT072	114	CAGCGAGTGTTCATGGTGTAGAGCGGTGTGCCTGTCCCAGCGGGGACAGG CACACGCTGGCGCGGATGCCAGTTCCCATCCGGCGACCGGCGCGCCGCGC GCCGGTCGCTGCCTT
>BbSIPHT073	308	GGGGGCTGCCTGCGGGTGCCTGTCCCAGCGGGGACAGGCACCTTGGGGCA GGCACTCGGGCGTCTGCTGGTGCCTGTCCCAGCGGGGACAGGCACCCCGG GAGTCAGCAGGTGCCTGTCCGCGTAGGGACAGGCACACCGGAAGTCGGCA GGTGCCTGTCCTCGTGGGGACAGGCACCCCGGGAGTCAGCAGGTGCCTGT CCGCGTAGGGACAGGCACACCGGAAGTCAGCAGGTGCCTGTCCGCGTAGG GACAGGCACACCGGAAGTCAGCTGGTGCCTGTCCGCGCGGGGACAGGCAC CAGATTTAT
>BbSIPHT074	99	GAATGACGGGGTGCCTGTCCCCGCGGGGACAGGCACCCAGGCCGCAGGTG TCGCTCCCCGTGCGGGGACAGGCACCCAGGCCGCAGCGGGGACAGGCGCC
>BbSIPHT075	109	GCGCGCGGCCTGGGGAGCGACAGCCGCGGCAAAGGGAGCCCGTCCCCGTT GGGACAGGCACAAAAGATAGCGGGAGCCTGTCCCCGTTGGGACAGGCTCC CGAGTATTGG
>BbSIPHT076	87	CTGTCCCCGTGGGGACAGGCACCCGCACGGGCAGGCGGCACCCGCACGGG CGGGCGCCAGTCCTCATGCGGGACAGGCGCCCTTTGCC
>BbSIPHT077	97	ACAGCAGTGTGCCTGTCCCGGCGGGGACAGGCACACTTGCCGCTGGAGTC GCTCTCCGTGTGGGGACAGGCACCCCACCGGCGCGCCGGGAAAGGGGG
>BbSIPHT078	111	CCATACGCCGGCTGATCGGCCGGCAGGCCGCGCCCACGCTGGATGCGCGG TCGCCGGGCCGCGCCGGCGCAGGCGCGGCCGCGCGCCGTCACCGGCGTGT CATGGCATTTCC
>BbSIPHT079	175	ATTTACGGTTGCCGTTTGCGCTTTTTGCGCTCCTGCAGCCGGCAAGCAGTA CCGAAGCGGCTTGAGAAGACATTCTCAAGCCGTTTTTTTGCGCCTGTCCCG TGCGGGGCAGAGGCCGCTTTGGCGGCGGGTGCCTGTCCCGGCGGGGACAG GCACCCCTGCCGCAGGTGTCGCTT
>BbSIPHT080	139	CGCTGCCTGGCGCCGGCAACAGGTGCCTGTCCCGACATACAGGGTGTGCC TGTCCCCGCCGGGACAGGCACACTGCCGTACACCATCGACGCTTGACCGC CAAGACGGTGTCTGTCCCGGTTGGGACAGGCACCTGTTGC
>BbSIPHT081	130	TCCCGCCGGGAGCGTGGTGTGCCTGTCCCGGCGGGGACAGGCACACTGCA CGGGCAGGCGCCAATCCCCATGCGGGACAGGTACCCCGCCCGGCGCAAAA AAAAACGGGCCGCGCATGCGGCCCGTTTCAG
>BbSIPHT082	112	TTAAAAGTCGCCCTGCGCGGCCCTGGCGCCGAATATGTCCACGATAGGGC GCGCAGGGCATCAATAGGCAAGACTGGGTAATACTTCGGCGGTTTCGGGG TCTTGGCGGGCGC
>BbSIPHT083	104	GAATCAAGGCAGCAAATCAAGGCAGCAAAATCAAGGTAACAATCCTGCG CAAGCAGGCGCAATAACCCCGGCGAGGGCGCATCGGCGCACCGCACGGG GTTTTTT
>BbSIPHT084	72	TGCAAAGACCGGCCGTAAGGGCCGGTCTTTGTTGTGATTTTCACTATTTCA ACGTGTGTATCAATGGAAATGT
>BbSIPHT085	80	CCGTAGCGCCCGGCGCCGGTGCCGGGCGCGTCCATTGTGCGAAAGGGCCG CGATTTTTGCGGCTTTTCCGCACAATTTGTC
>BbSIPHT086	61	GATTCGGCGTATGCCGATGGCCCGTAAGGGCGCAACTGGAACCCGCCACA TGGCGGTTTTTT
>BbSIPHT087	199	GCGTCTCACGTTGAGCGCCGTGCGCGTGCATACCCACCTATTGGTGACCAG CTAGATGCTGCCATGAAGCTGGCGGCTGCGCTGCAGGTGGCAGGCGTGCC GTTGCCCGATGAGGTGGCTCGCTGGATCGATCAGTGCAAGGCAGTGAAGC ATAGGTACCCGAAATCCAATGAATGAACGGGCCCATAGGGGGCTTTTTT
>BbSIPHT088	173	GACACAGGAGCCTGTCCCGGCGGGGACAGGCTCCTGCTGAATCCTATGCC GGGGTGGCTGTCCCGGCGGGGACAGGCTCCTGCTGAATCCTGTATGCCGG GGTGCCTGTCCCCGTCGGGACAGGCACCCTTGCCGCAGGTTTCGCTGCCCA TGCGGGGACGGCCTCCCGGAGGG
>BbSIPHT089	141	ACAGGAGCCTGTCCCGGCGGGGACAGGCTCCTGCTGAATCCTATGCCGGG GTGGCTGTCCCGGCGGGGACAGGCTCCTGCTGAATCCTGTATGCCGGGGT GCCTGTCCCCGTCGGGACAGGCACCCTTGCCGCAGGTTTCGC

Name	Length	Sequence
>BbSIPHT090	140	GGAAAGCCGCCATCGCCCGGCCCCCCGCCCGGCATGGGCCGGGCGGGGG GCCGAAGGACGAAGGATGTTGCTATTAAACGTGATAGCCCAAAATTTGTG TACAAATTTGATCGTGCGCCGGCCCGGAGTGTGGCGCGCCCG
>BbSIPHT091	43	AGGAAGCGGGCGCGCGCAGGCGTGCGCGCGCCTGGCCGGGAGGC
>BbSIPHT092	155	ATATACCCGGGAGCCTGTCCCCGGACGGGGACTGGCTCCCGCTGACTTCC ACCTGCCGGCGCCCCAGGTGCCAGTCCCCGCTGGGACAGGCACCCCATGC TTCTACTTGCCGGCGCCCCAGGTGCCCGTCCCCCTGCGGGGACAGGCACCC CGTCA
>BbSIPHT093	67	GGCATGGCAGCGGGCGGGCGCGCCCGCCCGGCGAACGAACCGGCGCC CTGGTGGCGCCGGTTTCTT
>BbSIPHT094	121	CGCGATTCGGCGCACAATGGGATGGTCCGGCCGGCGTGCAGGCGGCGGGC CGGACACGCAGCCCCATGCCGGCCGTCCGGCTTGACGCGGGTCAAGCCGG TGGATCCGGCAGGGCGTTATGG
>BbSIPHT095	80	GCGGCGCCGGCCAGGCGGACCGCGACATTGCTATGATGGCCGGCTTGGCC CGCGTCCGGCGGGACGGGCCGCCCATTCGTC
>BbSIPHT096	71	AGCACACCGCCAGCGCGGTGCTGCGCTGGCGGTTCCGAACGGGACCCAGT GCTGCCTGCGGCGACGCGCCGC
>BbSIPHT097	81	ACGAAAGCGGCCTGGCGCGTCTTGCCGCGCCAGGCCGCAATCACTCGCCC TCAATCGGGCGCGCAGGTTCTGAAGCCGCCGC
>BbSIPHT098	153	GGGCTTTTCGCCCGGCGCCAGGCGGCCTGGCGCATGATGCGTCATGCGCC GCCGCAAACGGGCCCTGACTTCGCGGCATGAGCGGGTTCGGAGCCTCGTG CGTTTGGCCTGTGCCTGTAGCGGCCGGCCGGCATTACCGGCCGGCCGCCTG TAT
>BbSIPHT099	209	GCAGTGGCCAGGACGGACGACGGGGTGCCTGTCCCGGCGGTCGAGTGCGC ATGGCATGCCACGGTGTGCCTGTCCCAGCGGGGACAGGCACACCAGGTCA GCGGGGAGGACAGGCACCCCGGATTCAGCGCGTGCCAGTCCCCCTGCGGG GCAGGCACCCCGGATTCAGCGCGTGCCAGTCCCCATGCGGGGACAGGCAC GCGGGTTTAC
>BbSIPHT100	120	TGAAAAGCACATTCCATGCCAGCCCGCCCGCGCCGGCATGGCGCGCGCGC ACGGCGCCGCGGCGCACCGCAGCGGTGCACCCCGGCCCTGCTCCGCCTTG TTTCGGCGCATGGCGCGCCAG
>BbSIPHT101	85	TGACAAGGCCGCCTGCGGAACGACCCGCGGCGGCCTGGCGAATAATAACC GAGGCCGCGCCACAACGGGGCGCGGCCCGCGGCGCC
>BbSIPHT102	87	AAGGAACCCCCGGCGCGGCCGGGGGTTCCTGCCCATGGCAAGCGCAATGG ACACAGCGGCGCGAGCGCACAGGCGGCGCCGGCACGGT
>BbSIPHT103	137	AAGAAAGCGCTGGCATAGGCAGCGCTTTCAAGAAATGCAGGCGCTGGCG GGCATGGCAAAGCCCCGGCCAGACGCGCACGGGGGCACGGCCGGGGGTA CGCTTACGCGCGCCCGCACCCGCCGCACCCGGCAGCGGGC
>BbSIPHT104	229	GGCAGATGCCGCCCCCCCGGCAGGCGGGCGTCGATGGCATACGGCAAGTG TGCCTGTCCCGGCGGGGACAGGCACCCAGCCGCAGATGTCAGTCCCCATG CCGCTGGCGCATGACGGCGTGCCTGTCCCGGCGGGGACAGGCACCCTGGC CGCAGGTGTCGCTCCCCCTGCGGGCACAGGCACCGCGTCGCACGCCCAAA AAAAACCCGGGCCAGCGGCCCGGGTCTGTA
>BbSIPHT105	113	AACACACGCGCAATGTGCCTGTCCCAGTGGGGACAGGCACACGCACGGGC AGACACCATGCAATCGTGCAGGCGCCCGTCCCCATGCGGGACTGGCTCAC AGCGCAAGCGCCCG
>BbSIPHT106	101	GGAGCGACACCTGTGGCAAGGGTGCCTGTGTCCCCGGACGGGAGCTGACA TCCGCGGCGAGGTGTGCCTGTCCCCACTGGGACAGGCACACCGCAGTGTT GA
>BbSIPHT107	179	CTGAAACCCTCAGGTGCCTGTCCCCACGAGGACAGGCACCTGCTGACTTC CGGGGTGCCTGTCCCCACGAGGACAGGCACCTGCCGACTCCCGGGGTGCC TGTCCCCACGCGGACAGGCACCTGCCGACTTCCGGGGTGCCTGTCCCTACG CGGACAGGCACCTGCCGACTTCCGCCCCC
>BbSIPHT108	62	AAAAAACCGCCATGTGGCGGGTTCCAGTTGCGCCCTTACGGGCCATCGGC ATACGCCGAATCT
>BbSIPHT109	124	CACAACCTTCCGGTGCCTGTCCCCGCATGGGGACTGGCACCGGCAAACTT CCCAGGTGCCTGTCCCCGCTGGGACAGGCACCTGGGTACCTGGGCCCTGG GCACCTGGGCACCCGGGCCCCTGCG
>BbSIPHT110	82	GCAAACTTCCCAGGTGCCTGTCCCCGCTGGGACAGGCACCTGGGTACCTG GGCCCTGGGCACCTGGGCACCCGGGCCCCTGCG

Name	Length	Sequence
>BbSIPHT111	112	AATAGACGGTGGCGTGCGCCGCGCTATTCCGTCGCCGGGCGACGGCAGCG CGCGCATGGCGCCGATCCAGGTAAGACAGCGCACCGCGCCGGCCGTTCCC GCAGCCGGCTCGC
>BbSIPHT112	72	CACAAAGCGGGTTGCGGCAGCGCCGCAACCCGCTTTTTCCTGCTGTGCCCC ACCCGCGCGCAATGCGCGGCGC
>BbSIPHT113	104	ATACGACGCACCGGCCCTCAGGCCGGTGCGAGGTGAAGCAGGAGAACGC ACCGCCACGCCGGTCAGGCGTCGCCGCGCGGCGCGTCCCTTGATGCCGAC AGCGGT
>BbSIPHT114	133	CACATAGGGTGCCTGTCCCCGCGGGGACAGGCACCCTGTCAACGCCATGA ACACTTGACGGCCAAGACGGTATCTGTCCCGGCGGGGACAGGCGCCTCGC GACGGCCTCCCGCCGCCCGGCCGGCCCGCGCGCC
>BbSIPHT115	102	GACAGAGTGTGCCTGTCCCCGCTGGGACAGGCACACTGCTCTACGCCATG AACGGCTGACTGCCAGGACGGTACCGCCCCCAGCGGGACAGGCGCCGCAT GCG
>BbSIPHT116	105	GGGCCGCATGGGGACTGGCGCAGGCTGACTTCCAGCCTGCCCCCCGATGG AAACTGACATCCGCGACAGAGTGTGCCTGTCCCCGCTGGGACAGGCACAC TGCTCT
>BbSIPHT117	69	AAAAAAGGAGCTTAGGCTCCTTTTTTCGTTTTGGCGCGCGCCACGGCAATC CCACCCACAGCCAACGCTC
>BbSIPHT118	143	GAACACTCGGGTGCCTGTCCCCGCACGGGGGGACAGGCACCCGCTGAATCT TGTTCCCGCACGCGGGACAGCCACCCGCTGAACCCATGCCGGCCGCCTCC CCCGACGGCGCCTCGGCAGCCACGAAAAACCCGGCCGAAGCCGG
>BbSIPHT119	253	ACTAACTCCGGTGCCTGTCCCCGCAGGGGGACTGGCACCGGCAAACTCCT TGCCAGTCCCCGCATGCGGACTGGCTCCGGCAAACTCCTTGCCTGTCCCCG CATGCGGACTGGCACCGGCAAACTTCCGTGCCTGCCGTGCAGAGGGACTG GCTCCGGCAAACTCCTCCGGTGCCTGTCCCCGCTGCGGCAAGGGTGCCTGT CCCCGCGGGGACAGGCACCCCGTCATACACCATGCGCACTTGGCCGCCAC GC
>BbSIPHT120	157	TAAGAAGGCGTGCCTGTCCCCGTATGGGGACAGGCGCCCGCCAACTTACC GGGGTGCCTCTTCCCGCAGGAATCCGGGTGCCTGTCCCAGCGGGGACAGG CACCCCGTCATACGCCAGCACGCTTGACCGCAGCGTACCGGCGCGCCGGC GCCACTGC
>BbSIPHT121	91	CAGGAATCCGGGTGCCTGTCCCAGCGGGGACAGGCACCCCGTCATACGCC AGCACGCTTGACCGCAGCGTACCGGCGCGCCGGCGCCACTGC
>BbSIPHT122	81	ACAGCAGGGTGCCTGTCCCGGCTGGGACAGGCACCCTTGCCGCGGGTGTC GCTCCTCCCCAACGGGGACGGGCGCCCGCTTC
>BbSIPHT123	85	TCAAACCGCAGTGTGCCTGTCCCAGCGGGGACAGGCACACTGGCCGCAGT TGGGACAGGCACCCACGCACCCGGAGCCTGCGCCCG
>BbSIPHT124	160	AAGCTGCCCGCCGGCCCTACCTCACCGGCGAGCGGTGCGCGGCAGACCTC CCCGCGCCGAGGTTAAGCACACCGACCCTCGTTATGAACATCGGGTGCAG TGACGGGCCGGCTGGCTCCCAGTGCCGTCACCCTCAACACCGCCTCCGGG CGGTGTTGTTG
>BbSIPHT125	62	GGATAAGACCCGATGCGTGCCGGCAACGGGCCGGCAACGCATCGGGTCG GAAGGGAACGCGCC
>BbSIPHT126	160	GGAAAACTGATCCCTCCGGTGCCTGTCCCCGTAGGGGGACTGGCACCGGC CAAATCCGACCTCCGGTGCCTGTCCCCGCAGGGAGGGTGCCTGTCCCCGC CGGGACAGGCACACTGCTGTACGCCATCGACGCTTGACGGCCAAGACGGT ATCTGTCCCCC
>BbSIPHT127	108	CCCTCCGGTGCCTGTCCCCGTAGGGGGACTGGCACCGGCCAAATCCGACC TCCGGTGCCTGTCCCCGCAGGGAGGGTGCCTGTCCCCGCCGGGACAGGCA CACTGCTGT
>BbSIPHT128	176	ATGCGGCGCCGCAGCGCGGCGCCGCCGTGCGCGCGCGCCGGCGACCGGCC ATCCCCGGGGCACCGCGGCGGTGCTGCTGCGCACCAATATCCGGCGCGCA ATCCCCCTTTTCGTGCTGCTTTAAAAAAACGGAGCTGGTTGGTTCGTTTTA ATTATTTTCGCGCGCCGCGCTTTTTT
>BbSIPHT129	265	CGATCACTTCGGCGGGCGCGCCTTGCAGGGCGCGCCCGCCGCATTTAGTA ACAACTAAATTTGAACTTGTATCGCTCAGTGTCTTGTCCCGGAACCGGACG CTCACTAGAATTAGTGCCATACGAGGGGGTGGTACACTACATCTTGTGGTC GCTAAACACGCTGGACCCACCGGCCACGCACCCCGGACCATTCGTCCACG CGTAGCGATTCCCCTCACACCGAATACTTTTGCACCCGGCCAGGCGGCCCG CCTGGCCTTCCTC

Name	Length	Sequence
>BbSIPHT130	155	CTGCGCCGGTCGCAGGCAAGCCGGATGTGCCCCGGTGACTGGTGTTCTGA CACGTAGAACGTAGCGGTAGGCGACCTCTGGCGAGGCCGCCTATCGGCGC CATTTGAAATGGCTCGCGCCGACGAGGTTGCGGGGGCCATATCAAATGGC TGTGCT
>BbSIPHT131	110	CCATAAGCGAACCGGCCCTGGCGGGCCGGTTCACGGTGCAGCCCGGCCGT CGGGCCGGGACCGCGGGGCCGGTTCACGCGGGAACGGGCCCAGGCCATC CCGCCGTGGCGG
>BbSIPHT132	154	GGACAACCCGGCGTACAGGCTGTTCCTGTAGGCCGGGGGGAGGAGGTGCC TGTCCCCGCAGGGACAGGCACCAAAAATTCAGCGGGTGCCTGTCCCCCTG CGGGGACAAAAATTCAGCGGGTGCCTGTCCCCCTTCGGGGACAGGCACCC GAGTA
>BbSIPHT133	103	ACACCACGGTGCCAGTCCCCGCAAGGGGGACAGGCACCGCAAATCTCCAT GTGCCTGTCCCCCTTGCGGGGACTGGCACCCAGCCCAGGGCATCGAGCCC CGCC
>BbSIPHT134	78	AGTAAACCGCGGCGTCCGCCGGCCTCGCCGGCGGACGCTGTCCGCGCCTT GCCTTGCGCCTTACCCGTCTTTCATCAAA
>BbSIPHT135	68	CGACGCCGTTGCCTCCCTCGCGGCGCCGCTGCGCTGCGTCTGGCCTGCCTG GCAGACGCGGCGTTTCTT
>BbSIPHT136	40	TGGAAACGCCGCGCATGCGCGCGGCGTCGGCAATCGGGCGG
>BbSIPHT137	113	GCCAGGCCGGCCGGCCTGGCCCGGCGCCCCGGCCCATTTCCCTGCCATGA CCTGGCCCGCGCCTGCCCGGCGGCGGGCCGTCGGCGCCGGCGCGCCGGCA GGCACCGGTTGCAT
>BbSIPHT138	83	CGGCATGTACCGATAGGGCTGTACGCCTATCCCGCCTGGCAATTGGGGCC AGCGCCTGCTAATTTGGCGGCGCTGGCCTTTCGC
>BbSIPHT139	67	CAGAAAGCCCGGTTCATGCGAACCGGGCTTTTTGCTGCCTGCCCGGACGC CGGCGCCCGATGCCCTTG
>BbSIPHT140	64	TGCAAAGGGCGCACCGCATGGTGCGCCCTTTTTCATGCCATGGCCGCCGA AGCGGCCACGATGCC
>BbSIPHT141	158	GGAGGACGGGTGCCTGTCCCCGGATGGAGCATGGGTGCCTGTCCCCGGAA GGGGGACAGGCACTCATGCAATCCGCGGGAGCCAGTCCCCCTGCGGACAC AGGCTCCCCAAAGTTTGCGGGAGCCAGTCCCCCTGCGGGGACAGGCTCCC GGGTATGGT
>BbSIPHT142	42	GTCGGATTGTGCGGCGTCTAGGGTTGGCTGGCCGCGTTTTTAA
>BbSIPHT143	87	GAAAAACCCCGCCAGGCGAACCTGGCGGGGTTTTTTTCGGGAAGGCCGCC GCGGCAAGCCGCGGCGGTTTCGGCCGGGAAGGCGGGAC
>BbSIPHT144	222	CCGCTGGCACATAGAACTTTGGGCCGCATCCGGACTTGTCCGGATGCTGCT GTCAAAGACCGCTGGAGCGCTAGCCGGTTGTTTGGCGAATGCTTAATCCC GAAGCTGGTTCAGTTCGGATCCAGCGTAGACGGCGTCCCCAGGAAGATTT CTTTACCCGAAGAACTCAACCAGGCGCCGCATTCGGTTGACGCGCAAGGC TAGCGCCCCAAGCGTCTTTTGA
>BbSIPHT145	97	CAAAAACCCGCTGCCGCCGAGCGGATCGGCGGACAGTTCGAAGGTGCGTC AGTTTACCCGAAGCCCCGCGCCGCTGCGGCCGCCGCCGTCCGCGCCGC
>BbSIPHT146	111	GTCTTGCTCGCCGGCCATGTTCGCGGCCGCGACGGGCAACGGCATCACGC CGCGCGGGTCCGATCATTGCCAAATGGTAATGAGCGGGTAATACCGCCTG CGCGTCTTTTGA
>BbSIPHT147	86	TGAAAACATGACGCTCCCGTGGCGCGCGGCGCCACGGGGGGCGGCTTGCG GCCCTGTTTTCCCGCGCACGCGCGCCGCGCGGCCGCG
>BbSIPHT148	119	CCCGCGCCGAACTCCCGCCATCGAACGGCCGCCCGCAGTCTGCTTTGGCC ACCATGCAGCAAGCACAAGGGGAATCGCATGAAAACGCTCCATCGGACG GCCCTGGCTGTCCTGTTTTCC
>BbSIPHT149	202	GGTTCGGATAGGAACGCCCGCGCGGGCGGGCCGGAAGGGTTCTGGTCTGG CCGCGCAGCGTACCCCTTCTGGGGCGGCATTGGTGCAAACACCAATGTTCT CACGTCGGCTCTGTTCTTATCATAATCGAAAATAGAGACAAAGTTCCGAA AATCGAGACTTATTTTTCAGGGAGCGGCCCGGCCGGACGCACCTGGTTTTC G
>BbSIPHT150	101	GATATTCATTAAACCAATTAATTTCATAATACTTGAACAAACCCCGATACT GGCAGCATGTTCAAGCAAGGAGGATTGGGTATGCCCATCATCCGGTTTTA T

Name	Length	Sequence
>BbSIPHT151	210	GCGGGCAGGGAGGGCGGCAATGCCTTGCATCTTAGCGTGCGGCCGCTGGG GTGCCTGTCCCCGTGGGGACAGGCACCATGCCACCGTGCCACCGCCCGGA AACGCCAACAACTGATATGCTTTTTTTACCATCAATTGAATCTGTACTGAT TTCGGCATGCTGCGTAGCATCAAAGCCGTGCTCCGGGACCTTCCCGGAGA CCGTTTTTGA
>BbSIPHT152	147	AAAAAACGTGCCGCTTCATGTCCGCGCGGCGCGAGGCCGTCACTTTTTTCG GGGCGAGATGCTAAGGTGTCTGCCACCCTGGCGCGTGGGCGAGGCGTTGC GGCGCCTGCCTGCGGCCGGTCGGTCATCAACCCATAGCGAGGCAAGC
>BbSIPHT153	40	AGAAAACCCGCCCTCGAGGCGGGTTTTCCTTTTCTGTCTCC
>BbSIPHT154	111	GGTACGGGCCTGGCGGGGAAGCGGAAGGAATGGCACGAATATAGGCCAA CCCGGGCATGCGCAGGGTTGGGCGGGGCGGCCGGCGGGGCGCGGCCGAT TGCGGCACTTTTCT
>BbSIPHT155	169	CGGCCATCCCCACGAGGGGGGGCCTGTCCCCACGGGGACAGGCGCCCCTC TCTCCGACATTTCCTGCTTTACGTGGAGCAAATACCTGACGCTGGCGCGGA CATCTGCGAGTTACTCTGGCTTGCCTTCGTAGCAACCTCTGCCGGCGTTCA TCCGCCGGTTTTTTTTAT
>BbSIPHT156	66	CCGGGACGGCGCGCACGCGCAGTGTGCGTGGAACGCATCGGCTCGTCGCG CGAGCCGGTGTTTTTCT
>BbSIPHT157	80	GCGTGGGCGTGGCACGCGTGTGCCGCGCAATCCTGCATTATGGCCAAAGC CGCGCCGTGGCCGGGCCGCTGGCGGTTTTGA
>BbSIPHT158	233	CGCCGGCGTGTGCCCATGACGAACGGCCGCCGCGAGCGGCCGTTTGCTTG TCTGGCGGCGGATCGGTGCCTGTCCCTGCGGGGACAGGCACGCCGGGGAA GTGGCAGGCGTGGAGGCGGATCGGTGCCTGTCCCTGCGGGGACAGGCACC CCGAGAGTTTGCGGGTGCCAGTCCCCCTGCGGGGACAGGCACCCGGAAAG AAAAAAGCCAGTTGCGTGAGCAACTGGCTTTTCG
>BbSIPHT159	237	GGGGGGCGGTTCGCGCCTGTCCCTGCGGGAAGTGGCAGGCGCGGAGGTGG ATCGGTGCCTGTCCCTGCGGGGACGGGCACGCCGGGGAAGTGGCAGGCGT GGAGGTGGATCGGTGCCTGTCCCTGCGGGGACAGGCACCTCGGGGAAGTG GCTGGCGTGGAGGCGGATCGGTGCCTGTCCCTGCGGGGACAGGCACCCGG AAAGAAAAAAGCCAGTTGCGTGAGCAACTGGCTTTTCG
>BbSIPHT160	145	GGAAGCGCCTTTCGGGTGCCTGTCCCCGTGGGGACAGGCACCCTGGCCGC AGGTGTCGCTGCCCATGCGGGGACAGGCGCCTCGCGTATCGGGCTGCATG TAAAAAACCCTGCCCGGCCGTATTTGGCCGGGCAGGGTTTTTTTAC
>BbSIPHT161	85	CGAAAATGTGTCCGCGTGCGCCGCGGCGCGCCAGCTTAGGTATGCTTGCA TTCTTGAGCGAAAGAACATGTAACAGGCAGGGAGCG
>BbSIPHT162	55	AAAAAAGCCCCCTCGGGGGGCAGCAAGCCCGCACAGCGGGCGCAGCGTG GGGGCCT
>BbSIPHT163	82	TGCCGCCGGGACGCCCCAAGGCAAAAAAGCCCCCTCGGGGGGCAGCAAG CCCGCACAGCGGGCGCAGCGTGGGGGCCTTTTTT
>BbSIPHT164	53	AAAAAAGCCCCCTCGGGGGGCAGCAAGCCCGCACAGCGGGCGCAGCGTG GGGGC
>BbSIPHT165	54	AAAAAAGCCCCCTCGGGGGGCAGCAAGCCCGCACAGCGGGCGCAGCGTG GGGGCC
>BbSIPHT166	82	TGCCGCCGGGCCGCTCCAAGGCAAAAAAGCCCCCTCGGGGGGCAGCAAG CCCGCACAGCGGGCGCAGCGTGGGGGCCTTTTTT
>BbSIPHT167	120	ATCGCAGTCTCGGAAAGACGGATATTCCCAATATACCCGTTGCCCGGCGG TCCGGCGCGGCGTCCGATGGGGTTATGGATGGAACCGGGCGGCGTCCTCG GAGCCGGCGGTGGTGTCGCCC
>BbSIPHT168	179	ATGCCCCGAGGCTACCTGTCCCGGCCGTTTGCGTCGAGGGCGGCCGGACA GGGCGGGGCCATTTGCAGTAACCTAACGCCTGCGCGTGCCGCGCCCGCGG CCCTGGCTGCCTCGCGAGCCGCCGCGGCACGCCCGTCATCGACGCTTCCTG CCCAGCCTCGCGCGGGCGGGAAGTTTTAT
>BbSIPHT169	54	CCCCCACGCTGCGCCCGCTGTGCGGGCTTGCTGCCCCCCAAGGGGGCTTTT TTGC
>BbSIPHT170	46	ATAAAAGCGTCGCGCCGGGGCGGCGGCGCAACCATTACGCCCGCACA
>BbSIPHT171	51	GGAAAACATCGCCGCGCTGTGCGGCTTCTAGGCCGCGCGCCGCCCTGGGC CC

Name	Length	Sequence
>BbSIPHT172	268	TCCTGGGTAGGGTCTTGCGCAATAAGCGTCGACGACGTGCGCGGCGTACC CGCGCGGGAAGTGTCAGGCCGTGACGTGGAATCCGTTGTTGATGTGGAAT TCGACTTAAATGATCGGTATTCGGTATTAATTTGTTATTTGTAGTTATATAT ACGAATATTCATACCCGGATTTGCCCTAAGTTGGTGCGTTTTCGACGGGTG CTTTCGATTGCCCGGGGCCTCCGGCACATGAAGGAATGCGCGGCAACGCC GGGCCCGCGCTGCGAT
>BbSIPHT173	114	CCGCCAACATGCGCCGAAACCTACGCCTGGCTTACAGGCTTGCAATTCCG GTGGACTTTGCCGACAATGTCACCTGATTGCCCCGGTTCCGACCCGAGCCG GGGTTTTGTTTTGG
>BbSIPHT174	38	CCGCAACGCCGGGCCGCGCCGCCTAGGCGGCCAGCCTGA
>BbSIPHT175	75	GGAAAAACCCAGTGCGACGGCGCACAAGGCCAGTGGCGTGGATATGTTCG GCCAAGCCAGGGCTCGGGTCGGCGCG
>BbSIPHT176	82	GGAAAAGCCCTGGCCGGCGGCAGTCGCCGGGGCATGGCGGCCTGTTCCGG TGTCGCTCCGATTATAACTTTCTCTCAATTCAG
>BbSIPHT177	151	TCAAAAGCAGCCCCGGCCATCGGCCGGGCGCTGCAGGCCTCAGTGCCGCC TTGAATATAACTTTTATTCAATTCAAGCGGCGCTGGATGCTGCCCTCCGGG CGCCTGGCCTGGCCGCCTGAAAAACACGATGACAAGACGGAGACAGAGTT T
>BbSIPHT178	159	GCAAAAGCCTGCCGCCGCGGCTGGCCTGCATGGCCGGCCAGGCGTGGCGA CGCCCACTGCAAACAACACTTTCGTTCAACTCCGGGGTCATGCTATACATA GATTTGGTATATGAATAACTCTATTTTTTTGTGTCGTTTATAAATATTATTT GTTGATC
>BbSIPHT179	50	GTAAAATGCGCGCCCGGCCATGCGCCCGCCGGCATCGCCGCGCCGCGCGC A
>BbSIPHT180	149	CGAAAATGCGGCGGCGCCCATGCATGGCGCGCGCAAGCAACGGCATATGC GGTTGCGAGGGGGCGTAATCTTACAAGACAGAGGCGGGTAAGATCGAAT GGCTGTATGGGCATACAGCCATTCGCGCGCCTGGGGGCGAAGCCGTGCGC G
>BbSIPHT181	71	TCCCTTCGTCTCCTGACGCCGCGTAGCGGCACAAAAAAACCGGCAAGCCG CAAGGCTTGCCGGTTTTTTTTG
>BbSIPHT182	178	TGAAAATGGCCTCCGTTACCGGAGGCCATGAGCATGCCTGCCCGCGGGCG AGCATCCCGACGGCGTGCCGGCCGGCACGGGGGCTGGCCGGCGGGGCCG CGAATCAGGACGACGGCGTCGTGGCGCCGCCGGAGCTGCCCGGCGCCGGT TCGGCGGGCGAGGGAGCCGGCGCAGGCGCG
>BbSIPHT183	154	GAAAAAGGACAGCGCCCAGGCGAGCGCTTGGGGGATAGTCAGCATGGCG GGGGCTTAGGAGAGCGAGCCGGTGAGGCTCAGGGACTTACGTACGAGAG GAAAGCCGCTTTTGCGCCGGCGGCTGGGCGGCGCCGGCTCGCTCTCCTAA GGGCTGC
>BbSIPHT184	38	TCCAGAAAACCCGTCATGCGAAAGCAGGCGGGTTTTTGC
>BbSIPHT185	79	GCCGGCCTGCGCGGCGCGCGGCGCGCGTTGCGCAGTTGTTGCACACGCAG CCTAACCCTATGGCGGGTAAGGCTTTTTTC
>BbSIPHT186	155	GAAAAAGAGCCTGTCCCGACGGGGACAGGCTCTTTTTCCCGGCGCGGATA GATACCGTCCCGGCCGTCAAGCGTTCATGGCGCATGACGGGGTGCCTGTC CCCGCGGGGACAGGCACCAGGATTCCTGCGAGGACAGGGGCCTCATGGCG GGGTAT
>BbSIPHT187	290	CAAAAAGGCCCTTGAGCAATCAAGGGCCTTTTTCTTTTGCGGCGCGGCGG CGCCTGTCCCCGCGGGGACAGGCGCCTGCAACCGAGCCACGTCACCCGGC CTTCCATCGGCCGGCCACGCCCGGTGCCCGCCCCTGCGGGGACAGGCACC CCCGATCGAGCCACACCGTTCCCGGTGCCTGTCCCCGCGGGGACAGGCAC CTGCAACCGAGCCACATCGGACGCATGGAGTTCGAGTTGCGCGCACGCTA CGAATAACAGGCCTGTCCTGACATAATTGTCAGTTTTCCGG
>BbSIPHT188	89	ATTTACGGTTGCCGTTTGCGCTTTTTGCGCTCCTGCAGCCGGCAAGCAGTA CCGAAGCGGCTTGAGAAGACATTCTCAAGCCGTTTTTTT
>BbSIPHT189	193	CGAAAACGCCAGATCCGCGCGGGACGTGGCCGATTGATGGGGTGGAGGC CATCTTAGAAATGACGCCACGGTTTGAATAGGGCGCCATTTGGCGTACCG CCCCTGCACCATCCTTCAGATGGCACTTTTGGATGAGGGCAATCTTTTGTA TGCTGTAAAAAAATCCACTGACGTATTGTGTAATGAGCAAATCT

Name	Length	Sequence
>BbSIPHT190	214	TGAAAAGCCGCCATGCCTTGTGCATGGCGGCTTTTTGCTTTGGGCGGCGCG CGCGCTTGCGGCGCTTGCCCTGCCACGGGTTTGAGGGCCGGCCGTATGCCC TTACACTTGCCGTACTTTTCGCAGCCACTTTGACTTTTCGAGGTGAAGGCG GACTGCCGTGGCGTGCCTTGCCGGCGAGTGGGGCGCGGGGCGGCGGCGCG CCGCAGGAGCCT
>BbSIPHT191	79	ACAAAAGGGCTATGCCGTGGAGCATAGCCCTGGTAATGCGTTCGGCTGGT GACCGAAGAATACGGGACGGGTTTGGAGCC
>BbSIPHT192	101	CGAAAATCAGGCCTGGCGCGCCGCTGCGCGGGCATCGCTTATCGATGCAA TGAAAGAAGATGGGGGCGGGATTTCCATCTTCAAGAGGGGACGGAGCGA CGG
>BbSIPHT193	75	GCAAAACACGGTGCAATGCGGCTACCGCCGCTGGCATGCCTGAACCGGCG ACGGCCATTATTCGCGCGAAATCTTT
>BbSIPHT194	108	TCAAAAGTTCCCGCGCGAGGCGGGAGCCGGGCCGACGTGCTCCCATGGTT TTACCAAGATCCGGCTATCGGCGGCATGACAGCGCGCAAGCAGCGGGCCG GTCGCCGAC
>BbSIPHT195	138	GCGTCTCGTCTCTCCAGTCACTGCGGCATTCTAAGTGTTCTCCGGGCCATG TCCCGTCCAAGAACAGGGCCGATTTCCCGTCTTTTGTGCGAAAGGTCTTTG CATGTCCCGGCGTTGCCGCCGCGCGACGGATTTTTTA
>BbSIPHT196	114	CGAAAACGCGGGCATGCCGCCCCCAAGGGGGCCTGGGTCTTCAGGGCTGC AAGCGCGTACAGCAAAAAGAAAGGCCCGCCTTGCGGCGGGCCCTTTGTCA TGGCCGAAGACCGGC
>BbSIPHT197	79	GCAAAAGTGCCTGCGCCGGCGCCCGCGCGGGAGGCTGGGCGTATGATGCA GGTTCGACCGCGAGCCCAAGGAGGCGCGTC
>BbSIPHT198	329	GCAAAACGGCTGGAGCCGCGCAGCGTTCCACGCCATAAGCATGAAGCCGC ACGGCGGTGCTCATGCGCCTGCGATCGTGCAAACGGCGCGGGGGCACCGT CGCGTCCGCTTGCGGCTTAGTTCGCACGGGGCAGGGCCGCTCCGCCAAAA CTCATGCAAACGCCCCGATCGGCCTTTTGCCGGATTCCGGCAGGCTGGGC GCGCGGCAATCATTGCCTCACCGGCGCTTGAGTTCTGTTGCACAAAACAA AGGGGAGCCAGGCGGCGGTCCGCCTGCCGGTCGGAAATCGGCGAGCGTGT TCCGTCGCTTTACGGAATCGGAGGAACCCT
>BbSIPHT199	88	ACAAAACAAAGGGGAGCCAGGCGGCGGTCCGCCTGCCGGTCGGAAATCG GCGAGCGTGTTCCGTCGCTTTACGGAATCGGAGGAACCCT
>BbSIPHT200	123	GGAAACTCGCAAACGATGAAAAAAACGAATCGATTGGATGATTTTAAGTC ATTACGAGATTTGCCGCGGGCTGCCTATACTGCGGAAAAATTTGGTCGCA CCGGCCATCCGGTGGACGTTTTGA
>BbSIPHT201	85	GCCGTACTGACTGGAAGGCGCGCCCCGCGAGGGGCGTGGCCCGACAGTCA AAAAAAGGCCGCTGATGCATATCAGCGGCCTTTTGC
>BbSIPHT202	128	GGGGAGGGGACAGGCACGGCAGGCGCGTCACCGCTCCTTGCCGGCTCATA TAAATGAAAATGCGACGCATTATTTATTAATTCATCCCGGCCCGCAATCGC GGCGCACAGGGCGCCATATCCCTTTTGC
>BbSIPHT203	178	GCAGGGCTGTATCGAGGCGTTAGCAAGCGTGGCCGATGGCAGCGCCGGGC CCGGGCAAGGGCGGCCGCCCGGCAACATCTCGAGCGAGTGTCGTGCCGGG CGGCGCCGCCGCGCGGCTTATTGCAGAACGATGTTGGCCTGCTTGATGAG CTGGCTCACCACGGGCTGCTCGTTTTTCA
>BbSIPHT204	99	AAAAAAGCCGGCCTGGTGCCGGCTTTTTTTTCGAGGGCGGCGCGGGCCGC GGGAGGCGCGCTACCATCAGCGCCATTGTTTCTGCTCAGGGGATGCCGGT
>BbSIPHT205	91	CGAGGGTTCGCGGTCCAGGCTCGGCGAGGCTCTTCCCGCCGCCGCCGAAC CGTGAAAAAAAGCCCGCCGGCGCGATGCTGGCGGGCTTTTCT
>BbSIPHT206	31	AAAATCGCGTGGTGACGCGCGGTTTTTTTTGT
>BbSIPHT207	43	GAAAAAGCCGCCAGCATAGCTGGCGGCTTTTTTCTTGCCGTCAG
>BbSIPHT208	103	ATTTTTGGTTTCTGAAGCCGCTGGCGGAGGAAGGGACGCGTGTTAACTGA AGTTTGGTAAAGAATCCCCGAAAAAGCCGCCAGCATAGCTGGCGGCTTTT TTCT
>BbSIPHT209	43	GTTTGGGGCAGGCATGAAAAAAGGCGGCCGCGGCCGCCTTTTCC
>BbSIPHT210	124	GGAAAAGCGCCACGCGCCGCGAGGCGCGCTCCCGCGCCGTCCAGCCGCGC GGCCGGCGAAAAAAAGCCCCGCTGGCGTGGTCGTCCAACGGGGCGGCTGC ATGGCGGGCGCAAGCCCGCCATGCA
>BbSIPHT211	65	AAAAAAGCCCCGCTGGCGTGGTCGTCCAACGGGGCGGCTGCATGGCGGGC GCAAGCCCGCCATGCA

Name	Length	Sequence
>BbSIPHT212	79	GGCCCATGGCATCCTGGCGCCGGTCGGCGGCCAGGGTGGCCGGCAATGAA AAAGGCGGCTCATGCGAGCCGCCTTTTTCA
>BbSIPHT213	91	AGCAAGCGCAAGGCACGTAGCGGCCGTCCACGCACGCTTGTCCTTCATCG CGCACAAAACAAAATGCCCGCAAGCGATTGCGGGCATTTTGT
>BbSIPHT214	141	AAAAAAGCGCGGGGCGGCGATGCGCCGCAAACGCCGTTCCCGGGCTGCC GCGAGAGGGGGGCGCCGGATTCGCATTCGGTTGCCAATTGACACGCCCGC GCCGGCGTGCGCCGGCCTGTTCCGGGCGGCGGAAATACTGCTT
>BbSIPHT215	124	ATTGCCCGGTGGCCGCCGTTGTTATACGGCCACCACTGCACATAAGAAAA ATTCTATTTATTTATTCGCCATATTCGTTTGATTTATGTGGCGCGCGTGGCG CGCCGCCGCAGGGGCTGTTTTGC
>BbSIPHT216	67	AAAAAACCCCGGCGCTTGCGCGACCGGGGTTTTTTGGCACATCACGACAG GTATCGTGAAGTATGTAT
>BbSIPHT217	151	GAAAAAGCCGGCGCGCGCACGCCATTGAACCGGCCAGGCGGCCCCGTCCG GGGCCGCAGGCGGTAACATTGCACCGTAGACCGGGCCGCAGGGCCCGGCC CGCCGCCGTTTTCCGGCGGCGTTCCACGTTTCAATCACGAAGGTATTCATT C
>BbSIPHT218	70	ACAAAACCGGGTCCGCTATACTGCGCCCCGCGATTTACTTGAGTGGCCCC GAGGCAGCAATGCGGTGCAAG
>BbSIPHT219	179	GAAAAAGGGGGGCGGAACCGGAGGGCCGGCGCGCGCCGGCGGCGTGCGC CGGCCGGAAAGCCGGTCGTGGGCCGACGCCTGCCCGAGGGCGGGCGCGG GCCAAAAGATCAATCATATACAATTCGCGGCGCGGGTCGCATGTTTGAGA CGCGGCCGATCCCGGCGTCCGCGCGCGAAGCC
>BbSIPHT220	61	CAAAAAGGATGCCCGCGGGCATCCTTTCTGCTTGCCGTATTCGCTGCCCCG AAGGGCAATAC
>BbSIPHT221	76	CAAAAAGGAGCCGTAAGGCTCCTTTTTTGCGTTCCGGCGCCTGCCGCGGG GGCGGGCGCGGCCCGACGGCGCCGGGT
>BbSIPHT222	92	ATAAAAGGCGGGGCATGCAGCCCCGCCTCAAATAATAACGACGCCCGGCG CCAAATGGTATCGGCGCAGTGGCGACGCGATAAGCAATGGTGC
>BbSIPHT223	138	CTTTTTCGTATACGGGTTAACCTTAATACCGTACGCGTTGCAACCTCTTTA GGATACGGTTCCACCGCCAGTGCCCGCCGTGATCGCGGGCACTGGATCGC AGTCCTGATGGCAGGCATCACGCCGGCCACCGTATCTC
>BbSIPHT224	157	GCCGTGCCCCTCGGGAAGAATGGCCTGCATGGGTGGCGGCGTTGCCGCTA TCCATGCAGGCCTGGAACGGTGTCGCGGCAATCGACCCCGTTCCGTTGGG CCACGGGGCGTAGCTGACGGCTGCTTCTCGGGATCGCCGGCTGGCGATTC AGTTTTTC
>BbSIPHT225	204	GCCGTGCCCCTCGGGAAGAATGGCCTGCATGGGTGGCGGCGTTGCCGCTA TCCATGCAGGCCTGGAACGGTGTCGCGGCAATCGACCCCGTTCCGTTGGG CCACGGGGCGTAGCTGACGGCTGCTTCTCGGGATCGCCGGCTGGCGATTC AGTTTTTCTCCTGTGCTTTGAAAGGCGGCGCCGCTTGCGGGCCGCTTTTTTT TCG
>BbSIPHT226	69	CAAAGCCCGGCCGCTCAGGTACGGCAGGCCGCGACGCCGCCTGGTTCTTC GCGAACCGGCGGCGTTTTCC
>BbSIPHT227	82	CCGCGCCGGGCTGCCCGTGGGCAGTCCGGCACGCTTCCTCGGCCCGTGCG CGCGTGGACAGGCGCGCGGGCGCCTGGTTTTGC
>BbSIPHT228	103	AGCTACCTCGTCGCGGTCGCAAAGGCTTGATCATACGCTGACGCAGACAT TGCCGGCACATGGGTGGATCGCGGCGCCCGCAATCCGGCATCGCGGCGTT TTCC
>BbSIPHT229	80	CAAAAAGGCTTGGCGATTCATCGCCAGGCCTTTTTTGTTTTCGCAGTCGCA GCTTTGCGCGCTTGGCGCGGCGGGGCTTCT
>BbSIPHT230	68	GCGAGGCGGCGTAGCCAGCCGCACCGTGCGCAGCCGCCTGCCGCGATAGG GCGCCAAGGCGGCTTTTTC
>BbSIPHT231	211	AAAAAAGGCCCGACACATCGTGCCGGGCCCGTACGGGGCACGCAATGCC GCGTCAATAAGGCTCGCGGGGAACACGGCATTGGTGTCAGAGTACTGGCC CTGCAATCCGCGGACAGGGGATCGTTTGTCCATGGATTGGCGCAAGGCCG TGATCTGAGCGATCAGTCTAATCCTGGAGGTCTGACAGCGTCCTGAATCG GACGGCGTCTCCA
>BbSIPHT232	55	GCCCCCACGCTGCGCCCGCTATGCGGGCTTGCTGCCCCCCGAGGGGGCTTT TTTGC
>BbSIPHT233	49	AAAAAAAGCTCCCTGTGGGAGCTTGAAAAAGCCGGATTGGCGCAGCGGTC
>BbSIPHT234	68	GCAAAGCTGAACTACAATAGCGGGCTGTTCGGGCCCCGCCTTCAGGCGGG GCTGACCAGCCCGTTTTCT

Name	Length	Sequence
>BbSIPHT235	55	CAAAAAGGCCCCCTTGGGGGGCAGCAAGCGCGCGCAGCGGGCGCAGCGT GGGGGCG
>BbSIPHT236	54	CAAAAAGGCCCCCTTGGGGGGCAGCAAGCGCGCGCAGCGGGCGCAGCGT GGGGGC
>BbSIPHT237	231	GTTCTCGGCAATGGTTTTGTCAGAAAGCTGAATTTTAGCTGAACGTTTCCG AGTGTTCGCCAGATTTTGCCTGCAAGATGCGTACGAATATGCGGAATCAG GGCTGCGGCGTCCAGGGGCGCGGGGGCGTTGCCTAAGCTGTCCATGGTCG CCGCCCGACAGGCCGGGCGGCCGGGCACGGAGATCATGCATGGCAGGCA CATTGGCGCCCATCCGGGCGCTTTTCTTTTGG
>BbSIPHT238	120	CGCCTGACCGCATGGCCGCCTCGCGAGGCGGCGTTTGCCCGGCCGGTTGA GCGGCCGGGCGCGTGCGCGCCGCGCGGACAGCACGAAATCCAAAGGCCA CGAGTATCTCGTGGCCTTTTGC
>BbSIPHT239	52	AAAAAAAGCGGAACCCTCTATCAGGTTCACCGCGTGCGGCCGTTTTACGC AAA
>BbSIPHT240	115	TTAAAAGCCGCGTCAGTCTTTGACTTCGCGGCTTTTTGCTTTTTGCTTTTGG CGTTGCCAAGCCAAATAGCACCAGGATCCGTTTTCGGCGCTGGGCCAGAT ACTCGGCATGGTTG
>BbSIPHT241	44	ACAAAACGCGGCCAGGCAGCCGCGTTTTTCGTTTCCAAATCCAGG
>BbSIPHT242	148	AGATGCTGCTGCGCCCGACACGAACTACGACCGGAATCTTCCGCACTCAG TCGCGTTAATGATTTCATGCGTCGGGTTGGTATCGCGTCATTGCGGTAAAT ACTCAACAAGGCACCAGACAAAACGCGGCCAGGCAGCCGCGTTTTTCG
>BbSIPHT243	148	GCAAAACGGCGCCTTGCGGCGCCGTTTTTTTCGTTGCGTTGCACAACGGGC AGGACTAAGATGAATCGCATGCGGTTTCTCAGCGCGCTTTTCTCAAGGAG AGAATCATGGGCGACGCGACACGCAGCAACGGTGAGCCGGCCACCTCT
>BbSIPHT244	51	CGGTTTCCCTGGGCGCGGCAATGCGGGCAGGCCCAGGGCCTGCCTTTTTTG T
>BbSIPHT245	61	AAAAAAGCCCGCTCGGGGCGGGCAGGGCAATAAAAAAGCCACCCGGAGG TGGCTTGCTTTTT
>BbSIPHT246	175	CTTTGCCCTCCCGCCGGGTGGCGGGCGCTGCCTCGGTGAGGCGTTGGGAG AAATATAGCAACGCTACAGATTCAATGCAATAGCATCGCTACAGTATTTTT GTAACTGACGCAAAAAAGCCCGCTCGGGGCGGGCAGGGCAATAAAAAAG CCACCCGGAGGTGGCTTGCTTTTTAG
>BbSIPHT247	75	ACGTGACCCGATGCCCATGGCGGGCGCTGTGCGTCCGTCAGAAACGAAGC CCCGAGCCGATGGCCGGGGCTTTTGC
>BbSIPHT248	81	CTGAGCCATTCCCGTCCGGCCCGCCACTGCGCGGGCCTTTTTTCTTTCTTAG CCCTGCCGACTGTGTCGGCGGGGCTTTTGT
>BbSIPHT249	86	TAAAAAGAACTGCCGCCATCCCTTGTGGGGATGGCGGCAGTTCGACTGCG CACGCGAACGCCGGCCGGGGGGTGGGCCGGCCGATGC
>BbSIPHT250	66	ATAAAAGCCCCCGCCAGGGGGCTTTTTCATTGCTTGGCCCAAGCGGCCAA TTCCGCAGGAGCACGAC
>BbSIPHT251	83	GGCCCCGTTGTCGCCGCGGCCGGCGCAGTTCCGGCGGGCCGCATGAAAAA AGCCCCATTGCCTCGCGGCAGTGGGGCTTTTTGT
>BbSIPHT252	96	GCAAAACCGGGCCCGGATCATCCGGGCCCGGTTTTTTTTCGGCTATCCGCG GGTTCGGCGGTTCCAGCCCGGCAATGCGTGGCCGGCGGGGGCGGGT
>BbSIPHT253	67	CCGGCTCAGGCCGGTGGCCAGCGACCCGCAAAACCGGGCCCGGATCATCC GGGCCCGGTTTTTTTTCG
>BbSIPHT254	53	AAAAAAGCCCCCTCGGGGGGCAGCAAGCCCGCATAGCGGGCGCAGCGTG GGGGC
>BbSIPHT255	53	AAAAAAGCCCCCTCGGGGGGCAGCAAGCCCGCATAGCGGGCGCAGCGTG GGGGC
>BbSIPHT256	53	AAAAAAGCCCCCTCGGGGGGCAGCAAGCCCGCATAGCAGGCGCAGCGTG GGGGC
>BbSIPHT257	133	AGAAAACCGCGCGACTGCGCGGTTTTTTCCGACCGAGATACTTAGCGTTTT ACTCGGGAATTTGCTATACTTGACCCAATTACTCGGGTATTGATGCCCGGT CCAGCGCCCGATGCGTCGGGTTTTCTTCCGGA
>BbSIPHT258	30	GAGAAAACCGCGCGACTGCGCGGTTTTTTCC
>BbSIPHT259	100	CCGCCGGCCGGCGCGCCGCCCGGACAAGGGCCCCGTCATGGACAGGGCCC TTGTCGTTTCGGGGCCGCGCGGCCGGGGTATCATCGGTCCTGGCATTTTGC
>BbSIPHT260	99	CGAGGCCGCCGGGCCGCGTCCGCGCACCGCGTCGGGGGACGGTCGAGCCC GTCCGCCCGTTGTTGGCCGCCTGTCGTGCATCGACTGGCGGTTTTTTTTT

Name	Length	Sequence
>BbSIPHT261	54	AAAAAAGCCCCCTTGGGGGGCAGCAAGCCCGCATAGCGGGCGCAGCGTG GGGGCC
>BbSIPHT262	145	AAAAAAGCCCCCTTGGGGGGGCAGCAAGCCCGCACAGCGGGCGCAGCGT GGGGGCCTTTTTTTGCCGCCGGGCCGCCCCAAGGCAAAAAAGCCCCCTTG GGGGGCAGCAAGCCCGCACAGCGGGCGCAGCGTGGGGGCCTTTTTTT
>BbSIPHT263	167	TTGCCGCCGGGCCGCCCCAAGGCAAAAAAGCCCCCTTGGGGGGCAGCAAG CCCGCATAGCGGGCGCAGCGTGGGGGCCTTTTTTTGCCGCCGGGCCGCCC CAAGGCAAAAAAGCCCCCTTGGGGGGGCAGCAAGCCCGCACAGCGGGCG CAGCGTGGGGGCCTTTTTT
>BbSIPHT264	61	AAAAAAGCCCCCTTGGGGGGCAGCAAGCCCGCACAGCGGGCGCAGCGTG GGGGCCTTTTTTT
>BbSIPHT265	78	GCCGGGCCGCCCCAAGGCAAAAAAGCCCCCTTGGGGGGCAGCAAGCCCG CACAGCGGGCGCAGCGTGGGGGCCTTTTTT
>BbSIPHT266	61	AAAAAAGCCCCCTTGGGGGGCAGCAAGCCCGCACAGCGGGCGCAGCGTG GGGGCCTTTTTTT
>BbSIPHT267	78	GCCGGGCCGCCCCAAGGCAAAAAAGCCCCCTTGGGGGGCAGCAAGCCCG CACAGCGGGCGCAGCGTGGGGGCCTTTTTT
>BbSIPHT268	61	AAAAAAGCCCCCTTGGGGGGCAGCAAGCCCGCACAGCGGGCGCAGCGTG GGGGCCTTTTTTT
>BbSIPHT269	78	GCCGGGCCGCCCCAAGGCAAAAAAGCCCCCTTGGGGGGCAGCAAGCCCG CACAGCGGGCGCAGCGTGGGGGCCTTTTTT
>BbSIPHT270	54	AAAAAAGCCCCCTTGGGGGGCAGCAAGCTCGCACAGCGGGCGCAGCGTG GGGGCC
>BbSIPHT271	61	AAAAAAGCCCCCTTGGGGGGGCAGCAAGCCCGCACAGCGGGCGCAGCGT GGGGGCCTTTTTT
>BbSIPHT272	45	AAAAAAGCCCCCTTTGGGGGCAGCAAGCCCGCACAGCGGGCGCAGC
>BbSIPHT273	105	GGAAAAGAAACGGACGCCAGCTTTTGGCCGGCGTCCGTTGGGTTGGAGGG TTGTGCGCCCCTGCCTTGCGTGGCGGCTTGCGCCGCCGTTCCCCTTGGGCT GCTGC
>BbSIPHT274	107	GATGCCGCGGCAACGCGGCACCGTACGGCGCGCTGTGCCATGCGAAATCC GCGCACGGCCGCGCCGGGTCTTGCGGTGCGCCGGCCATTCCGGTCGGCGG CCTTTTCC
>BbSIPHT275	72	CGGGGCCAGTCATCATCACGCAATGCATCGCGCACCGGCAGAGCCCGCCA GGTTCTGGCCGGTGCGCTTTTAT
>BbSIPHT276	191	GCCGCTGCCTGCCTCCGTCGTTGGAATTCGTTATGAGGGCAGTCTAGGCAA ACCGGCGCCGACACTCCATAAAGCAAGGTGAATAAGTAATATGCCGCATT GCTTATGTTATGCCGGGACGGCGGTCGATTCTCCGCGCCAATGCGGGTTTT CATGGGGCGCGCAGCATGCCTGGCGCGCGGAGGATTTTCG
>BbSIPHT277	70	GAAAAACCACGCCCAGGCCCGCGCCGCCCAGGCCGCCGCAAGGCGGGCG CAAGGCCTCGCGCGGCCAGTCC
>BbSIPHT278	70	CGCGCCGCGACGCGATGTCGCCGCCGGCAACGCCGGCCTGTCGCAAGATG CGACAGGCCGCTGTTTTTTAT
>BbSIPHT279	68	CCCGCCATGAAGGAGGCACCGTGTTGATTCGGTGCCTGTCCCGACAGGGA CAGGCACCTGTGCTTTCGA
>BbSIPHT280	47	GTGTCATCCTCCGGAAGGTGGCCACGTCTCGTGGGGCGGGCCTTTTGA
>BbSIPHT281	104	GGAAAACCTCGATGGGCGCGTCTGCGCCACGGTCGAAAACCGTCGATTTT ACCGTGGCGCGGCGCGGCCGCGGCCGCGGCATCCCGGGGCGGGCCGCGCC GATGC
>BbSIPHT282	127	GGAAAACGCCTAGCCGTGGGCCTGGACCGGCCAGGCCTTGCTCGGGAAGG GCGTCAGGCGCGAGGCGCGCGCGTTGGAGAGGGCAGGAAGCGGGCGCGC GCAGGCGTGCGCGCGCCTGGCCGGGAGGC
>BbSIPHT283	43	AAACAAAAAAGCCCTGGACCACGTGGCGTGTCCAAGGCTTTTCC
>BbSIPHT284	61	ATGAAGTCCGCAACTATACACGGATTCCGGCCGTTTTGGTAGCGGCCGGC GGAAAATTTTCG
>BbSIPHT285	104	CACCCCGACGCCGCGGCCGGCTGGCCGCGGTCCGGGCCATCCGTCCGCCA GGGCGGGAAAGCGGCGATCCGCCGCGGCCAGCGCCGCGCCGGATCGTCTT TTTAC

Name	Length	Sequence
>BbSIPHT286	154	TTTGGGTGTGGGAGCGTGGATTATGGATGGGCCGCGCGCAATCCCAATGC CCGAAAACAACCGTCTTTTCGGCGCTATTTGATCCGTATCGTTCCTGCGTG CGCGGCACGCGGCCGCCGGGCGCCATCCAAGCCGGGAAGTGGCGGCTTTT TTGG
>BbSIPHT287	46	CGTCCCTGGTTTGGGGCGCGGCGATCGGTCGGCTCGGTTTGTTTTTT
>BbSIPHT288	40	ATAAAATCCCCCGCCATTTGCACGACCAAGGATACCCAAGG
>BbSIPHT289	216	CCGGGTGCCTGTCCCCGCGGGGACAGGCACCCGCCAACTCCTGTGCCTGT CCCCGCATGGAGACAGGCACCCGCCAACTCCTGTGCCTGTCCCCGCATGG GGACTGGCGCCCGCCAACGTCTCGGTGCCTGTCCCCACAGGGAACAGGCA CCCGGGCCTCCCTGTTCCTGCTCCTCCACCTGATCACGTGCCGCCCACCGG GGCGGCATTTTTTTGC
>BbSIPHT290	101	GGGCGCCCAGTGTTTCTGTTTTCACCACCGAGGAGTATCCCCGCATGACAC AACACGTTATGGCGCGCGGCGCCGAAGGAGGCGCCCGTGCTTGAGTTTTT C
>BbSIPHT291	69	ACGGCGCGGCGAACGCCTTGGCATGCGTTGCTAGACTTTTCCGTGGGCGG CGATCCGCCGCCGGTTTTCC
>BbSIPHT292	121	AGAAAACGACGCCAAGGGTTTTCCCCGGCGAAGAAAAAATTGACAAGCC GAATGGCCTGTACGATAGTTATCAAATGATAAACAACGGATCGTGCGGGA TCCAATCTGACCAGGAGACATCA
>BbSIPHT293	305	GGAAAACCCTTGACCCGGGTTTCCCCTGACCCCCCGAGGGCCGCGCTACG CGGCCCTTTGCGTACCTCGCTGCCCGCTCATCCGCGCCAATACTTGAGTGT GAGTGGTTGCTCACATGATTGTCACAGGTAAAAAATCTCCGATTCATGATG TTGCGCATAAATGTCACGAATTTCAATCTTCTTGCCACAAAGTTGCGGCTA TCATGCCTTCATAAGACGCCGTCGAGTGCTCCCCGGGCAGTTTCCGGGCTT CAAAGCCGAATGGCGCGTTCCCTCAATCGAAAGAACACGCAAGGAGTCGG AG
>BbSIPHT294	91	AGAAAATGCCGGCGCCGCCGAGGTGGCGAACGCTCGCCGGCCGCAGTGC GCAAGTCGCGTGCCGCCCCGCGGCAGGGCCGTTGCGGTCCCTG
>BbSIPHT295	231	ACAAAACCGCCCAAAATGTCGGACGGCAGTCGGAACGGGAGAATGATAA TTTCGGGGCACCGGCCGTTGATTTCAACATATATCAAATCAACACCATCCA TGCGGCGATCAATCGCTTCACAAAAATGGTGTATTGGCGGCCTATAATCC AGGTGCATCACAACCGGGATTTTCCCGGCAAACCAATGCCGCGCATCCGT CGACGGCCGGGCGGCACACTTGCAAGGCCGCC
>BbSIPHT296	109	CCATGCCAGCCCCGAGCCCCGCCATGCGTCGCGCCGGACCGGCCCAAGGC AAAAGCCCGTCCAGCACGACGGCGCCCGCGCGGCGGGCACGGCGCGGGG CCTCTTTTTCC
>BbSIPHT297	121	TCCTGGCTCGGCGCGTTCGTCCCCCATGCGGGGGCAGGCGTGAAGTTTGC GGGTGCCAGTCCCCATGCGGGGACAGGCACCCGGCAAGAAAAAAGCCGG CATCTCTCGATGCCGGCTTTTCT
>BbSIPHT298	113	CCGCGCTACACTCACGCCCGCCACGCCATTGCGGCGGCGCGGGCGTTGCC TGGCGCACGGCATGAACTGCCTTGTTTGTCGGCAAAGCGGCCATCCGGCC GCTTTGCATTTTTC
>BbSIPHT299	52	GGCGGGACGGATTCCGTAGTGACACGACAGCTTGTCTGTTGCGTCGATTAT GC
>BbSIPHT300	59	AAAAAAGGCCGGCCGCTCCCGGCGCGCCGGCCGATTTGCTATTCTTCGGTT CTGCCCCTT
>BbSIPHT301	98	GCCGGAGCGCGCCGCCTTCGGGGTTTTCCCCCAGCGCGCCCAGGCTTGCC GCCACGCCGGCGGCGGCGCCGCGCGGGCCGCGCGCGCGATAACTTTTTC
>BbSIPHT302	90	CTAAAACAATTCGGCCGATGGGCCGACTTACCTAGGAATTACCCCTAAAG CCCGCGCGGGACCGTGTAGCGGCGCGCACGGGGACTTAGAA
>BbSIPHT303	124	CAAAAAGCCCATCCGAAAGGATGGGCTGCAATAGGGTGGAAAGGCGGCG CCGTTGGCCGGCATTGACCGGCGCATGAGAACCCTGCGAATCCCCATGCG CCGGGCCGACGGCGGGCAGCGACTAC
>BbSIPHT304	114	CTTCCGTTGGTCGAATTTCCCGGGTGGCGCCATGTGCGTTGCCAGCCCACA AGCCGTACAGCGAAGGGGCGCATAGCGAAAGCCGATACCGCAAACGTGG TGTCGGCTTTTTTGC
>BbSIPHT305	182	TAAAAAAAGGCGAACCCGATTCGCCTTACTCCGTGCCGGAAACTGTACCG GATCAGCCGCCGGAGTTTACCACCAACAGGCCGGCGAACCACCCCCGCCG TGCACGGAAGAAGGCGGCGTTCGGGGTATGCTTGCAGGCTTGAGCGCGGC TGTGCGCAGCCCCATTTCCCGAAACCGAGCATC

Name	Length	Sequence
>BbSIPHT306	53	GTTTTCTCCTGGATGAGGGCCGGCCTTGCGCCGGCGGACAGCGTCCGGTTT TCG
>BbSIPHT307	215	GTTTTCTCCTGGATGAGGGCCGGCCTTGCGCCGGCGGACAGCGTCCGGTTT TCGAACGTTGGAGAGAGCTTAAACCTTCCCATGATGTGAAGGTCAAGCGT AGAATGGAAGCGACGCCTTCCATGCCGGAAGGCGCGCAACCGGCCCGCCG TCGGCGGGCCAGGGACACTGCCCATGAATATCGGAGACGCCGCCCAGGCC TCGGGCATTTCGGCC
>BbSIPHT308	117	GAATCTCCCACACTCTTTTACATACGTGAATTCCTGTAGTCTAGGAGGCCC CAGGCAGTGATAATCACTAAAATCATTCCGAAACGCCGGCCTTTCCACTG GCCGCCGCGGGTTTTCT
>BbSIPHT309	270	GCACGCTCTCCCCGGCCGGATTGCCCGGCCGCACCGCTCGATGACTCGAC GCCATGTGCCCCTCCCTGGATGGCGGCCGGCCCTGGCCGGGCTGGTCCGC GCCTTGAATGTCCATTATTTAGACGATATGTCTATTTATTGGACAAATTAT AGGGCGGCAAAAATGACACGTCAACGAATGGCGAGTGAAAAAGAAAGCG CGCGCAGCGCACGCGCCCCTTGCGGGCGCGTGTCGCCACGTCGGGGCGTC AGCCCCCGCGGCCGTTGTCCA
>BbSIPHT310	36	ACGACTCCAGGCGGGCCGCCAGGCCCTGGAATTTTCC
>BbSIPHT311	93	TGAAAAGCGTCGCCGCCTGCCGCGGCCGGAAAAAGCGATATTATGGCCCG CCTCTTACCTTCACCACAGATCTGTCACGCAACAGCTACCCATC
>BbSIPHT312	99	CCGCCGGCGCATCATGGTTGCGACGGAATGGCTTTTCTTACATGTTTCCAG GATATGTCCGTATTTCGGGCGATGCCTCGGTCGCGGCGCCTGCTTTTGT
>BbSIPHT313	301	GGGCCGCACCAGTCTCGCCTGGCTCGCGGCGACAGGCCAGCAGCTCTCGG CCTGACCGTACGTAGTCGCCTTCCGCGATCAACGCGTCGCGGCAGATATCC TGCAAATGAAAGAGAAATGACACTCGATGATGTTCAAATTCATACAGGAA ATTCACAAAAGGACTGATCACGACATCCCTGCCGAACGTCTACACTGCGC TAGTGCTCGTCTGGCTGACCTTGCCTCGCCCGTTTGCCAGGCTGGCACATT CCTACAGCGAAGTTGATATCCCGGCCGAGGGGCGCCGGGGTTTTTTTTAT
>BbSIPHT314	77	GCAAAACGCGGGCGGGCATGAGTGCCCTGCCCGCGTGTGGAGGCAAGCG CAGATCCGGCGCGCGAGGCGCCGGCAGGC
>BbSIPHT315	48	TGAAAACCATACGCGGCCGGATGGCGTTCCGGCCCGGGAGGAGCACCTG
>BbSIPHT316	76	AGAAAACAGCCGCCATTGCTGGCGGCTGCGGAAGGACTTGCACCTGCGGC AGAATCCAGGACACCACGCGAGCCTGA
>BbSIPHT317	124	GCAAAAGGCCATGCGCCGGCCCGGCACATGAATGCCGCCACCCGCCATGA CCCTGCGCACCGCGCGTCAGCCGGTACTCGTCCATCCCTGCACACCCGCGG CAAACGCCAAGCGCCCGCTCGCCA
>BbSIPHT318	75	GAAAAACGGCTGAGAGGGCGTAAGCCCGCTCAGCCGTATGCGGAATGTTT TCTAGCTTACGCGGTATTGTACATGG
>BbSIPHT319	44	CTCGTGAAAAATGGCCGCATTCCCCCGGGAGTGCGGCCATTTTGC
>BbSIPHT320	67	CCTTGGGAAACCGGGGCGCCGCGCAGGCGGCCTGCCCCAACGGCGCGATT GTGCGCGCCGCGTTTTAA
>BbSIPHT321	81	GGTAATCCCTTGAATTGGTTGTGAAAAACGGGAAATTCTAGCACTTAAAG TTGCCTGGTGGCAACTGGAGTTGCGTTTTTCC
>BbSIPHT322	223	GGAAAAAGGCCGCTGATCGCCGCGCGCCGGTCCGCGTAGGGGCAGATGCC GCCCCCCCGGCAGGCGGGCGTCGATGGCATACGGCAAGTGTGCCTGTCCC GGCGGGGACAGGCACCCAGCCGCAGATGTCAGTCCCCATGCCGCTGGCGC ATGACGGCGTGCCTGTCCCGGCGGGGACAGGCACCCTGGCCGCAGGTGTC GCTCCCCCTGCGGGCACAGGCACC
>BbSIPHT323	65	AAAAAACCCCGCGGGGGTTCGCGGGGCTCAAGCCAAACGCAGGCCGGCC CGAAGGCCGCGCCGCGC
>BbSIPHT324	70	TGAAAAGGGTTGTGGCCGCCACGGCCACAACCCTTGCGCTGGCACCGCCG CCTGAAGGCGGCCGGCACGTT
>BbSIPHT325	69	AAAAAATGGGCTGCACGCCCAGCCCACCGTATTGCCTGGCCCCCGGCCGG GCGCTGCTGACAGCGCCGCT
>BbSIPHT326	324	CAAAAATCGGCCGCGCTTGCCCGCGGCCCGCTGGTGGCCGCGTGCTCATT CCAGGTTCCCCTGGCGCGCCGCCCACCCGGCTGGCCTGCCCTGCAAAAAG GCAAACCCGGCGCATTTGCGACTCACGCGCCAAATGGCGTCAATCACAGC AATCACGCGCCAAAGCGTGCACAACGTGCATTCAAAAGCCGCTGCGCGCA ACGCCGTGCTGCACGCGCACTGCCCTGCAAATGCACTGCGCGCCATGAAA ACCGTTGATTTTACCAGATAAACCGGAAATTTGCCTGATACGTCGGGAGCT TACGGATAAGCGGGACACATCCGA

Name	Length	Sequence
>BbSIPHT327	136	TCAAAAGCCGCTGCGCGCAACGCCGTGCTGCACGCGCACTGCCCTGCAAA TGCACTGCGCGCCATGAAAACCGTTGATTTTACCAGATAAACCGGAAATT TGCCTGATACGTCGGGAGCTTACGGATAAGCGGGACA
>BbSIPHT328	78	CAAAAAGCCGCTCACTTGCGTGAACGGCCCTTTTGTGTCGACGGTGTCGAT ATCGCCGACGATACGGCAAAAATTCTGT
>BbSIPHT329	72	AAAAAACCCCTGGTCCATTGGATCCAGGGGTTGAGAGGCAGGGCGCCGCG CGTGGCGGCGCCTGTCGGCGGCG
>BbSIPHT330	32	CCAAAAGAAGACCGCCTCGACAGGCGGTCGCGC
>BbSIPHT331	158	GCGCCTCCACAGGTCGACGGCGCGATGCGCGCCCGTTACGCCATGATAGG CAATCGGCGCGGCGCATGGAGCGCGCGGCAATCGCTATACTGGGTTCATA TCAGGAACCCGGCAATCTTGAACCGCCGCACCGTCCCCCTGGCGGGCGCG GTATTTTCA
>BbSIPHT332	98	ACAAAACGGCAAGCGGAGCACATGCCCGCCTGGCCACAGTAGACTTCCGG CCGCCATTTGTCCAATATATGAAACCTGCACCAATGATATTTTAAAAAA
>BbSIPHT333	84	CGTCTTGGCAGGCCAGCGGGACGGGGTGCCTGTCCCGGCGGGGACAGGCA CCCCGTCGCTTTGCCCCCCGAGGGGGCTTTTTTGC
>BbSIPHT334	53	CCCCACGCTGCGCCCGCTGTGCGGGCTTGCTGCCCCCCGAGGGGGCTTTTT TGC
>BbSIPHT335	42	AAAAAAGGGGGCGTCCGAGGACGCCCCAAATCCACCCTGCAAC
>BbSIPHT336	147	TTAAAACGTCGGGGTTTCCCCCCTTTGGCCGATGGACATGCCGGCGTCAAA GGCGCATGGGGCGGTCGTTTGCGAACAGCGGCGGCATCTGCGCCGGGCCG ATGTCCATGCACCCGCCTGGCGGGGCGGACACGACGCCGGACGGATT
>BbSIPHT337	126	TGAAAAGGGCCCGCCGGGGCTCGCGCACGGGCAAGCCCGCGCCATGCGAT CCCGCGGCCAGGCCGCCTGCGATACGGAACCAGCGATGAAAGAAGAGCC GCGCCGGCCGGCGCCGCCCCCCTTTCGT
>BbSIPHT338	77	GGAAAACGGCCACCCACTGGCGACCGTCTTCCATCATATCGCCGTCGGGG CGTGCGGCGAAACCTGCGGGCGGCCGGC
>BbSIPHT339	58	GCAAAACCAGTGTACGCGCCGTGTTGCGGCGCAGCACGGCGTCGAATTAT TTGGTTTCA
>BbSIPHT340	316	CGGCACGGTCAGAAGCAGGGGATGTCTGCAGGTGCCCGTCCCCCTGCGGG GACAGGCACCCGGCGAGGTGCGGCGGATGCGGGCGCCTGGGCACCTGCTT GTCACTACGATAGGCGGAGGTGCTGCCGAGGCAGGACTCCACGGCGGGTG TCTGTCCCCGCACGGGGACTGACACGCGCTGCAGGCAGCACGGTCAGAAG CAGGGGATGTCTGCAGGTGCCCGTCCCCCTGCGGGGACAGGCACCCGGTG AGGTGCGGGCGCCTGGGCGTCCCCCCAAAAACAAAACCCCGCGAGCCGTG AGGCTGCGGGGTTTTGA
>BbSIPHT341	69	CGGGCCTCTTTTCAATACATACGTATGGATTTATATTATAGGCACGCCCGT TCCGGCGTGACTTTTTTCA
>BbSIPHT342	137	TAAAAAGACCGACGCCGCGACCGGCGCCGCCCCATGCAGAGACCAGCAG TCCCGGCGCGCCATCGGCGCGCGCATCGTTCGAGTCCCCGCCAGCCCGCC CCGCGCCGCGAACGCCTGCCCGATCGCTAGGGCGCCCGC
>BbSIPHT343	98	TGAAAACGCGGTTGCGCAGAGGGCTGCCCGACCCGCAAGGTTATGACCCG ACTATGACAGGATTATGAATCCCCCCCGGAAGACCGGCAAGCCGGCCCC
>BbSIPHT344	41	GTCTTTGTTCCGGTCGCGCCGTGGCGGGCGATAGGCTTTTCT
>BbSIPHT345	57	ATGGTTTGTCTCCTGAAAATTCCAGTTGCCGGCGGCGCGTTGCCGGCCGCC CTTTTCT
>BbSIPHT346	359	CAAAAAGGCACCCGCAAGGGTGCCCTTTCAGGCGAAAAAGACGCTTCTCC AGATATCTTTTCACGATAAAAGGGCACCCTTGCGGGTGCCCTTTTACTCTG AAAAGGCATTCTTTCAGATACTTTTTCACGATAAAAGGGCACCCTTGCGGG TGCCCTTTCACTCTGAAAAGGCATTCTTTCAGATATCTTTTCACGATAAAA GGGCACCCTTGCGGGTGCCCTTTCACTCTGAAAAGGCATTCTTTCAGATAC TTTTTCACGATAAAAGGGCACCCTTGCGGGTGCCCTTTTATCGCGGAAAGC CGTGCCAGGCACAAACGCGTCGACGTTTGCATGCCGGACACGGCCCCGGT CTCGA
>BbSIPHT347	294	ATAAAAGGGCACCCTTGCGGGTGCCCTTTTACTCTGAAAAGGCATTCTTTC AGATACTTTTTCACGATAAAAGGGCACCCTTGCGGGTGCCCTTTCACTCTG AAAAGGCATTCTTTCAGATATCTTTTCACGATAAAAGGGCACCCTTGCGGG TGCCCTTTCACTCTGAAAAGGCATTCTTTCAGATACTTTTTCACGATAAAA GGGCACCCTTGCGGGTGCCCTTTTATCGCGGAAAGCCGTGCCAGGCACAA ACGCGTCGACGTTTGCATGCCGGACACGGCCCCGGTCTCGA

Name	Length	Sequence
>BbSIPHT348	129	AAACAAGGCAGCTGCAAGAGTAACCCCGCCATGCAAAAAGGCACCCGCA AGGGTGCCCTTTCAGGCGAAAAAGACGCTTCTCCAGATATCTTTTCACGAT AAAAGGGCACCCTTGCGGGTGCCCTTTTAC
>BbSIPHT349	228	ATAAAAGGGCACCCTTGCGGGTGCCCTTTCACTCTGAAAAGGCATTCTTTC AGATATCTTTTCACGATAAAAGGGCACCCTTGCGGGTGCCCTTTCACTCTG AAAAGGCATTCTTTCAGATACTTTTTCACGATAAAAGGGCACCCTTGCGGG TGCCCTTTTATCGCGGAAAGCCGTGCCAGGCACAAACGCGTCGACGTTTG CATGCCGGACACGGCCCCGGTCTCGA
>BbSIPHT350	162	ATAAAAGGGCACCCTTGCGGGTGCCCTTTCACTCTGAAAAGGCATTCTTTC AGATACTTTTTCACGATAAAAGGGCACCCTTGCGGGTGCCCTTTTATCGCG GAAAGCCGTGCCAGGCACAAACGCGTCGACGTTTGCATGCCGGACACGGC CCCGGTCTCGA
>BbSIPHT351	96	ATAAAAGGGCACCCTTGCGGGTGCCCTTTTATCGCGGAAAGCCGTGCCAG GCACAAACGCGTCGACGTTTGCATGCCGGACACGGCCCCGGTCTCGA
>BbSIPHT352	327	AAACAAGGCAGCTGCAAGAGTAACCCCGCCATGCAAAAAGGCACCCGCA AGGGTGCCCTTTCAGGCGAAAAAGACGCTTCTCCAGATATCTTTTCACGAT AAAAGGGCACCCTTGCGGGTGCCCTTTTACTCTGAAAAGGCATTCTTTCAG ATACTTTTTCACGATAAAAGGGCACCCTTGCGGGTGCCCTTTCACTCTGAA AAGGCATTCTTTCAGATATCTTTTCACGATAAAAGGGCACCCTTGCGGGTG CCCTTTCACTCTGAAAAGGCATTCTTTCAGATACTTTTTCACGATAAAAGG GCACCCTTGCGGGTGCCCTTTTAT
>BbSIPHT353	91	ACGCATGCTCCAAAGAATTCAAGACGAGTAATTGTAAGCGGAAGCGGCGC CCTTGCCGCGCCCGGCGGCACATGCGCGGGCCGTCTTTTTCA
>BbSIPHT354	43	AAAAAACAAAGCCCCTTGAACAAGCAAGGGGCTTTGCCGGTACA
>BbSIPHT355	81	TCAAAACACACCGCAACGGATTGCCCGCAGCAGTGGAGTGGTAAAACCTG CCGGCTTCGCCGGCCCTTGCGGGGTGTCCCGT
>BbSIPHT356	41	GTAAAACCTGCCGGCTTCGCCGGCCCTTGCGGGGTGTCCCGT
>BbSIPHT357	201	AAAAAAAGGCCCCTCTGGGCCCGTTCATTCGTTGGATTTCGGGTACCTATG CTTCACCGCCTTGCACTGATCGATCCAGCGAGCCACCTCATCGGGCAACG GCACGCCTGCCACCTGTAACGCAGCCGCCAGCTTCATGGCGGCGTCCAGC TGGTCACCGATAGGCGGATATGCACGCGCACGGCGCTCAGCGTGAGACGC C
>BbSIPHT358	82	ACAAAAGCCCCGCCGACACAGTCGGCAGGGCTAAGAAAGAAAAAAGGCC CGCGCAGTGGCGGGCCGGACGGGAATGGCTCAGA
>BbSIPHT359	77	GCAAAAGCCCCGGCCATCGGCTCGGGGCTTCGTTTCTGACGGACGCACAG CGCCCGCCATGGGCATCGGGTCACGTCC
>BbSIPHT360	52	CCCCCCGGCTCCACCAATTACAAGGGCTCGGATAATCTCCGGGCCCTTTTT CT
>BbSIPHT361	83	AAGCGCGGTTACCTGAGCCCGTCAAAGCGGCTCAATTCTTTGCTCAAATGT AACTAAGGCCGGCGCAAGCCGGCCTTTTTTTCC
>BbSIPHT362	75	TCATGACGCGGGGCCGCCGGGTGGTCCGGCGCGCCGTACGTGTACCGCCT TCGGGTTCGCCCGAAGGCTTTTTTTT
>BbSIPHT363	150	TACGAGTCCTCTTAACTGAAAAAACGGCCCGAAGAAACACAAACAAGAA CCGTTTTCATTAATGTAGTTCGGTTCGGCTTACAGATCCCTTAAGCCGACC CCAATTTCAGGACTCATACGGAAAAGGGGCGCCCGTGGGCGCCCCTTTTC C
>BbSIPHT364	98	GCAAAAGGGCTGTTTCCCACATCGGAATCAGCCCCCCCCCCCCCCCCCCC TAAGACCTAAGATCGTGGCTCCATAACTCTTCTGGCGCCAAGACGCCC
>BbSIPHT365	52	AAAAAACAAGCTCGCGCATGCAGCGAGCGTTCCATTATTTCGGAGACCAC CCC
>BbSIPHT366	72	TTGCACGCAGGATACCGGAGCGGCTCGCCGCTCCGGCATGAAACGGCAAG AGCCTGGCTCTTGCCGTTTTTTT
>BbSIPHT367	32	TTTTCCCCTTGGGGGCTGCAAAGCCCCTTTTAA
>BbSIPHT368	62	TTTTCCCCTTGGGGGCTGCAAAGCCCCTTTTAAGACCCGCTCTAAACCAGC GGGTTTTTTTTT
>BbSIPHT369	68	GCTCCAGCTCCTATTGAAATGATCTGGTACAGGCCCGGCCCTGGCGGATC GCGCCTGCGCTGTTTTTTG
>BbSIPHT370	111	GGAAAAGGCACGGGCCGGTCATGCGACCGGCCCGGCTGGCTACAGGCTG ACGGCATTATGCGCCTGCCGGCGCATCGCCCGGTGCGGCGTCCGGCGGGC GGACGCCGCACGA

Name	Length	Sequence
>BbSIPHT371	75	GCAAAAGCGCCCCGAACGGGGCGCTTTTGCTTGGGCCTCAACTGGCGCGC CGGCTACGGCGACCATGCCTGGCGGC
>BbSIPHT372	89	AAAAAAGCCGCGCACTTGCGCAGCGATTGCGCAAAAGCGGCGTACCCCGG GTGCCGTGGCGGGACCGCCGCGTCTGGCGCGCCGGCCGCA
>BbSIPHT373	42	GAAAAAGGCTTGCCGACGGCAAGCCTTTTTTTTCGCCTCCGGC
>BbSIPHT374	162	AGAAAATGTCCGTGGCCAGCACGGACATCTGTCTTAATCTGGCAACACTG CGTCCCGGAATTCGACATCGCGCGGTACAGTGCCGCCGCATTCCTCGCATG CATGGAAAAATGACAACGCCCGGCCTGGGCCGGGCGTTCGCAGGCACTGC CGGCGGGCAGCC
>BbSIPHT375	274	TGAAAACCAAGGCCTGAACGCCGCGCGCCAGGCCGCCGCTTGCTGCTGTG CACGACGCCCGGCCGCCCGCGGATAACCCCTGCGCCATCGTGGGAAACAC CGTGGAAACCCTGGTTTCTGAAGCGTTTTGACGGATTTCTTACTGCTGTTT GTCGGCTGCTTGACAATCGAGCATGCGGCAGCGCAAAACCTATGAAAAAC CCTCATGGCAATTTGATAACCGGGTCCGCATACTCGCCGCAACTCTTAAAA AAAACCAAAGCGGATGCGGCGCA
>BbSIPHT376	63	ACAAAATTGCCCGCTTTCGCGGGCAATTTTGTTTTCGGCTGCCGGGCGCCG GCGCCCGCTTCGC
>BbSIPHT377	92	AAAAAAGCCGCCCCGAAGGGCGGCTTTTCACGCGCGGGCACGAGGCCCGC GCGTCAGCTTGGAATCGAAGCGGAACTTGCGTTCCGCAGCCGA
>BbSIPHT378	123	CAGGCGTCCGTCAAATGCGCAAACCTTTGGATTGTAGCATTTACGGAATTA CTTCCGGCTGAACCCAAGCGCTCGGAAACATGATTCAAAAGCGAAAAAAG CCGCCCCGAAGGGCGGCTTTTCA
>BbSIPHT379	91	CAAAAAGCCCATAGGCGCTGACGCATATGGGCTTTTTTGCTGGCTTTCTCT TGCGGGTCTTTCCCGCACCAGGGCCTGCCTGGACAAATCCT
>BbSIPHT380	126	CAAAAAGGCCTGACCCGGGGTCGGGTCAAACCTGAGATTGTATCCTGTTG CCGGCCGATGCGCCGTTACCAGGGCAAGAAAATCCGGGCGCACCGGGCG ATCCGGCGCGCCCGCCGTCCGGCTGGCG
>BbSIPHT381	107	GAAAAACGAGCCCCGAGACAACCGGGCGCCCTGCTACGCTGTTGCCGCCG CGCCGGATCCATTCTTGATCATTCTCGGACTCGGCAGGCTGGCGTGTCTTG CGGCCGG
>BbSIPHT382	72	AAAAAACGCGCCCCGTGGGGCGCGTCTGGACCAGCCTGCCGGGGCCCTGC GGGACCCGGCGTGGCCGTGGCCT
>BbSIPHT383	81	TTGTGCTCCGGTTCTTGAATTTTGGTAATCGTGAATGCTAATGCAGGGTGG CGCGCCGCCAGGCGCGCCGGTCAGATTTTCA
>BbSIPHT384	112	GTGGTCCGGCCAGCCTGGAGGGCCCGACCCGCGATAGTGGATCGGTGACC AGATCAGGGCTGCCGAAATACCCGGGAAACCCCTGTAGAACCGAGCGTCT GCGGGGTTTTTCT
>BbSIPHT385	165	GAAAAAGCCGCCCGGATGGGCGGCCTTTTTTTTGCGGCACGGCTGGGTTC AGGCCTGTTCGGCCGGGACCTTGGCGGCTGGCGCGGCCGTCTGCCCCTCG GCGGGCTTGTCGCTGCGCATGGCGTCAAGCCAGCGGCGCGGCGCGCTGGA GGCCTTCCACCAGCCG
>BbSIPHT386	55	CGGCTTTTTAAGCGACAACGGCTGGCCCGTGAGGTGCCAGCCGTTGTGCG TTTTGC
>BbSIPHT387	91	GTGGGAAATACCTTGCTATGAATTGAATACGGTAGGAATCAGGGTTTGAA ACTCAGTCCCCCAGTAACTGGGGTTGGCCACAGGGTTTTTCA
>BbSIPHT388	121	TCGGCATTACATGGGCTGATACCCATTTCGCCTGTCTCGACAGGAACAGTG ATGAACAGGTGCAATCCGTCGGCGCAACGATCCCCAGGCACGCGCGCTGC AGCGCCGCGTCTGCGTTTTCT
>BbSIPHT389	86	AAAAAAGAGGCCACCCGAAGGTGGCCCGAATTCCAGCTCTGCTCTGTCAC TCAACAGTACTACAGTATGGGCTGTCGCAACAGCCCC
>BbSIPHT390	76	CGGCCAGGCCCGGATTGCTTGTGGATAACTTATTTTAGCGCACCGGCGGCC CGGGGCGCCCACGCGGGCTTTTTTGC
>BbSIPHT391	53	GCAAAAGCCCCCCTGGGGGGCAGCAAGCTCGCGCAGCGAGCGCAGCGTG GGGGC
>BbSIPHT392	82	AGAAAAGCGCCTAGACCGCCGGCGTGCGCGGCCGCGCCGGTATGGCGCCG GCCGCGCGGCGCGGCTGGAGGGAGGAAGCGCGC
>BbSIPHT393	118	CGAAAACGCGCGCAAGCACGCCCAGGGCGCGCGGGTCAGACATGGGTAG CAGGCAGATGCATCATGGCGGGGATTTTAAGCCCCAGCGCCGGCCTGTCC CTAATCGCCCCGGCCGCCGG
>BbSIPHT394	83	AGTTTGTTTCCTCCAGGGATTTCGGCGCCGTGCGTAACGGGCCGGCTGTGC GGTGGGCGTACCGGTGTGCGGCGGCCTTTTTGA

Name	Length	Sequence
>BbSIPHT395	45	GCTGTCTCCTCGACACACGCGGCGCGGCCGCGCGATGGTGTTTTCC
>BbSIPHT396	70	ACGGTATGCGAGGATGGATGAACAATGGGCGCGCGGCGGTCAGGCCCGC CGCGCCATGCGGCCTGTTTTAT
>BbSIPHT397	80	GGAAAAGGACTTAGCGACTGCTCGCTAAGTCCTTGATCATTCTGGTGGGCT GTAAGTGGCTCGAACACTCGACCTACGGAT
>BbSIPHT398	70	GAAAAAAGCCGCTTCCATGAGCGGCTTTTTTCTTGCTGGCCTTGTTGCTGA CCTTGCTTGCTGAATCTGGT
>BbSIPHT399	236	AGGCTGAAGTCCGCGCGCCGACGCACGGCATCCGCCCGCGCGGCGCGACT CGGAAAACCAGCTTAGCAGGCTGCGGGATCGCGCAGGCAGGCTGGCGCG GCGGTTTCCCGATTGTCGGGAAACCGCGTCGCGGCAAGCGGATATCGAAG GCGGGAAGATGGCCGCGCGCAACAACCAGGCGCACCGCATGAAAAAGGC CAGCAAGAAAAAAGCCGCTTCCATGAGCGGCTTTTTTCT
>BbSIPHT400	89	TCAAAATCGGCCGGGTGCAAAAGCCAAGCCGCCCGAGGGCGGATTGGCG GGATGCGGGCCCGGCCACCGCCGGACGCGTGCTGCCGAGGC
>BbSIPHT401	73	GCAAAAGCCAAGCCGCCCGAGGGCGGATTGGCGGGATGCGGGCCCGGCC ACCGCCGGACGCGTGCTGCCGAGGC
>BbSIPHT402	131	GAGCATCCCCCGGCTTTAGCCAAAAAACCATTAAAAGACGCTATTTTGAC GGATTTGGGCCCTGAAATAAACTCGCCCCCTCATTTTGAGGGTGCCGGCAT GCCGTAGGCGGCGCTTCAGGCGCCATTTTGC
>BbSIPHT403	41	TCTGCCTGTGCGCCGGCACGCGCTGCGCGAGCCGGATAGCGA
>BbSIPHT404	220	ACGACTCTATTATTCTTGCGAAAGCCGGAAATCTACCAATAGACACCGGC ATAACCGGGCTATTGTAACCTCCAGGGAACAATGCTGCCAGCCAAGTCCA TTGCCAATAAATAATGCGTCCCTTTTATTGGCGGACAGGCGGACCGCTCCG ACTCCGGCCCGCCGGGGAAAGGGCGGGCCGGAGTTGAGGGAGTTAGAGC CGTAGGACGGCTCTTTGGTTC
>BbSIPHT405	434	CGAGGTGCGGCGGATGCGGGCGCCTGGGCACCTGCCCGCGGTCACGGTTG GCGGGTGCGATGCCGAGGCAGGACTCCACGGCGGGTGTCTGTCCCCGCAC GGGGACTGACACGCGCTGCAGGCAGCACGGTCAGAAGCGGGGGATGTCT GCAGGTGCCAGTCCCCCTTCGGGGACAGGCACCCGGCGACGTGCGGCGGA TGCGGGCGCATGGGCGCCTGCCGTCACGGTTGGCGGGTGCGATGCCGAGG CAGGATTCCACAGCGTGTGTCTGTCCCCGCACGGGGACTGACACACGCTG CAGGCAGCACGGTCAGAAGCGGGGGATGTCTGCAGGTGCCCGTCCCCCTG CGGGGACAGGCACCCGGTGAGGTGCGGGCGCCTGGGCGTCCCCCCAAAA ACAAAACCCCGCGAGCCGTGAGGCTGCGGGGTTTTGA
>BbSIPHT406	62	AGAAAACCAAACGGCACCACCAAGTGCCTTTTTTGTGCCCGCCACCCAGC ACGAGCCGGAGCA
>BbSIPHT407	126	AGCAACCCCAGCGCCATCGGTGACGCATGGCCGGCCGTACACGGCGGCAC AAAACGTTGACATAGCAACTATATGCTCATAGGATGAGCACCTCGCCCGC CGCCAGCCCGGCCGCGGCGCGTTTTCC
>BbSIPHT408	78	GGAAAACCCCCACGCTGCGCCCGCTGTGCGGGCGTGCCTGTCCCCGCCGG GACAGGCACACTGTCGTACGCCATCGACG
>BbSIPHT409	116	CCAAAACGGCAACCCCGCCCCATCTTAGCTGCCGTGCCTGCCTTGCGGACT CTACCGCATGGCGCAGGCTTGCGCCAAATCAACGGCGACGGCCCGCCCGC GCCACAAGACGCAGCA
>BbSIPHT410	49	GGATGTCTCCGTTCAGGCGCGGCCCCGCTGGCTGCGCGTCCAGGTTTTGT
>BbSIPHT411	180	GTGAGGTTCCGTAAAACGTAGAACGATTGTATCGCTCAAAAAACGGAGAC AGGCCGCGCAATGGGTACGCGCCGCCACCCCGTTTCCGACCGGACGACCG GCGCGGAGTTCCGCGCCGCACTGCAAGACGCCCGAAACCGGGGCGCCATA AGCAAAAAGGCCGCTTTCGCGGCCTTTTTGC
>BbSIPHT412	53	ATGGTTCCTAGATAAAAATATTTGCTTTGCCAGGACATCCGGCAAAGTTTT TTC
>BbSIPHT413	53	ATAAAACTTCGGCGCGCCCGGCTTCGTGCCCGCCGGCCGAGCCCCTGCCCT GCC
>BbSIPHT414	123	AGCACCTCGGCCGGGGCGCGTTGCCGCTCCGGCAAATCCGCACGACAGGC GCGCGAACGCGCGTTCCACGCATCGCGTCATCCAGACGCTGTCCGGCATC GCGCCGGGCGCAGGCGATTTTTTC
>BbSIPHT415	193	GGTGATACAGTTCTGGCCCCGAACCGGTGCTTTCGGCAAGTACGCTTGCGC GGAAGGTAAACGGGAAACAGGAAGGCCCTGCCCAGCCTGTGCTGCCCCCG CAACGGTCCCCATGGCGCGCGACGCGCCGTGGCAGCCCGATACCGGCCGG TTCGACAGCGAGGATGCGCCGCCCGCGGCCATGCTCATTTTCG

Name	Length	Sequence
>BbSIPHT416	36	AAAAAAGACAAGGCGTCGGCGCCTTGCCGACCCGGCG
>BbSIPHT417	229	GCCCGTGGCAAGACGTGCCTGTCCCCGCCGGGGGAGCGACATCTGCGGCC AGGGAGCCTGTCCCTGCCGGGGAGCGACACCTGCGGCCAGGGTGCCTGTC TCCGCCGGGGGAGCGACACCTGCGGCCAGGGTGCCTGTCCCCGCGGGGAC AGGCACCCGTCAAGCGACGGTATCTGTTCTTTCCCGCCTGCAGCCCACGTT TTGAGACGGCCCTTACGGGCCGTTTTTCT
>BbSIPHT418	77	ATAAAAAGCCGCGCCAGCGGCTGCCATCACCCACTGCGCGACTGGGCTGC CTTCAGTCTGCGCATCAGGAGACAAACG
>BbSIPHT419	70	AACGTGCTCCAGAAAAAAACGAAGGATCAGGCCTGCGATCGCGCGCCGA AGGCCCGCGATCCCAGTCCATC
>BbSIPHT420	247	GAAAAAGCCCTTCGGCGTAATGCCGAAGGGCTTTTTCCATTCCCGGGTGCC TGTCCCCGCACGGGGACTGGCACCCGCTGACTTCTGGGCCTGTCCCCGCAC GGGGACTGGCACCCGCTGACTTCTGGGCCTGTCCCCGCACGGGGACTGGC ACCCGCAAACTTCTCAGGTGCCTGTCCCCGCCGGGACAGGCACCCCGTCA TGCGCCATCTTCTGCAGGGACTGGCACCTCTTCCCCGCGAACCGCA
>BbSIPHT421	58	AAAGAAAGAGCCAGGGGAAGCAGGCATATTAGCCCCTGGCGCCAGGCGG CAAGTTTTGT
>BbSIPHT422	164	CGCTCGTCCCTCGAACCAAATGCGCAACGCCTCGTGGCGTTGGTGGCGTTT GTCTTCCGGGTGGCGATGATTCCCTTCCCGGCCCTGGACGTATCTGTCGCC TAGAGCCGAAGCTGTGAAGTGCAGCGCCAAGCCCGGAACCCGTAACTGGA CCGGGCTTTTTGC
>BbSIPHT423	102	CGAAAAGCGCGCCGCCGCGGGCCGCGCCGGCAAACCCCGCCGCTTGTTGC GGGCCGGCAACAGCGTGGGCGGCGCGGCCCCATCGCCGGCCTGGCCGGAC GCC
>BbSIPHT424	88	GAAAAACCGCCCGCCATCCACTGCGGCGCGTCACGCGCTACATCAGCGGT TTTGTTAAAGTGAGGGCTACAACAACGCACGAACCAGCA
>BbSIPHT425	113	CTCCAAGGCCGCCCGGCATGCCGGCGCGGCCTTTGTTCTGCGCGCGACGTT GCCTGCGCCGCTACGGCGGCGGCGTAACTTGATACCATGCGGGTTGATAT ATCCAGGTCTACT
>BbSIPHT426	192	GGAAAACTGGCCCCCGCTGGGCCTCGAGCGTCCGTAAGCAGTCAAGCGCG GTGGGCGCGCCATGCGCCCGCCGCGATGTTTTCATGCATCGGAAGGTGCA TACTTTTTCCACCGGACCGCGGCCTGTCACAGACCGATAGAAGAGCCGGA TAACCCGATGGCGGAAACGCCGTCCTTAGATTCAAAGGAAATA
>BbSIPHT427	52	GGAAAACCGACGCTGGCCGGCGGCCGGGCGCACACTTGGTTACGCGGCCG GCA
>BbSIPHT428	170	AAAAAATGGGGCACCGATATGCTAGGTGCCCCGCAAGCATCCGTCTCATA GGGGAGGGGAAAGAGGAGAAGCCGAGACGGATGACAAGACTGCAACCGT TGGCATGGCCAAGTTCGCGGTCTACGTATGTTTAGGCGCACGAGCGGCGA TTTAGTTTCCAGCGTGCCAGGG
>BbSIPHT429	152	CAAAAATGCGGGGTTCCCCGTCGGGGCCCGCATGTCATGGCGTAGCAGGG TGTGCCTGTCCCAGCGGGGACAGGCACACCCTGCCGCAAAGGCCCCCAAA AAAACAGACCGCCGTAGCGGTCTGTCTTGCATGCGCGGGCGCCGCGTACG GCG
>BbSIPHT430	282	GGGCGAGCGGAGTGGAATTGGAGGGGTTTGCGGCGGGGGATTTTGAGTCG GTGATTTACAGTGTTTCTACAGGGAAATGGGGCGAATTCTAGCGTAGGTT GAAATAGGTAAAACCAGTTTCAAATCAGTAGGTTAGCGTATTTTAACTGA ATTGGTGAAAACTGGTCACTGTACCTGTACTGTATCGAAATTTCGGTACAG CCTTGAAGGGGAAGGGCCAGATCGGAATGTCCGTTATCACGCAAAAGGGG CAGCGCTGACGTCAGGCTGTCAGAGGTTTTGG
>BbSIPHT431	144	GGATAGATGCGTCCCTGATTATATTGACCCATAGGCGGACTTCACTTTCCC ATTTGCGCACCAACGGCCGGCGACCATTGCGCTCGCGATACAATGGCCGT TTCGCCCGGCGCGGGCCGCCCGTGGGCGGCCCCCAACCTTTTAC
>BbSIPHT432	54	GCAAAACCCCGGCTGGCCGGGGACATTGCCCGGCAACGGCGAGGCCCCCT CTATA
>BbSIPHT433	77	ATAAAATGGCCAGACTGACTGTCACGGCCTGAGCAAATATCAGCCAAAGA CGGCGCATGATTTTCGATAGGGCGAGTA
>BbSIPHT434	189	GGAAAACCCGGGCCGAGCGCCCGCGCAGGCGCGGCTCAAGTACACTCGG AGCGCCCTGCGCAACGGCGCCGCCGGGGCGAAGATTTCGCATAGCGAACG ATCGCAGCGCAACAAGCGCCGGCCGGATGAGGGGACGGCTATGCTTGGCC CAGGGCCGACAGGTCGTCGGTAGCGAAGGAGGCCGGATGAC

Name	Length	Sequence
>BbSIPHT435	58	GCATGTCTTCCTGTAACAATGGATCGCCGGCATTACAACAGGTCGGAGCC GGCTTGCCT
>BbSIPHT436	80	CCAAAAGGCCCGCCCCGCATGCCGGCCGGCGGGCTCCGACCCTTGCGCCC GCCCCGTTCAACCGTCTTTCGAGCAGTCAAC
>BbSIPHT437	195	GCGAGTTTCCTTGTTCGTATTCGTTTCGACGGCTGCGCCGCGCACGGCAAG TCATAACCTGGGCAGGGGTAGCCATAACCCGAGGCCGGGTTTACCCCCAG GTAAACGATACCCCTCCGGACATTGGCTAACGCTTTTGTGCATGACAAAAT CCCGAACCTATAAAGGGTTGTACCGCAAGGCTCGCCCTTTTTCA
>BbSIPHT438	93	TGAAAACGTGAAGACCGCGCGCAGCGCACCGCCCCCTTTTACAGAGCCGC GCGCAGTTGCGGCCAACTTGCAGCGGTCTGGCGCGTACGGCGCC
>BbSIPHT439	84	GGGCATTCTTCGACCTCCAACCACCCCGCGGATCGGCGCAGAACCAAAGA AAAGGCCCTTCACTCATAGGTGAAAGGCCTTTTTT
>BbSIPHT440	74	AAAAAACAGCCTCACGGCCACACCGTTCGCGCGCGCAAGCGCGCACAAAC GGCACAAACCCATCTGTGTCGCGGA
>BbSIPHT441	200	CGAGCCAGCCACGAAACGGCGCCCTGCGGGGCGCCGTTTTTTTCGTCCTG GGAATGTGACGGTTGCGTTGCAAGAAGGGGGACCGTCCCTCGTAGTTAGC TTTTATTTATTGCGCTGGATCAAGTAATCAATTTCCTAACGGCGGTCAGCT CACCTATGATGAGCAAGCGCCGTGCCGCCACCGTGTGGCGCGCGTTTTCA
>BbSIPHT442	74	ATGGGGTCAGGCTTGTTGCGGGTGACCCTGCATTATGCAGGATGGGGCCG GATCCTCCAGCCCGTGGGGTTTTGC
>BbSIPHT443	107	AAAAAACGGCCAGACTTTGCGTCTGGCCGAATCGATCGGGGCTAAGCGGG GATAGCCGCCCCGATACAAACCGGCAAACGCTGCGGGGTTTGCCGGTCTG TCCTGGTT
>BbSIPHT444	260	GTGAGCCGGTTGTACGTCCTTTATGCAACTCAAAATAGGGACCACGATGG TCGTTCAAACGGCCCTTGATCGAGACAGAGGCAGTGGGCCATGAGCGTTT GCTGTCGGCCAAGGGTAAATATGGCCGCGATGGCCTGTGATGCCGTGCAG GTGTTTGTAGAGGAGCGGGCGTCAAGGCATGCTTGAGTCAGGTGCCCAAA CTCTCATCGGGCCAGGCAGTCATTCGCGGGGATAGGCTCGACGGCCCATC CTCCTTTTTCG
>BbSIPHT445	54	ATGTCTCCTGTCGTTATGGTCTGCGCTTGCCGGTCGTGGCCGGAGCGCGTT TTGC
>BbSIPHT446	116	GGGCCTGAGACGGGCGCGGCCAGGTGCCGCGCCAGAATTTAACCGGGCGC CGGAACAAGTTTGCCGAGCCCGTTGTCTATGAGGGGGGAAGCCGATTGCG CGGCCCCCCTTTTTTCA
>BbSIPHT447	114	TCTTTCTCCCTGGTATACGGGAATTAGATCGGCGGCGCCGCCGATTATCAA TGCGCCCGTGCCTTGTGTTCCGTACTTCGGGAGGGTGCCGCAAGCGCCATG CGGCGCCTTTTGC
>BbSIPHT448	116	AGACCCATGTTCCGGTCCAGTATGACGCCAGGATCGGCGCTTTGCGAGAG CAGCGCGCCCGTACCAGTTCCAGGCTCAGCAAGACGAAATCCAGGTGCGG TTCGGCCTGGATTTTGC
>BbSIPHT449	59	GCTGTCTCCTGCCTGTCCGTTGCGGCGGCGCGCCCGCCCGGGCGCGCCGCC ATTTTTTCT
>BbSIPHT450	117	GCTGTCTCCTGCCTGTCCGTTGCGGCGGCGCGCCCGCCCGGGCGCGCCGCC ATTTTTTCTGGCTGTTGGTACTAGTGTTCCGTATCGCGGAACTTTTTAGGCG TATCCGCGATTTTGC
>BbSIPHT451	156	GGGCTGCCTCGCCCATGCGGCCTACCCCAAATGGGGGGTGCGGGCCGAGG GCTCGCCGGGCTAGGATAGGGCCAGCACTTCGCCGTGCTTTTATTCGCCGG CTCCGCGGCAGTCCAGGGGCATCTGCTGACGGCAACGGGGCCGGCGGATT TTTTCG
>BbSIPHT452	86	AGAAAACTGGCGGTCCGGCAGCCATCTGCCGGAGGCGCGGCCCGCGCCGC TCGCCGCTTGCGACCATCACCTGAAAGGATCCAGATC
>BbSIPHT453	66	CGAAAAGGGAAAACCCGCATTGAGGCGGGTTATTTCAATCATCTATCATT ATATTGTGATTCATATA
>BbSIPHT454	53	TGAAAACACGGGCGGCCGCATACCAGGCCGTCCATCCAACGGGGAGTTAC AGAC
>BbSIPHT455	147	CGAAAACCCTGCCCAGGCAGGGTTTTTTTTCGCCTGGAGCAGGGCGCGAC GGGGCATCCGGAGACCTGTGTAATCAATAATAATATTGCGAACAAGTCGC ATTCATAATAAAACACGAGTTTTTCTTGGCCATACTGGAGAAATTTCG
>BbSIPHT456	142	TCAAAACATCCCTGCACGCGGGATGCTCCCGTGCGTGGCCGCCACGCAGC GCCCGTGCCGCCGCGTATAGCCTGCTTACCAAGCGGCCGGGCAAAGGGGG GCCGAATGGGGCGCACGGGCGGTTACAGTAACGGTTGTGAGAT

Name	Length	Sequence
>BbSIPHT457	55	ATCCGTCCTACAATTCCCCGAGGTCGGCCCGCTCGTCGGGCCGGTTGCTTTTTTGT
>BbSIPHT458	162	ATGGAAGTGCGGCCATGCCACGCGCACGGGCCGCTTATATGCGTAGCCCGAAGGCTGGCGGCGAGTATAGCGAGCGGCGTGGCCGGAGCGGGCGAAACATATTACAAAAGACGACAAAATGTTCAAAAACAGCGGGACAGGCCGCAAAAGCGCGGCCAAGCGC
>BbSIPHT459	88	CAAAAAGCGCCGCCCAGGCTTGGGGCGGCGCGCCCCGGCCGCTTCTAAATGCTTGAAATCTCTAGGGTAATCCCGCAAGGGTTTTCCCT
>BbSIPHT460	65	TGTCTCCTCCACCCTCCTCCTTTGGTGGATTTAGCCCGAGATCGCAAGATCTCGGGCTTTTTTTTC
>BbSIPHT461	226	GAAAAAGGGCCGGAGCCAGGGGCCCGGCCGGGATTGGGAAATCGGAACCGTCATCATCGGCCGGGCGAGCCGGTCTTGTCCATAGCAATAAGTACTGGTCATACCAGTTTGTGCGAATGCACATGGCGACGAGATAAATATTTATATCTGGTTTTCAATCACATAGAATGCCGGTTCGAGTAGAGTCAGTAGCAGCGGATATCGTGGTCATACCAGTGGGGCAAGTC
>BbSIPHT462	93	CCGGATTGCATTCGTCCATCTGGTGGACGATTTTGGCTACACAAAATTTGACATCGATTTGACGCCCAACAGGTATGCTTGCGCGTCTTTTTTT
>BbSIPHT463	100	AAAAAAGGTGCCTGTCCCGGTTGGGACAGGCACCTTCGTCACCCGCGGGGCGTTGTGGGCCCGCGGGAGATGGTGCCGGGGGAGCCAGGCACCGGAGGGGA
>BbSIPHT464	82	GGAAAATGGCCGGGCCGCAATGCCCAACGGCGATGAACATCGCGCTCATCGCCGCTGCATGCCGGCGCCGGGGCGTCGCGGGC
>BbSIPHT465	56	GATGCCTCCTGGCGGTGCGGCCTGGTTCGGCGCGGTGGCCGTCCGGGCTTTTTTTCG
>BbSIPHT466	62	CAAAAACCACGCAGCGCATGTCGCGCCGCGTGTCTTAACCGGCGATCCATTCGCCAGATTGTT
>BbSIPHT467	70	AAAAAACGGCGCCCGTCCCTCTGTACAGGGCGGGCGCCGGCACTTACGGCTGCGGCGAGGGGGGCGCCGGG
>BbSIPHT468	129	CCAAAACCAACCCGCATGGCCGGCGCCAGCGTCGCGGCCTCAAGTGTTAAATTTCGGCGAACCTGCGGAAAACCACATCCAGCCAGGCGAGCGCACGTCCCGCTGTTTCGTACCTTGCCGAGCCCTCGAT

Sequencing sRNAs from the *B. bronchiseptica*

To identify small RNAs transcribed during growth, we isolated RNA from mid-log phase cultures of *B. bronchiseptica* RB50 selected for short RNAs of 45 bp or less for sequencing by the SOLiD sequencing platform [259]. The reads obtained from SOLiD sequencing were then mapped onto the *B. bronchiseptica* chromosome (L. Weyrich and J. Park, unpublished work). Regions with fewer than 50 reads mapping to a given location, or those that mapped to the same strand and location as known protein-coding genes, rRNAs,

or tRNAs were excluded from further analysis (See Materials and Methods for further detail).

Table A-2: Candidate sRNAs in *B. bronchiseptica* identified by SOLiD sequencing. The name, length, sequence, and fold-expression compared to the 16S rRNA of sRNAs identified by sequencing are shown.

Name	Length	Sequence	Exp.
bbsRNA001	98	TAAGAGCTTGAGTGCTCGTGTCAAGTGTCCACGCTTATCGGTTGTTGT TATATAGCTGCTGGATCGGTGGCTGCTGATCCGAGAGAGAAAGGTTT CGC	12.86
bbsRNA002	439	GTTGCAAATTCTGCAAAAGTCGTGTTATAGTTTCGGGCTTGGCAGTT GCCGAAACCCAAGGGTTAACCCTGATGTCGATAGGCGCGGGTGCTG GGGTTTCGAGAAGTGCCAGGTGAGAGGCGCAGGCCTTGAGGAAGGG GGTGAGTCAGGTGCGGAGCAGGCGAGGCCGACTTGACAAGCTGAAA AAACTTCTTCATAATCTCGTTTCTCTGCTGCTGAAAACGCAGCGCCG GTCGGAAGGCCGGTTGGTAGGACCGCTCTTTAACAATTTAACAACCG ATAAGTGTGGGCACTTGGTGCGAGCGCACGATGGCGGCTACGGTCGT CATCGAAAGCAAAGCATTATCAAGTGCTCACTGAATGAAGTAAGGTT TTTTAAACCTCACTTCCTTTGAGCAAGCGATACGGATCCTGGTTTCGG TCAGGGTCCACACACAGA	2.11
bbsRNA003	303	TAACCCGCGCGCAGTATCGGAATTCATGCAGTATTCATGGTTTGTCC CTGCCGTGCTCGTGACTTGGCCATACATTCTCTGAACCTATGTCTTAC GGCATATGGGTTGCGGGAGTGTAGAGCTGGAGTGATCGTCTACTCGT AGACGAACCCGATGCTCTTCGGATCGCGACCACCTTGAACCTCAGGG TTCGAGATGCCGGCCTTGACGGCACAGGCGGGGCATCTATCCAGGCA GGCCTGTATCGCAGGCCTGCCTTCGGATGTCTCTCTCCCTCCCCCACG TTCCCTCACAGGTCCACCA	0.86
bbsRNA004	215	ATACTTGTCCGCCACCCTGGCACGCTGGAAACTAAATCGCCGCTCGT GCGCCTAAACATACGTAGACCGCGAACTTGGCCATGCCAACGGTTGC AGTCTTGTCATCCGTCTCGGCTTCTCCTCTTTCCCCTCCCCTATGAGA CGGATGCTTGCGGGGCACCTAGCATATCGGTGCCCCATTTTTTTGCCT GCTCGTTGCGCAAGCCGCCCGCGGC	0.93
bbsRNA005	202	CTACTATTTGCACATCGGATGTTGAAACGCAGACGGTGCGGTGCCAG AACCGCCGGCCCAAGGCCAGCAAACATGGATACCCCGACTGCCGCG ATTTCCGGGTCGCAAGGCCTGACGCAATCCCGACATGCCACGGTTGT CTCCTCCACCCTCCTCCTTTGGTGGATTTAGCCCGAGATCGCAAGATC TCGGGCTTTTTTTT	0.86
bbsRNA006	133	GCTGGCGGGCCCCTTCGCATGGTTCGGCGGTGAATCTGGTCAGGTCG GGAACGAAGCAGCCATAGTCGTTCAGAACCAGTGCCGGAGTAAGGC TCGCCTACCGGTATCCCTGAAAGGGGCGCCGTGACCGGAA	1.3
bbsRNA007	141	GCTCAGATCACGGCCTTGCGCCAATCCATGGACAAACGATCCCCTGT CCGCGGATTGCAGGGCCAGTACTCTGACACCAATGCCGTGTTCCCCG CGAGCCTTATTGACGCGGCATTGCGTGCCCCGTACGGGCCCGGCACG	1.11
bbsRNA008	171	ACTCTCAAGTCTTAATTATTACCGCTGGACAGTAAGCGCCAGCGCCG TTCATACCAACACAACATACACACGCAACCCGCGTCGCTGCCAGTCC CAATCGACTGCCAGGCGCACGCTGTGCTTCTTCCAGATTCAAGCCGT TGACCATGCGTCAACCGCTTAGCCGGTTAC	0.82
bbsRNA009	139	TTAGTGCTGCGAGTCAGTGTTAAGCGTTGGGTTTTGGCCCGACAGCT ATATATGTTCTTTAACAATTTGGAAGAAGCACAACGTAAAGTGTTCG TTTAGTAGTCGACGCGAGTCGATGAAGACGGATACGGGTTGTGAT	0.99
bbsRNA010	83	CAACTTCCCTTCGCGGAGTTAGGGGGGCGGCAGGCCGGTATGTGCTG CATTGCTGCTCTTGTCACTCAAATCAACGGAGATTT	1.56

Name	Length	Sequence	Exp.
bbsRNA011	68	TTGCGGCTTACTCTTAATCCGTAGGTCGAGTGTTCGAGCCACTCACA GCCCACCATGATTTGTAAGGC	1.89
bbsRNA012	134	AAGTCGAAAATTTTAGCCGAATCCGCTAGGATACTGCCCTAATCATA CGCAACTGTCATCGTACGGTTGCTGTCCCGGCTAAATCCACCCGTTT GTTTCACGAAATACGCTAGCAGGCGTATGCCCCTCGCACC	0.76
bbsRNA013	197	CTTGACTTGAAGTTTCCCGTGCACGGTTACTATGCTTTGTGGCCGTGT GCATAATTTGCGTCTCGCAGTGAACACACGCGGGATAGGCAACGGG GTATAAGCACAGGTCCTGAGGTGGCTTACGGCTTGCCGCACGCGTTT CGGTGGTTCAACCCCGCGACGCGGCAGGCGCCGGATCGGCCATCTCG CGGACCCCA	0.49
bbsRNA014	135	TCCAGCACGCAGGAGACCCAGTAGCGATCCTTTGAACCTGTTCTTGG CGTTGCTATACTGGCCCCGTTTTCCTGATATGCGGCGGGGGTTCGCTG CGAGTCACTTATCCAGCTTCAAGCTCAACGAGGAGAAACA	0.68
bbsRNA015	89	AGTCGGCTTGCCCTGTCCAAGGCTCGCGCGGGGGACAGGCATACAG ACAACCTATATAGACTCGAACAACCAACCAGGAGTACACACCA	0.99
bbsRNA016	59	TCATCGTAGGGGCCGGCCAGGACGGGGCTTCCCGAACTGCCGGCCTA TTTCTTGTCTTA	1.46
bbsRNA017	48	ATCAGAATATTAGGTCCGGGGGGGGGGGGGGGGGGGGGGCTGATGG TT	1.77
bbsRNA018	57	CGCTAGGCTTCTTAATCCGTAGGTCGAGTGTTCGAGCCACTTACAGC CCACCAGAAT	1.35
bbsRNA019	274	TATGGCCGCCCAACGCACCTTCCTCGCGATTGCCTACGACCGCTCCA ATGCGGCCGTGGCCGGCCTGCGCCTCGCGGTGTCCCCCACCGCCCTA TCGCTACTACGCCTACCGATCGGTTGCCGGGCCTAGTGCCCCGTGGT AGCTCGGCAGCCGTACGCTCGCCACACGCGTTTTCCCCCGTTTTATTA AACTGAATCCGGCCGATTTACGGCCAAACTGCCCGCGAAACGACCA ACGCAATGCAGTGGGCCAGCCCCATGCGCGGCCGTCACC	0.23
bbsRNA020	158	TTATCCACAAGCAATCCGGGCCTGGCCGTACGTCCGTTCTTGTTCTTT TTGCAGCCCCCGGCATCTGCCCATCTTAATGACTGAATCTACCTCCTC CGACGCGCCTGTTACCGACGCGCCTACGCGCACTTTCGCAGACTTCG GCCTGCATCCGTTAC	0.39
bbsRNA021	63	ATCTTATCCGGGCGGTATCGCTCCGCCCACTGGAATTCGGCCGCGAT GGCCAGGAGTCAACCA	0.97
bbsRNA022	176	GCTTGCCTGGAAAAGCACGCCCTGTCTTGGCCCGGAATGTGATGTCG AACGTGATTTGTCTTGCGCGATGATAGGCGCGAGGACAGGCCGGTTG AAAGAGCAGGGCAGGCAGGCAAGTAGCAAAGCAGTAAGTGTAAGG CTATACGGTTTGTCTTGCGTTCGCACGGAGTCTTGCC	0.33
bbsRNA023	140	TGCATCAATCTCAAGAGCGCTTGCGTGCTTTCGAGAGCCCCCGCCAC GGCAATGGTTTGCCGCGCTGTACCCGCACCGGCGGGCACGACAGTCC GGGCAATGACCCGCCCATCGCTGTCGCGCCACGCCGCTCCAGTCCG	0.41
bbsRNA024	56	GAGAAATTTCCCTATCGCACGGTGATTGGACGGAGTATTTCCTGATC CAACGGGCA	1
bbsRNA025	45	ATGACGAATACGACCCACGACCTCAAGCCCGGGTACTACTGGTAC	1.23
bbsRNA026	109	CTAGTACTCAAGGATCCCGCACGCGGATCCACCTGCCCTACAGGAGG ACGCCCCCCCATTTACCCGCGCTCATGCGCACCCTTTTTTCTCTTTTC GGAAGGAGAGCAGA	0.51
bbsRNA027	182	GCGAGGTGTTGTTCATCGAATATCCGGCGAACACGCTGCCTCCATAT CCTCACTACTACGGATTTTGACCATGGCTACCAAAGCAAAAGCGCCT GCCAAGAAAGTCACGAAGATCGCCGCCAAGGCTCCTGCCAAGACTC CTGCCAAGGCTCCCGCCAAGAAGGCTCCCGCCAAACCCGCCG	0.29
bbsRNA028	467	ACGTGGCCGAGCGCGGCGGTTCGGCGGTGTCGGTGGTGGTGGCGAC GATGCAGGACATCTCGGCCAGCTCGCGCAAGATTTCCGAGATTGTGT CGGTGATCGACGGGATCGCGTTCCAGACCAACATCCTGGCGCTGAAC GCGGCGGTGGAAGCGGCGCGCGCGGGCGAGCAGGGCAAGGGTTTCG CGGTGGTGGCGGGCGAGGTGCGCTCGCTGGCGCAGCGCAGCGCCCA GGCGGCCAAGGAGATCAAGGTGCTGATCGAGGACTCGGTCGGCAAG GTGGGCACGGGCTCGCAATAGGTCGAGCGCGCCGGGGCGACGATGC AGGAGATCGTGGCCTCGGTCAAGCGGGTGACGGACATCATGGGCGA GATCTCGGCGGCCTCGGACGAGCAGTCCAGCGGGATCGAGCAGGTC AACCGCGCGGTGTCGCAGATGGACGAGGTGACGCAGCAGAACGCGG CGCTG	0.11

Name	Length	Sequence	Exp.
bbsRNA029	62	GCTCCATCCTGCCGGCAAGCCCCGTGCGGGAACCAGGCCGCGACCA CTACCAGGAGACGCCG	0.8
bbsRNA030	168	TTACTGCTTTGCTACTTGCCTGCCTGCCCTGCTCTTTCAACCGGCCTG TCCTCGCGCCTATCATCGCGCAAGACAAATCACGTTCGACATCACAT TCCGGGCCAAGACAGGGCGTGCTTTTCCAGGCAAGCCTATAATTCTA CTATTTTTCGACCGCGCCACTCAAGC	0.29
bbsRNA031	36	CGGAAATAGCGCATAAATGCGGCCTATTCAAGCTTT	1.31
bbsRNA032	93	CTGCACTGCACAAACCAGATTCGGCATACACGCGCTTTTGATGCGCA TTCGACGCAGTCTTTCCGGCACGGCTACACGCTGCCGGCCCATGCC	0.49
bbsRNA033	247	ATTGCTCAAAACCCCGCAGCCTCACGGCTCGCGGGGTTTTGTTTTTG GGGGGACGCCCAGGCGCCCGCACCTCACCGGGTGCCTGTCCCCGCA GGGGGACGGGCACCTGCAGACATCCCCTGCTTCTGACCGTGCTGCCT GCAGCGCGTGTCAGTCCCCGTGCGGGGACAGACACCCGCCGTGGAG TCCTGCCTCGGCAGCACCTCCGCCTATCGTAGTGACAAGCAGGCGCC CAGGCGCCCGCATC	0.18
bbsRNA034	36	CCTTGGATCGGCCCTCCTGGGGGTCGCATCGCCTAT	1.23
bbsRNA035	78	CGTGGAAGGTTGGCAGAGTGGTTGAATGCACCGGTCTTGTTCAATCG AGCGCCACGCGGGAAACCGCATGGCTGTAAG	0.56
bbsRNA036	95	ATTGAACGAGGCCGCCTGGGCGCAGCTAGCCGCCGGCGGGCGACAA GCGCCCATGGCCCCATCCAGACAACGGTTTTTATGTCATCAGGCATT GT	0.46
bbsRNA037	84	TTAACCCTAGGTTCGCCCATGCGTTGCGCGCGGCGAGGCCTCTTCCC GTTTCTTCCTTTGTGCGCTGGAAGGCGCCGTAATTTA	0.52
bbsRNA038	72	AGCAACATCCTTCGTCCTTCGGCCCCCCGCCCGGCCCATGCCGGGCG GGGGGCCGGGCGATGGCGGCTTTCC	0.6
bbsRNA039	71	TGTCTCTCAGGGTAAACCCCCCATTATCCACAGCCCTTCCTACCGCGT CCACCCACCCTTCCCTACACCCA	0.61
bbsRNA040	54	GCATTGCCCTGGCGTCCAGCCGCCCGTCCACCGCCCCTATGGAGCCC CGCCCCA	0.8
bbsRNA041	108	GTAAGCTAGAAAACATTCCGCATACGGCTGAGCGGGCTTACGCCCTC TCAGCCGTTTTTCGTTTGGGCAACGCCCGCGCATGTTCCGGGCGTGA TTCAGGTATAGAAA	0.4
bbsRNA042	50	ACGTAAAATGATTGAAAGTTAAAGAGATCTCCCGATCTCGGGGACA GACA	0.85
bbsRNA043	167	TTGGTTAGGGTTCCATACACTTTTGGATGACGCCTACGCCAATGTGC GAACAGTAAGTTTCTGGGACGCCAGATGCGCTCGCGACGAATTGGC GAACGCGCGGTTCGCCGCACCGGATCCCGACCTTGTCCGTAACCTGC TCGCAGACTGCCGTGCGAGCATCGCTC	0.25
bbsRNA044	55	GGTCAGGTACAGTTTCGGTACAGTGATGCGGCGCTAGCCCCATACGG AGTACTCA	0.75
bbsRNA045	93	AGCGAGGTTACAATGATGTCGTGCTGGCGGCGTGGCCGCCCGTCGCG CGGTTTGCGCGCCGCCTGGCGTCTGCCTGGCCGGTCGCGCGTTTCG	0.44
bbsRNA046	65	ATACTGCTCTGCGTCGCAGCTCCTTCCCTGGCTTCATATGCCTGGAAG TACCTGCGACGCAGCCT	0.61
bbsRNA047	57	ATCAGACGTAACCGACGTGGAGACCTGGGTCTGCCCTTGCCCCGTCC CCCATTACGC	0.69
bbsRNA048	60	GCGGTCGCTGTTACAAAATTAGCAGCTTTGCGCCGGGCGCTTGCCCA TCCCAGCCTTTCA	0.64
bbsRNA049	78	CATCAACAAAAAACACACAGCAACGCCCGGCCGGCGCACTGGGGAT TACCCGCGCAACCCGGCCGCCGCGACTCGCGC	0.5
bbsRNA050	81	CTTATGGATCCGCGTGCGACTTCCCCCGGCCAGAGCATTATTACCGC CATTCCAGCAGTGGCGGGTTAGCCTACGAATTCC	0.48
bbsRNA051	91	AGGCATCTTCCATGACCGTCATGACTGCCGGTACGTCCCGCGTCGCC GACCGTCCCGCTTGCGCGGACACCTCCGCCAGACGCCAGGGCGG	0.42
bbsRNA052	38	TGTGCAATGCAGCAAACGCGCCGGCGGTGAATTTTCCT	1.01
bbsRNA053	62	CCTATAATCGGGGTTAACCCTAGTAGCACGAACACATGCACTGGCCC TGGAGCCTACCGCGA	0.59
bbsRNA054	113	GTCCGATATGGTTTCCCCTAGTCTGACGGGATGCGCGACAGCCTTTC CTGGCTTCCCGTCCTTCGCCCCCACTCCGCCTTCCGTGCGCTGTGCCG GCGCTTGTTTGCTCGCGC	0.32

Name	Length	Sequence	Exp.
bbsRNA055	86	ACTCGTCGCCGGTGAGTGCTAACAACAGGCTCCCCGCGCCAAATACC GATCATTTTTCTTCTTTTCTTTAGAAACAGGAGTTCCTC	0.42
bbsRNA056	165	GAGATTCTTGGCCGGTCTTGCCAGAACGCTACCCGATATTTCGATTG ACAGCGCTTGTTCTCCGGGCGATCTTGCCTTCGAGACGCGTGGGTGT GAGTTGTTGCGACGTGCTGGGCCAGGGTTCCGCCGATAGCCCCCCGG CCACGCTCGCCTTTCGCCGCCGCG	0.22
bbsRNA057	58	CAAAAAATAGCCAGACCGCTGGGCAGCCGGACAAGCCGCGCTGCTC CCGAGGTCCGTG	0.62
bbsRNA058	75	CAGAAGCGGGATTAGCACCACCCGCCGACGTCGCGCGTTGCAGCCG CACGCGTTTTGCAAGTAACGGACTGGGAG	0.48
bbsRNA059	50	TTCCCACACAACCATCAGCCCCCCCCCCCCCCCCCCCCCGGACCTA ATA	0.72
bbsRNA060	64	GTAGTACTGTTGAGTGACAGAGCAGAGCTGGAATTCGGGCCACCTTC GGGTGGCCTCTTTTTTT	0.56
bbsRNA061	120	TATACGGCTTTCGAGGCTTGTACGTGCCTTCTGCCCCTGCTATCCGCT AAACGGGAAGCCCGTGTCGCGGAAGGTTTTCCTGCATCGTATCGAGT GAAGCACTCGTCAAGGTGAAGCGAT	0.29
bbsRNA062	37	GTTGAAGCAGGCGTCAAGCCAATACAGGGAGAGGCAA	0.94
bbsRNA063	111	ATCTGTTACACATGACCGGACCCCTACCATCCCCGTGCCCTAAAGTT TGCGCCCATGGTGTCGTTACACAGGCATAACCAGGCGGATCGCCCCCC ATCGGGCCGGTCCCGCG	0.31
bbsRNA064	65	TGCGCGTGCCGCGCCCGCGGCCCTGGCTGCCTCGCGAGCCGCCGCGG CACGCCCGTCATCGACGC	0.53
bbsRNA065	135	ATCCATGGTAAATTGCATTCAAATTGCGATGAAATGATACATTTCGC ATTAAGACGAAAGCAGGCTGGCGCGGGGTTCTCGGGGCGCGCCGCC TGGAGTTCAGCCCCCCGGCTTGCCTAAGTCTCAGGAGTCATG	0.25
bbsRNA066	86	AGCCGCGCCGGGCATACGCCGATCCGCAGAACTGAATATACTTTCCA CAGATCCCGCCTGTCCGGCGTTTCACCATTCGAGGTTGC	0.4
bbsRNA067	119	GACTTGTGAGGGGAACCTGACGATAGTGTCGGCACACACGAAATCA AGCGTTGCAATGTTTGTTCAGTGCCGTGCCTTCTCCGGCGTCCGCGTT ATTGTTCTACTGCCCCTCCTTGATT	0.28
bbsRNA068	93	AAAGAATTTCATATCCAGTTGTCCGGAAGGAGCCAGCCATGTCGCAA CGTCAACCCAACTTGTCGCGTCGTGCGTTCTTCGCCGGCGCAGCGA	0.35
bbsRNA069	201	ACTGTTGATTCACATATTTCCCATGTTCCAGCCCCGCGTCCGCGTTGC CATTGCCTGTACCTTCGGTGCGGCATCGGCGTGCGTGGTGCTGAAAG CCAAAAAACAGCCGCTCATCTGACCGGAGCTTGCTGTTCCGTCAGAT GGGCGACGGAACAGCGGCAAGCCCGGCGGCGCGCGGCCATGCGCGC ATCCCCTCCCTGA	0.16
bbsRNA070	39	TGCGCCCGCATCGGCTGGCGCAAGACGCGGCAGGGCCAT	0.82
bbsRNA071	220	TGATTGATTTTTCCGCAAAAGCCGTCGGATTTGCCTTGGCGCAGATC GTTTTCGCCGGCTTTCCTGGGTGCCCAGCCTGTCCTCCCGACAGGGG CGTCGCCCGTCCGCTTCTTCCGGCGACACCCGTATCGCTGCGACCAA GCCGCGCCCCCGAGCGCACACAGCCCTAGCGGCGTGAGTACCTGCCT ATCGTGCCCCTGCCGGGGCTTCTCGTGGCCCG	0.14
bbsRNA072	142	ACGTTTGGCATGCCGACCCGCGTGCCACCCTTGACGCATATCCGAGG CACGATCCAGATGCTGCTGCGCCCGACACGAACTACGACCGGAATCT TCCGCACTCAGTCGCGTTAATGATTTCATGCGTCGGGTTGGTATCGC G	0.22
bbsRNA073	76	TCGTTTCCCGGCGGGTTTGCCCGCTCAACCAGACAACAACAATCCAG TCACGGGTCTGTCCGCCGCCGTTGCGGCG	0.39
bbsRNA074	78	ACTCGATATTTCCACATATTACCCACCTATGACATCGCCGACCGGCC CTGCCCCGACGCGCGGCAGCGGGCCGGACCC	0.38
bbsRNA075	62	TAGAGCCAGACGGTTATGTGCCGAATGGAACACAGGGAACTCCATA ATTGTTGCAGGTTGGA	0.43
bbsRNA076	44	AGACAAAGAGCAGCCGGGTGCCGGACCACCCGCCACGAGGGGGA	0.6
bbsRNA077	36	CGGAGATCGCCAGGACCGCCTGGCAAGGAGCATCCA	0.73
bbsRNA078	41	GTCATTCCAGCCGTCGGTCGGGGCCAGTCCCGCCCGGGCGG	0.63
bbsRNA079	54	CTGCTGCTGTGACGACAGAAAGACGCCGGGCTGTCGCTGGGGCCAC CACTAGGT	0.48

Name	Length	Sequence	Exp.
bbsRNA080	75	ATACATACTTCACGATACCTGTCGTGATGTGCCAAAAAACCCCGGTC GCGCAAGCGCCGGGGTTTTTTCTTTGCG	0.35
bbsRNA081	45	ACTATGGTTCCCAGGGTCATCGCGCCATTACGCGAGCCGTTACTC	0.55
bbsRNA082	71	GTTTGCCAAACCTCATCCTCTGCTTGCGGAAACCCAATAGCCGTTGC CTGGCTGCGGCATCCGGCGGTAGC	0.35
bbsRNA083	204	CTATTCTTACTGTCACCAGGCATCAAGACGTTCACCGGCCTTGATGTC CAGCACCAGGCTGCGCACAGCGGCCCGGCATTCGTGCCGAAACACC CCTCCCGGCCCCACGGCCGCGGAACCAACATGCCGGCCCACAAAGC CGGTTCATTGCTACACCGTATCGGGCAGCGCCAGACGCCCACAAATC CAACCCGATCCGGTAGC	0.12
bbsRNA084	46	CGGACCGCCTGACGATGCAGCCGGTCGTGCCCCGTCGATTATTCCT	0.52
bbsRNA085	86	CGCGGCCCATCCAAGCCCAGCCGCGCGTTTCTGTGGAAACGCCGCGC ACCTGGATAGCGGCCCGGTAGGGGCCGGAACGGATGCGG	0.28
bbsRNA086	79	CTCGTCAGGGACGTGTCGGATACGCTGCCTATAATTGTGGGCCTCTA CCTGCCCTGGCGGGCCCTACCTGAACTCCTAG	0.3
bbsRNA087	148	TATATTGCTCCTGATGTCCTAGGGAAACCTGCGGGTATTGATCCTTG ACTTCGGATGTTAACCTGCGTGAATCTAGCCTGATCTGTCCACTGGCT TCCCATACAGGCTTGACCGAAGGCGATAGCGGCACTATTTATATAAC CGTGGC	0.16
bbsRNA088	195	CTTGTCCCGGAACCGGACGCTCACTAGAATTAGTGCCATACGAGGGG GTGGTACACTACATCTTGTGGTCGCTAAACACGCTGGACCCACCGGC CACGCACCCCGGACCATTCGTCCACGCGTAGCGATTCCCCTCACACC GAATACTTTTGCACCCGGCCAGGCGGCCCGCCTGGCCTTCCTCTACG CTTTTTC	0.12
bbsRNA089	50	ATTCAGCAAGCAAGGTCAGCAACAAGGCCAGCAAGAAAAAAGCCGC TCAT	0.46
bbsRNA090	71	CTTCCGTTGGTCGAATTTCCCGGGTGGCGCCATGTGCGTTGCCAGCC CACAAGCCGTACAGCGAAGGGGCG	0.32
bbsRNA091	127	TTAGAATGCCGGCTTCGGCACAATGCAGCCGTCCGCAACCCACGCGT AACATGGTATCCGGCATCCGGCAACCCTTGCGCCCCGCACACTGATA AGCCGGGAAGCGGCCAGAAACCCCAACAGGAGA	0.18
bbsRNA092	123	AATGGCTGTGCTGAATTCATTAGCGCAAGCCGCAGTAGCTGTCGCCA GTCAGGGACAGCGGCATGCGCCGATGAATAGCCCGGGTGCAATCCG CCTGTTTTGGGGGTTGCCACACGGGTTTAA	0.18
bbsRNA093	125	GCTTCTGTGGTTAGCCCAGCGGCAACACTGCAGCAGAAAGTGATACT CGATTCAGTATTATCACAGCATTATCGGCCTGTTTGTCAGGCAGGTG ATCAGACTGAATCCGGTATTTGGTGATTTTG	0.18
bbsRNA094	172	GCCGCGGCCTTGCAGGCGTCGCCACAGCGACCCTGCGAGTACCGCAT AAGCCAGCGCGGCCACGGCGTGAAATACAATGCCTGATGACATAAA AACCGTTGTCTGGATGGGGCCATGGGCGCTTGTCGCCCGCCGGCGGC TAGCTGCGCCCAGGCGGCCTCGTTCAATCAAC	0.13
bbsRNA095	37	GCTGTCACGCATGTCGCGGGAGAGAGCGGCCGATTGC	0.58
bbsRNA096	91	GTCCCCTGTATCTGCCGCCGCCCGAGCCCGGCCAGCCCGCGCCCACG GGGCAGCAAACGCCGCCCCCCTCCATCCCCCCTGCTCCGTCCCT	0.24
bbsRNA097	41	GATCAACTCGGGTCGGATCGCGTACAATGCGGCCCGTTTCT	0.52
bbsRNA098	144	CCTGTATCCTCCCGCGCGCCCGGATCGAACACTTCGGGGCGGCGCCG GGAATCGGGGCGGTTTCAGGAATGCCGTGCGGTCCGCGTGCGGGGA GGAAGTAACGGCCTTCCCGTGGTCTTGCGGTCCGCGCGCGCAGTGCT TCCC	0.15
bbsRNA099	34	GAATCCTCGTATCCGCCCGGCCTACGCCCGCACA	0.63
bbsRNA100	110	GTTCCGGTCCAGTATGACGCCAGGATCGGCGCTTTGCGAGAGCAGCG CGCCCGTACCAGTTCCAGGCTCAGCAAGACGAAATCCAGGTGCGGTT CGGCCTGGATTTTGCA	0.19
bbsRNA101	41	CCTTCCTTCCAGACGGCCTTCGCTCGAAGCCCGTTCCCCCG	0.52
bbsRNA102	40	ACGTTTGCTGGTCTGCCGATTTGCGGGCCGACTGTGCAAC	0.53
bbsRNA103	93	CGGCAGGGTGCGGCAGGGTGCGGCAGGGTGCGGCAGGGGCCGTCTG GCGGTGCTGCGGATGTTCGCGTGGGGGGGGGGGAGGCAGGGGAGCA G	0.23

Name	Length	Sequence	Exp.
bbsRNA104	111	TTCCAAATAGGCCAAAAGCACCAGAAAACCCCCCATAACCCGCGTTT TTGCCCGACGGGTCAGGGGGGATCATTGGCGTCGGTTCTAGACTTCG TTGGCAGAGGCCTGATG	0.19
bbsRNA105	48	ATCCATGCGCCGGCCGCGAGCGTCGCCCGGCGCATGGACGAACCGG CC	0.43
bbsRNA106	52	CCATAGATCGTGAAATCAATCGAGCCGGGGTTTCCCTGGATTTCGCG CTATG	0.39
bbsRNA107	181	CGCTTTTGCGGTCAAGGTTGTCGAAGTTGCTTTTCCGGCCGCAAGAC ACGCCAGCCTGCCGAGTCCGAGAATGATCAAGAATGGATCCGGCGC GGCGGCAACAGCGTAGCAGGGCGCCCGGTTGTCTCGGGGCTCGTTTT TCTTGAAACCCACGTCCGGGCGCCTCGGCCCTGACGCATGT	0.11
bbsRNA108	46	ATGCCGGGCCGCGCAGGCCGCCGCCGCGCGCCCGGTTCATGGCTTA	0.43
bbsRNA109	71	GTCTCATTTTTCCTGGCGTGCTGTCTGCCGGAGACACCGGGATCCATC AGTGTAGCTCGCCCAGCGGAGGA	0.27
bbsRNA110	127	GGTTTGACGCTTGCGGGCTAAAAGCGTGGCAAAAGCTGCGTCGTAA ATTGCTGGCAAGGGCAGGGAAGTTCATGCAAAAGAGTCTGGTTCTTA CCGCCGCTGATCGTGCGGCACTCGATGCGCGGGC	0.15
bbsRNA111	217	CCAAGCGACAGGCAGGCGGCAGAAGCGCGGAGCGGGCAGGGTATC GGGCAACCGGTGTTTGCAGGCTCCTGGTTCGGCGGCCAGATCAGCTG GCAAGAGGGAAATTTCCGCAGGATTTCAACCCAGGGTCGTTCGTTGC CAGCAGGCTGCGCACGGACCTGACGGGACCAAAACTGGCAGGAAAT GCGAACCTGGGAAGGGCGTGTTTTCAGTTAAG	0.09
bbsRNA112	96	GCAGGTCCGCGCCGGATTACCCATGCGGCACGGACTCGGAGATGGT AGCCCCTTACAACAAATGGCTGGGGCCTGAGCAGACCCCAAGAGGA GAGT	0.2
bbsRNA113	73	AATGCGGTCATTTTATAATTGACTGGTTTTTCTCTGCTTTTCCCCCTG GCCCGCCCTGTTTCGCGGCGCGGCC	0.26
bbsRNA114	128	AGCGCACCTCGCCCGCCACCACCGCGAAACCCTTGCCCTGCTCGCCC GCGCGCGCCGCTTCCACCGCCGCGTTCAGCGCCAGGATGTTGGTCTG GAACGCGATCCCGTCGATCACCGACACAATCTCG	0.15
bbsRNA115	82	TTTCCAGGAAACCCCTACTGATGGGTCGAGTTCAGCAGGGTTTCCTG GAGTCGTATCACCTGGGGCCGTGGTTTACGGTCCA	0.23
bbsRNA116	42	GGTCCGGCCGATGGCCAGGCCCGCCGCGCTTTCCCCCCGGCC	0.45
bbsRNA117	37	CGAAACTCTGGGGGCGGGGTGTATATGCCTGACGAGG	0.5
bbsRNA118	101	CAACAAAGGAAATCGCGGCGCTGTGCAAGCGAAAGTCCGATGTTAC AGATGGGCGGCCTAGCTGCCCGGTTTGAAGAAGCCTTTCTCTCTTGG GAGCCTCA	0.18
bbsRNA119	113	TTTGCAATATATTTTGGGTCGCAATTTTGGGCTTGGGTTCGGCAAGCG GCGCTGTCAGCAGCGCCCGGCCGGGGGCCAGGCAATACGGTGGGCT GGGCGTGCAGCCCATTTTT	0.16
bbsRNA120	170	TTGTACTCGTTCGACAATCCGGTCCCACCCGGAGCGGACGCGCTGTC ACTGTAATCCGACACGGGGCCTGCAAGCACTGGACGAACTCCCGCCC AGGCCCTAGTGTTTATAGCCACGCCCACGGCGCGGCCAGTATCCGGC AAACCGGGCCACAAGCCTGGCATAACAGA	0.11
bbsRNA121	74	AAAGGTTCCCGCGTATCCGCTATCCTTCGAGTAACGGCTCGCGTAAT GGCGCGATGACCCTGGGAACCATAGTG	0.24
bbsRNA122	130	TGCAGTTGATTCAAGCAGTGCGCAAACCCGGTTACTGTTGTAGATGG CAAGCCACAAACGCAACGTTGAAGCGGGTGCGTGTGCCCGCCGGTT GGTCAGTTGGTCAGCGTCCTTGATTCACTTGCATTTC	0.14
bbsRNA123	69	CAATAATGCAACGGCCGCCCGGCACCCGGAGTTCCTGAATGACCCAC GTAGTCACCGAAAACTGTATCA	0.26
bbsRNA124	72	GCCATGCTAACACCAAGGATAGGACTCACGTGACTCACCGTCCCGCT GCACTCTCGAAGCCCGCCTCCCGCC	0.25
bbsRNA125	121	AGCGAGGGGCTTTTCGATTTGCCTGTCCACCCACGCACCAAGAATGC CGCCAGGGCTCGCCTGGTCGAAGGCGGCCACGCGTGTCTTGTGTTTG TTCAATCCTAGGGGATCATCACTAGGA	0.15
bbsRNA126	45	TTCTTCGTTTGGCAGCTGACTCTTGTTTGCGCCAGACCAGGGATG	0.4
bbsRNA127	46	CCGTAGGCTTGGCGCTGCCCTCCGTGGCCCGGCGCCTGGCTGTCCT	0.38
bbsRNA128	35	GAAAGCGTCGCGCCGGGCGCACATACGGGAGACGA	0.5

Name	Length	Sequence	Exp.
bbsRNA129	37	CCCGTTGGTCTTCCCAGGCGTCCGCCTTTCGATCTCA	0.47
bbsRNA130	56	AAGCCAGCAAAAAAGCCCATATGCGTCAGCGCCTATGGGCTTTTTGT CGTCCAGGC	0.31
bbsRNA131	55	CTCAATCTACAGGCCTGTTTCCCACTCCATGGCCGCAATCTTCCTTTC GACGTCC	0.31
bbsRNA132	74	ATCGGGATGCAACAAACGACTGACGACGGATAGGCCGCCTCAGGTC ACTCCCGCAGATCAGACGGATGGAGGAA	0.23
bbsRNA133	74	ACCGTTTGCCTCACCCGCGCCAGCCCGCTTTCGCCATTTTGGCGGGG CGGGCTTTGCGGTAGAAACGGCGGCCG	0.23
bbsRNA134	116	CATTGCCCGCCGGCGAAACAGGCGGGCCTGACGCTGGCGGCGGCGC GCAAGCGCGCGTTGCCGACTGGCAAAGTATCGAGTGCCCTGGGTTGT CATTCTGATGGCCGGGGCTGCAA	0.15
bbsRNA135	43	ATTGGCACGTTGGCGCAATATCAACCAGCAGAAAGGTGACGCG	0.39
bbsRNA136	86	ATCCTGCGTCCGGCCGCGAATGGCGGCCCCTCGCAGCCCCAACATCC GAGCCGCGGCCTGCGCGACTCGACCTTTGCTTTCGCTTC	0.2
bbsRNA137	72	ATAGTCCGACCGCGCCCCCCGGCTGAAGGAACATGCGATGCCCGCC GTCACACGCCGCTCGATGACCCTGCT	0.23
bbsRNA138	234	ACACATTCAAGCCCCCGCAGGTGCATTTCTAGGGAAAACCCCGAGCA ACTGGCGGAAAACCAGGGTTTCCAGCGCCCTTGCCCTCCATTACCAT TCGGCCTTCAAAAATACGCAAGCGCCGATTTCAGAAAACCAAACGG CACCACCAAGTGCCTTTTTTGTGCCCGCCACCCAGCACGAGCCGGAG CAGGCAGTCTTTTAAATACATGCTGTAAAACAGGAGCCAACCAATCA	0.07
bbsRNA139	60	AGCTCTTTGGCGCCTCAACCATCTATTGTGGCCGGCAAACCTTGCAT CAGGCTTGCCGCG	0.28
bbsRNA140	52	ATTAGCAGGAGGCAGGTTCAGCAGTCTGGCGTGGACATTGTCCCGCG GCATT	0.32
bbsRNA141	36	GCGCATCGCCGCCGGCCGGCCTTCCAACCCGGCGCG	0.46
bbsRNA142	63	ACCCCATGAATAACCCGTACCCGAACGGGAATGCCGCCATGCGGCA ACCCGCCACGAGGAGGA	0.26
bbsRNA143	48	ATCGCCTGGACGGTTCCAGACCGGCAAGCCCGGCGGGCCGAGACGT TA	0.34
bbsRNA144	251	TCCCCGTGCGGGGACAGACACCCGCCGTGGAGTCCTGCCTCGGCAGC ACCTCCGCCTATCGTAGTGACAAGCAGGTGCCCAGGCGCCCGCATCC GCCGCACCTCGCCGGGTGCCTGTCCCCGCAGGGGGACGGGCACCTGC AGACATCCCCTGCTTCTGACCGTGCCGCCTGCAGCGCGTGTCAGTCC CCGTGTGGGGACAGACACACGCCGTGGAGTCCTGCCTCGGCATCGCA CCCGCCAACCGTGACG	0.06
bbsRNA145	61	AGTCCTGTAGGGCGCTGGCCGTTACCCTGCGCCTGATTAGCCCCGAC CAATCGCAGTCGGG	0.27
bbsRNA146	37	CTAACGGATACGGTGTATTAGTCCGTCCAGAAAACCC	0.44
bbsRNA147	123	AGCTTTCCCAAATTGTTGGTTCGATTGGGCCGGTTTTGGTTGGACTTT CCGGGGTTCTAGGCTTGGGTTAAACGAAACGACCTGGGTGGAATGA GCCTTCGGGCCTGGCATAGCCAGGCAGGG	0.13
bbsRNA148	97	GCTCCTGCGTCCCGCCACCTGGGTGCCTCCATGATCCGCTTCGGTCTG CCTTGCCCGCCCGCCGCCGCCTTTGCCGGCGGCCTGCCGCGCGCGCG CA	0.16
bbsRNA149	31	ACACCTAGACGAAACGTACTGGAGTAAGCCA	0.5
bbsRNA150	241	CACGTCGCCGGGTGCCTGTCCCCGAAGGGGGACTGGCACCTGCAGA CATCCCCCGCTTCTGACCGTGCTGCCTGCAGCGCGTGTCAGTCCCCGT GCGGGGACAGACACCCGCCGTGGAGTCCTGCCTCGGCATCGCACCC GCCAACCGTGACCGCGGGCAGGTGCCCAGGCGCCCGCATCCGCCGC ACCTCGCCGGGTGCCTGTCCCCGCAGGGGGACGGGCACCTGCAGAC ATCCCCAGC	0.06
bbsRNA151	107	CCGCACGGGGGACAGGCACCCGCTGAATCTTGTTCCCGCACGCGGG ACAGCCACCCGCTGAACCCATGCCGGCCGCCTCCCCCGACGGCGCCT CGGCAGCCACGAAA	0.15
bbsRNA152	116	TCCATCCGCGGTGGCATGGGGGGTCGATGGTCGACGCAAGCCAGCG TAGCCATCACCCGCTGCCCCACGGAACTATCGTTATGAAACCTTTAT GATAACGGTTTTATCCGTTTCCA	0.13

Name	Length	Sequence	Exp.
bbsRNA153	58	ATCAATCCAGGATGGCGTCTTCCGGGCTTGCCCGCGCCATGCATCGC ACCCCGAAGGG	0.27
bbsRNA154	72	ACAGCCTTAAGGGGCTGAACAGCCGCCTGGCGTTGAGATGCAATATT CGTTAGCGTCTCCGTCGGACGCGCA	0.22
bbsRNA155	37	CATGAGACGGCGGAAAGTCCGGGCGCGCCAGCGCCTG	0.42
bbsRNA156	101	AGTCGCTGCCCGCCGTCGGCCCGGCGCATGGGGATTCGCAGGGTTCT CATGCGCCGGTCAATGCCGGCCAACGGCGCCGCCTTTCCACCCTATT GCAGCCC	0.15
bbsRNA157	61	AACAACAATCATATATACCCTTCAGCCGGCGACCGAGTCATTTGAAA TAGTTAGGAATAAC	0.26
bbsRNA158	69	CCTATCCCTGAGGAGCGCTGCGACGACCCGGGCCGTGTCCACGGCAT GCCGGTCGCCAGGCTCAGGAAA	0.23
bbsRNA159	132	GACGCCTACCTGGCATTGTTGCGGCGGCCTCTCGATCACCGCCTGAA CGTTCTTGTGTTGTATGGCGATCGTTGGACGCATATTGGAACCGTCG AAGAGCACCTTGGTGCTTTTCAGAAGTATTCCCGCCAC	
bbsRNA160	46	GTATTTTGCGTGTGCATGCGGGCCTGGCGATGAAACGAGTGCTTGA	
bbsRNA161	44	ATTTTGAGTGTGCATGCGGGCCTGGCGATGAAACGAGTGCTTGA	
bbsRNA162	35	AGATCCAGGGCGTTCATGGGGCTACTCTACTTGCG	
bbsRNA163	152	CGCCGGCCTTGTCCACGCCGAGTATGCCGATGTGCCCGACGTGGTGA TGCCCGCAGGAGTTGATGCAGCCCGAGATGTTCAGGTACAGGTAGTC GAGATCGTCGAAGCGGCGCTGGATGGATTCGGCCACCGGGATGGAG ACGGCATTGGCC	
bbsRNA164	36	TGACCCTGTACTCGACCGCGCCCACGGGCATGTTCG	
bbsRNA165	68	AAGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC CCCCTTCAAACCTGCGCCGT	
bbsRNA166	79	CTCCAGGCGCGCCAGCGTTTCGGGGTCGGGCCAGAGGGTGCAGATG GCGCCCTTGGCCCGATGCAGGCGCTCGCGAACC	
bbsRNA167	120	CAGCGATGTCGCGAAGTTAACGCGCAGCGCATTCGCCAATGCGTTCT CAGGGGTTGAGATGCATTGGTCTAGTTGGTCGTCACGGGCAAGGCTC GCTTTGCCGACTTCGCCGATAGCAT	
bbsRNA168	128	CGAGATTGTGTCGGTGATCGACGGGTTCGCGTTCCAGACCAACATCC TGGCGCTGGAAGCGGCGCGCGCGGGCGAGCAGGGCAAGGGTTTCGC GGTGGTGGCGGGCGAAGTGCGCTCGCTGGCGCAGC	
bbsRNA169	128	CGAGATTGTGTCGGTGATCGACGGGTTCGCGTTCCAGACCAACATCC TGGCGCTGGAAGCGGCGCGCGCGGGCGAGCAGGGCAAGGGTTTCGC GGTGGTGGCGGGCGAAGTGCGCTCGCTGGCGCAGC	
bbsRNA170	146	CAGTCCGCGCAAGATCTCCGAGATTGTGTCGGTGATCGACGGGTTCG CGTTCCAGACCAACATCCTGGCGCTGGAAGCGGCGCGCGCGGGCGA GCAGGGCAAGGGTTTCGCGGTGGTGGCGGGCGAAGTGCGCTCGCTG GCGCAGC	
bbsRNA171	109	AAAAGCCCGGCTGGTGCCGTTCCTCATGGAATGAATCGCCACCAACC GGGCCATGTTCCAGGCCGACGGC- TCCATCCCAAGGAGGATGCCCAGCCGGCGCTGCTTGAC	
bbsRNA172	354	CTTATGCATAGCCAGAGGTATCACATGCAAAAACGGGTTCCCCTTTC ACCAGATGCGGTACACTGTTCCAGTTGCATGCTGGGGCACGTCTGCG TCCCCGTGGGCATGCCTGCCAACGAAGTCGAGAAGCTGGATGAACTC GTCAAGGGGCGCGTGCGCGTGGAGCGCGGCAAATCGCTCTACGGAC TCGACGATCCCCTGGACGCCGTCTACGGCGTGCGTTACGGCTCGCTC AAGACGCAGCTCGAGGATTCCAGCGGGCAGCTGCAGATCACCGGTT TTCATCTGCCCGGCGAAATCGTCGGCCTGGACGGCATGATCGAGAGC AAGCACGTCTCGAGCGCCGTCGCCCTT	
bbsRNA173	283	AAGACTCCGAGGTCTGCGTCATACGCCGGCCGGTAATCGACCGCGTC TCGACCCAGCTGCCATCGCTGCAGCAGCAATTCCGCCGCCTGTTGAG GCGCGGGTCACCCGCTCGCACCAGATGCTGGCGACGGTGGGCGCGA TGCGCTCGGTTCAGCGGCTGGCCGCCTTCCTGCTGAATCGGTCGCAG CGCTACGCGGCGCTCGGCTGTTCCTCGACGGAATTCGTGCTGCGCAT GAGCCGCGAGGAGATCGGCAACTACCTCGGCCTGACGCTGGAGACC GT	

Name	Length	Sequence	Exp.
bbsRNA174	120	ATGCTATCGGCGAAGTCGGCAAAGCGAGCCTTGCCCGTGACGACCA ACTTGACCAATGCATTCTCAGCCCCTGAGAACGCATTGGCGAATGCG CTGCGCGTTAACTTCGCGACATCGCTG	
bbsRNA175	41	GCAATTCTGGCGCGAGATTGCTGCCCAATCCCACCCCACCC	
bbsRNA176	63	GAAGAGAATGGCTACCAGGTAGTAGCGCACATCGAACTTCATCCGC GCGTCTTCAAAGTATTC	
bbsRNA177	276	CAGGTGGGCGGCGTGCGCGTGGGCGGCAGGTGCGCCGGGTTGACGG GGCACAAGCTGATCGGCGCGCGCGAGGTTTGCGGTGAAAGCGGGGT CATGCAAGGCGCCGACGGCCTGCGCGAGCGCCTGGAGCGCGTGGTG GCTGTGCGCAGCCGCGAGGGCTACATGGCCGATTTGGTCGCCACCGA CGATGGCTACCTGCTGGTGGAGAATCACTGTTCCATCTGTGCGGCCG AGCAAGCCTGGATGGC-TTCTGCCGCAGCGAACTGGACCTGTTT	
bbsRNA179	117	GCGAAAAGGGTTAGCAGGCACTTACCGGCTAACCATTACAAATACT GGTGGGCTGTGAGTGGCTCGAACACTCGACCTACGGATTAAGAGGC CGCTGCTCTACCAACTGAGCTAACA	
bbsRNA178	208	ATTGGTGCAGATGAGGTGATCGTGGTGATCGCCATCATTCAGCTCGA ACACCGCCTTGCCGGTATCGAACTGGCTACGCGTGAGGATGCCCGCC TGCTCGAACTGCGTCAGCACGCGATATACCGTCGCCAGCCCGATTTC CACATTCTCGGCGATCAGCGCGCGATAGACATCTTCCGCGCTGAGGT GGCGGAGGTCGGATTTACGA	
bbsRNA180	34	CCCTTACAAATCATGGTGGGCTGTGAGTGGCTCG	
bbsRNA181	124	GACAGATGCCGGTGGCCTTTATCTCCACCAGGACCTCTCCGGCCTTG GGGCCCTCGAGGTCGGCGTCTTCGACCGTCAACGGGGCACCAGCTTT CCAGGCGATAGCGGCTTTCGTCTTCATATA	

Comparing sequenced and predicted sRNAs

SIPHT-predicted and SOLiD-sequenced sRNAs were compared by identifying overlap in the genomic position of each candidate sRNA. Through this analysis, I identified 24 sRNAs in common between the two methods, three of which are predicted to be on opposite strands in the same position (Table A-3)

Table A-3: sRNAs common to both SIPHT prediction and SOLiD sequencing. The sequenced candidate sRNA and its correlated predicted sRNA are shown, along with the length of the sRNA, the strand on which it is encoded, and the genes encoded upstream and downstream of the sRNA.

Sequenced	Length	Strand	Predicted	Length	Strand	Upstream	Downstream
bbsRNA004	215	-	BbSIPHT428	170	-	BB4834	BB4835
bbsRNA007	141	-	BbSIPHT231	211	-	BB2062	BB2063
bbsRNA011	68	-	BbSIPHT036	41	-	BB4137	BB4138
bbsRNA018	57	-	BbSIPHT397	80	-	BB4132	BB4133
bbsRNA038	72	-	BbSIPHT090	140	-	BB2403	BB2404
bbsRNA060	64	-	BbSIPHT389	86	-	BB3975	BB3976
bbsRNA064	65	+	BbSIPHT168	179	+	BB1409	BB1410
bbsRNA072	142	+	BbSIPHT242	148	+	BB2149	BB2150

Sequenced	Length	Strand	Predicted	Length	Strand	Upstream	Downstream
bbsRNA089	50	-	BbSIPHT398	70	-	BB4168	tRNA-lys
bbsRNA092	123	+	BbSIPHT130	155	+	BB4377	BB4378
bbsRNA098	71	+	BbSIPHT304	114	+	tRNA-pro	BB0317
bbsRNA100	110	+	BbSIPHT448	116	+	tRNA-tyr	tRNA-gly
bbsRNA107	181	-	BbSIPHT381	107	-	BB3882	BB3883
bbsRNA119	113	-	BbSIPHT325	69	-	BB3248	BB3249
bbsRNA130	56	-	BbSIPHT040	399	-	BB3874	tRNA-leu
bbsRNA138	234	+	BbSIPHT025	130	+	BB4220	BB4221
bbsRNA148	37	+	BbSIPHT184	38	+	BB1687	tRNA-met
bbsRNA156	101	-	BbSIPHT303	124	-	BB2992	BB2993
bbsRNA020	158	-	BbSIPHT390	76	+	BB3983	BB3984
bbsRNA150	241	-	BbSIPHT405	434	+	BB4211	5S rRNA
bbsRNA151	107	+	BbSIPHT118	143	-	BB4016	BB4017

Analysis of identified sRNAs

Distance to nearest ORF

Many sRNAs identified by sequencing could be processed UTRs, attenuators, and cis-acting antisense sRNAs. Candidate sRNAs that are encoded further away from open reading frames (ORFs) on the same strand, however, are more likely to be trans-acting sRNAs. To assess the likelihood of our candidate sRNAs to be cis-acting or trans-acting, I compared the genomic position of the putative sRNA to the genomic position of the nearest open reading frame on the same strand. The results were grouped as less than 50 bp, between 50-200 bp, and greater than 200 bp away from the nearest open reading frame on the same strand. sRNAs identified as transcribed on the opposite strand of a known open reading frame are likely to be cis-acting sRNAs that target the gene from within which they are transcribed [260]. The majority of sequenced candidate sRNAs are within 50 bp of the nearest ORF on the same strand (Fig. A-1), suggesting that some may be cis-acting, or involved in translation regulation.

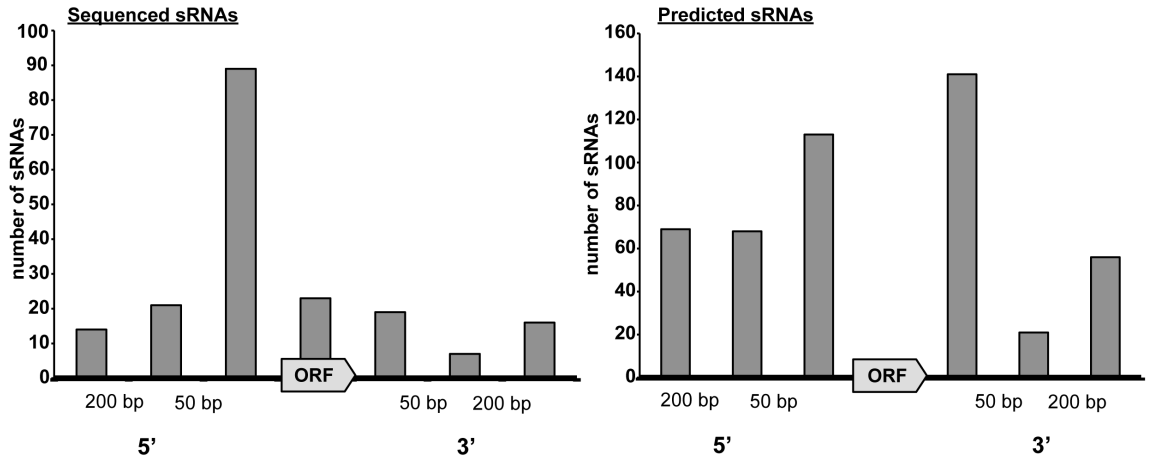


Figure A-1: Distance of candidate sRNAs from the nearest ORF. The distances of the sequenced (left) and predicted (right) sRNAs from the 5' or 3' end of the nearest open reading frame on the same strand are shown. Because SIPHT only predicts sRNAs from intergenic regions, no sRNAs are shown within an ORF, as with the sequenced sRNAs.

Confirmation of sRNAs by northern blot

We designed oligonucleotide probes complementary to selected candidate sRNAs for confirmation of independent transcription via northern blotting. RNA was isolated from mid-log phase cultures of RB50. 10 μg of this RNA was loaded onto a polyacrylamide/urea gel and transferred to a membrane. Then γ^{32}P-ATP-labeled oligonucleotide probes were hybridized to the samples, and the presence or absence of a transcript was visualized using phosphorimaging. We confirmed a transcript for bbsRNA007, bbsRNA011, bbsRNA28, bbsRNA150, and preliminary data suggests transcripts corresponding to bbsRNA23 and bbsRNA133 (Fig. A-2). These are estimated to correspond to around the predicted size for each sRNA, but further work is needed to determine the exact size of these sRNA transcripts and how they are regulated.

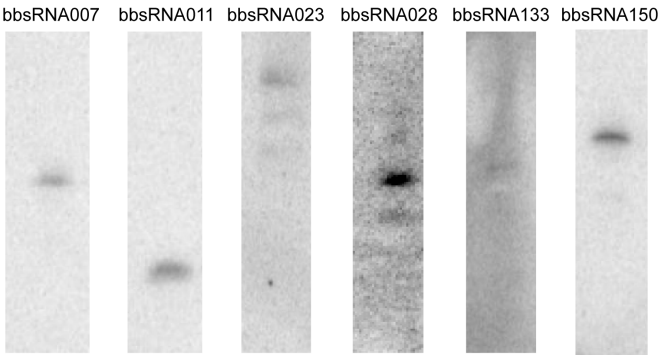


Figure A-2: Confirming sRNA transcripts with northern blotting. RNA isolated from mid-log phase cultures of *B. bronchiseptica* RB50 was transferred to a membrane following electrophoresis, and probed with a 32P-labeled oligonucleotide complementary to the indicated sRNA. sRNAs for which there is preliminary evidence of a corresponding transcript are shown.

Predicting targets for the candidate sRNAs

To identify potential mRNA targets for these sRNAs, the sequences of the candidate sequenced sRNAs were submitted to TargetRNA, which scores all the mRNAs of a given genome for their potential to basepair with a particular sRNA, and compares this to a randomized set of mRNA to calculate the probability that a given hybridization score would have occurred by chance from the genome [261]. The highest-scoring candidate targets of the six sRNAs confirmed by northern blotting are described below in Table A-4.

Table A-4: Predicted targets for selected candidate sRNAs. Target mRNAs for the selected genes are shown. When more than ten targets were predicted, only the ten highest-scoring targets are listed.

Name	Length	Sequence
bbsRNA007	141	GCTCAGATCACGGCCTTGCGCCAATCCATGGACAAACGATCCCCTGTCCGC GGATTGCAGGGCCAGTACTCTGACACCAATGCCGTGTTCCCCGCGAGCCTT ATTGACGCGGCATTGCGTGCCCCGTACGGGCCCGGCACG
	Gene	**Product**
	BB3888	hypothetical protein
	BB0593	CaiB/BaiF family protein
	BB3333	putative 8-amino-oxononanoate synthase
	BB0361	putative succinyl-CoA:3-ketoacid-coenzyme A transferase
	BB3287	hypothetical protein
	BB4702	probably acyl-CoA dehydrogenase
bbsRNA011	68	TTGCGGCTTACTCTTAATCCGTAGGTCGAGTGTTCGAGCCACTCACAGCCC ACCATGATTTGTAAGGC
	Gene	**Product**
	BB3472	phosphoserine aminotransferase
	BB3583	putative ABC transporter
bbsRNA023	140	TGCATCAATCTCAAGAGCGCTTGCGTGCTTTCGAGAGCCCCCGCCACGGCA ATGGTTTGCCGCGCTGTACCCGCACCGGCGGGCACGACAGTCCGGGCAAT GACCCGCCCATCGCTGTCGCGCCACGCCGCTCCAGTCCG
	Gene	**Product**
	BB3780	nicotinate phosphoribosyltransferase
	BB1265	hypothetical protein
	BB2404	hypothetical protein
	BB3865	cystathionine beta-lyase
	BB4740	hypothetical protein
	BB2989	fimbrial adhesin
	BB3090	putative conserved DNA-binding protein
	BB0078	ferric siderophore receptor
	BB1631	putative type III secretion protein
	BB4078	putative citrate lyase beta subunit
Name	**Length**	**Sequence**
bbsRNA028	467	ACGTGGCCGAGCGCGGCGGTTCGGCGGTGTCGGTGGTGGTGGCGACGATG CAGGACATCTCGGCCAGCTCGCGCAAGATTTCCGAGATTGTGTCGGTGATC GACGGGATCGCGTTCCAGACCAACATCCTGGCGCTGAACGCGGCGGTGGA AGCGGCGCGCGCGGGCGAGCAGGGCAAGGGTTTCGCGGTGGTGGCGGGC GAGGTGCGCTCGCTGGCGCAGCGCAGCGCCCAGGCGGCCAAGGAGATCAA GGTGCTGATCGAGGACTCGGTCGGCAAGGTGGGCACGGGCTCGCAATAGG TCGAGCGCGCCGGGGCGACGATGCAGGAGATCGTGGCCTCGGTCAAGCGG GTGACGGACATCATGGGCGAGATCTCGGCGGCCTCGGACGAGCAGTCCAG CGGGATCGAGCAGGTCAACCGCGCGGTGTCGCAGATGGACGAGGTGACGC AGCAGAACGCGGCGCTG
	Gene	**Product**
	BB3076	hypothetical protein
	BB2381	hypothetical protein
	BB2910	hypothetical protein
	BB3185	probable class-V aminotransferase
	BB2900	hypothetical protein

		Gene	Product
		BB0888	putative two-component system response regulator
		BB4143	hypothetical protein
		BB4520	acetylglutamate kinase
		BB0954	thiamine monophosphate kinase
		BB3721	putative cytochrome p450 oxidoreductase
bbsRNA133	74	ACCGTTTGCCTCACCCGCGCCAGCCCGCTTTCGCCATTTTGGCGGGGCGGG CTTTGCGGTAGAAACGGCGGCCG	
		Gene	**Protein**
		BB3128	hypothetical protein
		BB4659	putative ATP-binding component of hemin transport system
		BB3334	hypothetical protein
		BB0701	hypothetical protein
		BB4740	hypothetical protein
		BB0955	putative phosphatidylglycerophosphatase
		BB1318	hypothetical protein
		BB4241	putative ABC transport ATP-binding subunit
		BB4489	threonine dehydratase
		BB4089	glycyl-tRNA synthetase alpha subunit
bbsRNA150	241	CACGTCGCCGGGTGCCTGTCCCCGAAGGGGGACTGGCACCTGCAGACATC CCCCGCTTCTGACCGTGCTGCCTGCAGCGCGTGTCAGTCCCCGTGCGGGGA CAGACACCCGCCGTGGAGTCCTGCCTCGGCATCGCACCCGCCAACCGTGA CCGCGGGCAGGTGCCCAGGCGCCCGCATCCGCCGCACCTCGCCGGGTGCC TGTCCCCGCAGGGGGACGGGCACCTGCAGACATCCCCAGC	
		Gene	**Product**
		BB3277	ABC transporter, permease protein
		BB1603	hypothetical protein
		BB0326	cyclolysin secretion protein
		BB1375	hypothetical protein
		BB0094	TetR-family transcriptional regulator
		BB0782	putative type II secretion system protein
		BB3642	hypothetical protein
		BB0729	putative short chain dehydrogenase
		BB2382	probable ABC transporter ATP-binding protein
		BB2443	hypothetical protein

Putative SigE-regulated sRNAs

In *E. coli*, σ^E transcribes at least two sRNAs, RybB and MicA, which were primarily known for their role in targeting the mRNAs of outer membrane porins for degradation [77]. σ^E is a transcriptional activator that responds to cell envelope stress, particularly misfolded outer membrane proteins. When activated, σ^E directs transcription

of a number of genes that promote folding and assembly of outer membrane proteins, as well as degradation of aberrantly folded proteins [48]. Through transcription of RybB and MicA, σ^E can also downregulate the expression of porins post-transcriptionally, thereby reducing the burden that too many misfolded OMPs has on the cell [77, 80]. We previously determined that *B. bronchiseptica* SigE transcribes many genes involved in processes related to or located in the cell envelope (Chapter 4), and many other transcription factors, such as the heat shock sigma factor σ^{32}. To determine if SigE also transcribes any sRNAs, I compared the genomic position of the sequence and predicted sRNAs to the genomic position of predicted SigE-regulated promoters (Chapter 4). From this analysis, I identified three putative SigE-regulated sequenced sRNAs, and two putative SigE-regulated SIPHT-predicted promoters (Table A-5). Because the RNA isolated for analysis was isolated under non-stress conditions, when SigE activity is likely low, many putative SigE-regulated sRNAs would not be identified in this analysis. To enrich for SigE-regulated sRNAs, we have submitted a strain with high SigE activity (RB50Δ*rseAB*, described extensively in Chapters 3 and 4) for whole transcriptome sequencing.

Table A-5: Putative SigE-regulated sRNAs. The sequence of the indicated sRNA is shown, including the predicted SigE-regulated sRNA, in bold, with the -35 and -10 regions of the promoter underlined. Uppercase letters denote the coding region of the gene. Below, the predicted target

sRNA name	Sequence
bbsRNA047	**tgaattttcggctattgttcccatATCAGA**CGTAACCGACGTGGAGACCTGGGTCTGCCCTTGCCCCGTCCCCCATTACGC

Predicted target	Target Product
BB1808	probable GntR-family transcriptional regulator
BB0780	putative type II secretion system protein
BB2756	permease component of ABC transporter protein

bbsRNA125	**agaaattcgcgcgaacctataatcaaa**AGCGAGGGGCTTTTCGATTTGCCTGTCCACCCACGCACCAAGAATGCCGCCAGGGCTCGCCTGGTCGAAGGCGGCCACGCGTGTCTTGTGTTTGTTCAATCCTAGGGGATCATCACTAGGA

Predicted target	Target Product
BB1355	hypothetical protein
BB4346	putative AraC-family transcriptional regulator
BB0820	putative lipoprotein
BB3363	hypothetical protein
BB0157	3-deoxy-D-manno-octulosonic-acid transferase
BB1910	exodeoxyribonuclease VII small subunit
BB0595	hypothetical protein
BB4543	sensor kinase protein

bbsRNA058	**ggaacttgtgcgtggcagcatagtccaa**ttacCAGAAGCGGGATTAGCACCACCCGCCGACGTCGCGCGTTGCAGCCGCACGCGTTTTGCAAGTAACGGACTGGGAG

Predicted target	Target Product
BB2860	hypothetical protein
BB0889	putative two-component system sensor protein
BB0882	putative glycosyltransferase
BB3087	hypothetical protein
BB1081	LysR-family transcriptional regulator
BB2172	isoleucyl-tRNA synthetase
BB2803	putative heme export protein

BbSIPHT144	**ggaaattgcttcgctgtcggccaactaa**caggcCGCTGGCACATAGAACTTTGGGCCGCATCCGGACTTGTCCGGATGCTGCTGTCAAAGACCGCTGGAGCGCTAGCCGGTTGTTTGGCGAATGCTTAATCCCGAAGCTGGTTCAGTTCGGATCCAGCGTAGACGGCGTCCCCAGGAAGATTTCTTTACCCGAAGAACTCAACCAGGCGCCGCATTCGGTTGACGCGCAAGGCTAGCGCCCCAAGCGTCTTTTGA

Predicted target	Target Product
BB4638	putative polysaccharide deacetylase

	Predicted target	Target Product
	BB2666	putative transmembrane regulator
	BB3710	putative transport permease
	BB4407	phosphomethylpyrimidine kinase
	BB3277	ABC transporter, permease protein
	BB2235	phage-related hypothetical protein
	BB2831	iron utilization protein
	BB3485	phage-related hypothetical protein
	BB4270	putative acetyltransferase
	BB4135	hypothetical protein
BbSIPHT399	**agaactc**aaacatcgccacctcgc**tcaga**acagGCTGAAGTCCGCGCGCCGACGCA CGGCATCCGCCCGCGCGGCGCGACTCGGAAAACCAGCTTAGCAGG CTGCGGGATCGCGCAGGCAGGCTGGCGCGGCGGTTTCCCGATTGT CGGGAAACCGCGTCGCGGCAAGCGGATATCGAAGGCGGGAAGAT GGCCGCGCGCAACAACCAGGCGCACCGCATGAAAAAGGCCAGCA AGAAAAAAGCCGCTTCCATGAGCGGCTTTTTTCT	

Predicted target	Target Product
BB1081	LysR-family transcriptional regulator
BB4038	putative transcriptional regulator
BB1440	recombination protein RecR
BB2170	phosphopantothenoylcysteine synthase/decarboxylase
BB0883	putative glycosyltransferase
BB1530	N-carbamoyl-L-amino acid amidohydrolase
BB0648	probable regulator
BB1377	hypothetical protein
BB3081	putative membrane protein
BB3370	putative transmembrane transport protein
BB2766	putative short chain dehydrogenase
BB4949	biotin carboxylase
BB1855	putative integral membrane protein
BB1142	putative amidase
BB1556	ABC transporter, ATP-binding protein
BB2260	hypothetical protein
BB2810	cytochrome C-type biogenesis protein
BB2702	carboxymethylenebutenolidase
BB3687	aconitate hydratase
BB4810	probable transcriptional regulator
BB2421	hypothetical protein
BB3063	hypothetical protein
BB3300	RNA polymerase sigma-70 factor

I have confirmed SigE-dependent transcription for a promoter corresponding to that predicted to be upstream of bbsRNA058 (Fig. A-3). The nearest open reading frame encodes BB3108, a hypothetical protein, and the end of bbsRNA058 is only 7 nucleotides away from the beginning of the open reading frame. BB3108 also has higher expression in a strain with high SigE activity compared to wild-type. Taken together, one hypothesis is that this predicted sRNA is encoded as part of the 5'UTR of BB3108, and may or may not have regulatory function. Future work is needed to determine whether or not bbsRNA058 is transcribed independently of *bb3108* through northern blotting or primer extension to map the transcription start sites of each.

I have also identified a SigE-regulated promoter upstream of the gene encoding the sRNA-binding protein Hfq (Fig. A-3). However, in a strain where SigE activity is high, *hfq* expression is not higher than in wild-type, suggesting that even if SigE does transcribe *hfq* in *B. bronchiseptica*, it is possible that multiple levels of regulation prevent major changes in gene expression. In *E. coli*, Hfq influences σ^E activity through interactions with sRNAs that decrease OMP levels in the cell, but σ^E is not known to transcribe *hfq* [48, 262]. Consistent with this, by using the program we developed to predict SigE-regulated promoters in *B. bronchiseptica* (Chapter 4), I did not identify a sequence similar to a σ^E-regulated promoter upstream of *E. coli hfq*. I did, however, identify sequences with similarity to σ^E-transcribed promoters upstream of *hfq* in other bacterial species, such as *Burkholderia cenocepacia*, *Pseudomonas aeruginosa*, *Vibrio cholerae*, *Salmonella enterica* serovar *Typhi*, and *Legionella pneumophila*, but not in other species, such as *Yersinia pestis*, *Haemophilus influenzae*, or *E. coli*, as mentioned above. It would be of

interest to determine whether SigE transcribes *hfq* in *B. bronchiseptica* by either reporter assay or primer extension.

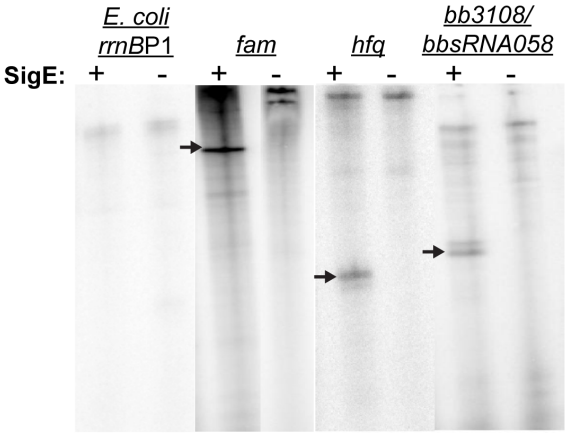


Figure A-3: SigE transcribes the promoter of either *bb3108* or *bbsRNA058,* as well as a promoter upstream of *hfq*. In vitro transcription was performed as described in the Materials and Methods section of Chapter 4. Briefly, His-SigE or buffer was incubated with *E. coli* core RNAP to form holoenzyme or a core RNAP control, then added to a pre-warmed mixture containing NTPs, α^{32}P-UTP, and the indicated DNA template. After 10 minutes at 30 °C, reactions were stopped, loaded onto a 6% acrylamide/8M urea gel, and after electrophoresis, the RNA was visualized by phosphorimaging.

Comparing identified sRNAs in the classical bordetellae

Although this work focuses on identifying sRNAs in *B. bronchiseptica*, putative sRNAs were also sequenced from *B. pertussis* Tohama I and *B. parapertussis* 12822. Preliminary results indicate that these data may be more complicated, which is why RNA from all the classical bordetellae are also being submitted for whole transcriptome analysis to provide additional evidence for candidate sRNAs in each species. At our level of stringency (see Materials and Methods), we identified only 43 candidate sRNAs in *B. parapertussis*, and we also identified 452 candidate sRNAs in *B. pertussis*, 208 of which are associated with transposases. Many of the early sRNAs were identified for their activity in limiting transposase activity [237], and the *B. pertussis* genome encodes many

transposable elements [139]. Further work is necessary to determine what role sRNAs may play in limiting transposable elements in *B. pertussis*.

I used two different methods to begin to determine whether any of the sRNAs identified in this analysis might be conserved between the bordetellae: 1) identifying whether the sequences of the sRNAs identified in a given genome are present in the other genomes (Fig. A-4A), and 2) comparing the sequenced sRNAs from all three genomes by sequence identity using ClustalW [263, 264] to identify which sRNAs are common to all three datasets (Fig. A-4B). After the whole transcriptome sequencing of *B. bronchiseptica*, *B. parapertussis*, and *B. pertussis* has been completed, we will be able to repeat this analysis, and ultimately identify conserved and unique sRNAs in the bordetellae. It has recently been shown that even evolutionarily divergent sRNAs can target the same mRNAs [80]. Therefore, even if sRNAs do not share significant sequence identity, they may still share a common target. Further work will be needed to determine the extent to which these regulatory mechanisms are conserved among the closely related *Bordetella.*

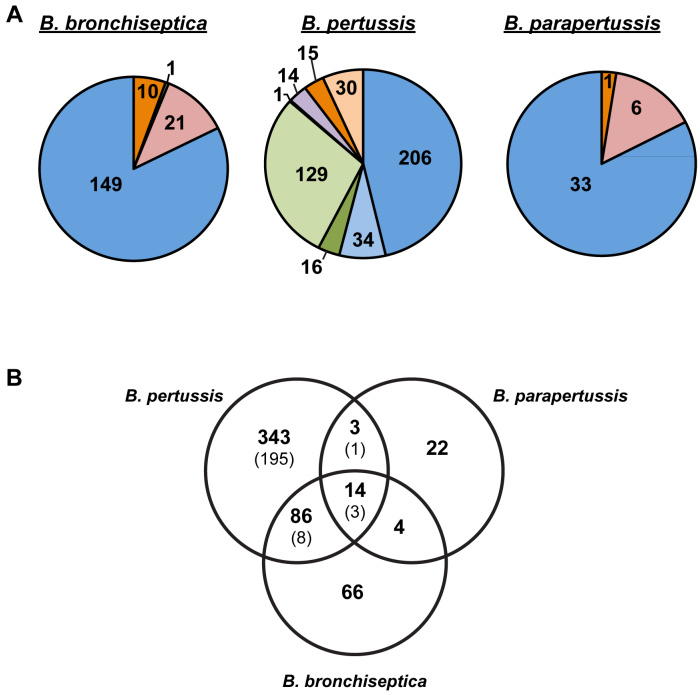


Figure A-4: Comparing candidate sRNAs sequenced from the classical bordetellae. (A) Identifying sequenced sRNAs from one genome in the other bordetellae: The sequences of sRNAs identified from *B. bronchiseptica (left)*, *B. pertussis (center)*, or *B. parapertussis (right),* were submitted to BLASTN to look for corresponding sequences in the genomes of the other bordetellae. Blue indicates sequences found in all three genomes, orange indicates sequences unique to that particular genome, purple indicates sequences common to *B. parapertussis* and *B. pertussis*, green indicates sequences common to *B. pertussis and B. bronchiseptica*, and light red indicates sequences common to *B. bronchiseptica* and *B. parapertussis*. For *B. pertussis,* lighter color indicates transposable element-associated sRNAs. (B) Comparison of sRNAs shared between the sequencing output datasets from each of the classical bordetellae. The sequences of candidate sRNAs from the classical bordetellae were submitted to ClustalW, and sRNAs with an identity above 70% were considered conserved. Parentheses indicate the number of the total sRNAs (indicated in bold) that are associated with transposable elements in *B. pertussis*.

In this work, we identified candidate sRNAs in the bordetellae, particularly *B. bronchiseptica*, and confirmed transcription of six by northern blot. Many are predicted to target virulence factors, and future work will determine if over-expression of these sRNAs does, indeed, affect the expression of their predicted targets and affect virulence-associated phenotypes in *B. bronchiseptica* (L. Weyrich, ongoing work). I have also identified five

putative SigE-regulated sRNAs, and confirmed SigE-dependent transcription upstream of one, as well as a SigE-dependent promoter upstream of the gene encoding the sRNA-binding protein Hfq. Future studies, including whole transcriptome sequencing of RB50 and RB50Δ*rseAB*, *B. pertussis* and *B. parapertussis*, as well as determining the regulation of identified sRNAs will provide additional insight into the role that sRNAs play in these important respiratory pathogens.

Materials and Methods

Predicting candidate sRNAs with SIPHT

sRNA identification protocol using high-throughput technology (SIPHT) prediction was completed in *B. bronchiseptica* strain RB50 as previously described [258]. To determine which sRNAs are present in both the SIPHT-predicted and sequencing analysis datasets, the genomic position of each sequenced and predicted sRNA was aligned, and overlap between these positions on the same strand was determined.

Bacterial Strains

Wild-type *B. bronchiseptica* strain RB50 was obtained from Jeff Miller and has been previously described [167]. All strains were maintained on Bordet-Gengou (BG) agar (Difco) containing 10% sheep's blood (Hema Resources) and 20 μg/ml streptomycin (Sigma). Liquid cultures were grown in Stainer Scholte media [168] overnight at 37°C with shaking until cultures reached logarithmic phase (OD_{600}~ 0.7).

RNA Isolation and Library Preparation

Bacteria for RNA isolation were collected from a liquid culture of *B. bronchiseptica* strain RB50 in logarithmic phase (OD_{600} ~0.7) and placed immediately into RNA Later (Invitrogen). Total RNA was extracted using the *mir*Vana miRNA Isolation Kit (Ambion) following the total RNA protocol that was provided. Isolated RNA was stored at -80°C until sequencing. RNA fragments smaller than 45 bp were collected and used for library preparation according to the manufacturer's instructions (Applied Biosystems) at the Genomic Core Facility at The Pennsylvania State University.

SOLiD Sequencing and Analysis

All sequencing was completed at the Genomic Core Facility at The Pennsylvania State University in University Park, PA, USA. Both SOLiD 3 and SOLiD 4 sequencing platforms were used to obtain 3,103,032 reads for strain RB50. 27% of the sequencing reads from SOLiD were mapped onto the reference genome of *B. bronchiseptica* RB50 with SHort Read Mapping Package (SHRiMP) using the default setting. Using the start and end position of coding sequences as well as the strand information (positive or negative) in reference RB50 genome, downloaded from National Center for Biotechnology Information (NCBI) database, we have excluded the reads that were mapped to genes, tRNAs, and rRNAs in the annotated direction. Reads mapping to the opposite orientation of the annotated gene, tRNA, or rRNA were maintained. The remaining reads were concatenated according to at least one bp overlap on the same strand. Reads that mapped to the genome sequence at least 50 times within the merged sequence region were labeled "sRNA candidates."

Genome Context and sRNA Conservation Analysis

sRNA candidates were reconfirmed using the reference genome sequence, and the reference genome sequence was then used for the further analysis. Upstream and downstream genes of candidate sRNA sequences were also extracted using Galaxy and by utilizing the start and end position of coding sequences of the reference genome RB50 [265-267]. Distances between the sRNA candidate and the nearest gene were calculated as the difference between the ends of the putative sRNA and the closest ORF on the same strand.

To determine which sRNA sequences from *B. bronchiseptica* are also found in the other classical bordetellae genomes, the sequence of each sRNA was submitted to BLASTN (NCBI) in order to search for the presence or absence of the query sRNA sequences in *B. bronchiseptica* (taxid:518), *B. parapertussis* (taxid:519), and *B. pertussis* (taxid:520) genomes.

Northern Blotting

Bacterial RNA was isolated by using the TRIzol Max Bacterial RNA Isolation Kit (Invitrogen) following the recommended protocol. The RNA quality and concentration were determined by 1.2% agarose-formaldehyde gel electrophoresis and readings of absorbance at 260 and 280 nm using a NanoDrop 2000 Spectrophotometer (Thermo Scientific). Northern blotting was performed by denaturing total RNA (10 μg per lane) and electrophoresis on a 6% polyacrylamide denaturing gel (National Diagnostics) containing 7.5 M urea. RNA was transferred to Hybond-N+ nylon membranes (GE Healthcare) and UV cross-linked. Membranes were prehybridized in 5X SSC/0.1% SDS followed by a 4 hour hybridization of [γ-^{32}P]-ATP radiolabeled oligonucleotide probes at 42-45°C. Washes of 15 minutes each with 2X SSC/0.1% SDS, 1X SSC/0.1% SDS, and 0.1X SSC were performed, and the blots visualized via phosphorimaging.

TargetRNA Analysis

To predict messenger RNA (mRNA) targets, each candidate sRNA was analyzed using TargetRNA, as previously described [25]. Briefly, potential targets for each sRNA candidate in the *B. bronchiseptica* genome were determined by the ability for basepair binding after RNA folding conformations. If available, the top twenty hits were all considered valid.

Appendix B

ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor σ^E in *Escherichia coli* by both direct and indirect mechanisms

Abstract

One of the major signaling pathways responsible for intercompartmental communication between the cell envelope and cytoplasm in *Escherichia coli* is mediated by the alternative sigma factor, σ^E. σ^E has been studied primarily for its role in response to the misfolding of outer membrane porins. This response is essentially reactionary, cells are stressed, porin folding is disrupted, and the response is activated. σ^E can also be activated following starvation for a variety of nutrients by the alarmone ppGpp. This response is proactive, σ^E is activated in the absence of any obvious damage to the cell envelope sensed by the stress signaling pathway. Here we examine the mechanism of regulation of σ^E by ppGpp. ppGpp has been proposed to activate at least two alternative sigma factors, σ^N and σ^S, indirectly by altering the competition for core RNA polymerase between the alternative sigma factors and the housekeeping sigma factor, σ^{70}. In vivo experiments with σ^E are consistent with this model. However, ppGpp and its cofactor DksA can also activate transcription by $E\sigma^E$ in vitro, suggesting that the effects of ppGpp on σ^E activity are both direct and indirect.

This Appendix published as: Costanzo A, Nicoloff H, Barchinger SE, Banta AB, Gourse RL, et al. (2008) ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor sigmaE in Escherichia coli by both direct and indirect mechanisms. Mol Microbiol 67: 619-632.
Barchinger contribution includes in vitro transcription experiments and demonstration that σ^E levels do not increase following gratuitous production of ppGpp.

Introduction

One of the hallmarks of bacteria is their remarkable ability to adapt to different environmental conditions and survive a wide range of cellular stresses. This ability relies on a sophisticated array of stress responses, many of which work by altering transcription so that genes required to combat a particular stress are induced, and those that may be deleterious under the stress conditions are repressed. These stress responses allow the cell to remodel its physiology to meet whatever conditions are encountered.

The bacterial RNA polymerase (RNAP) has 5 subunits, α, β, β', ω, and σ. The catalytic core of the enzyme (E) consists of the ββ'α_2ω subunits and is capable of transcription elongation and termination. The sigma subunit is required for promoter binding and specific transcription initiation. All bacteria have a primary sigma subunit, known as σ^{70} in *E. coli*, which directs the bulk of cellular transcription. Most bacteria also have one or more alternative sigma factors that direct transcription of specific subsets of genes. In most cases, transcription in bacteria is regulated in response to environmental and cellular cues by repressors or activators, proteins that bind to DNA near specific promoters and modulate the activity of RNAP at that promoter, or by alternative sigma factors, proteins that reprogram gene expression by replacing σ^{70} and redirecting RNAP to promoters specific for that particular sigma factor. *E. coli* has six alternative sigma factors, σ^{S}, σ^{H}, σ^{N}, σ^{E}, σ^{F}, and σ^{FecI}, which each respond to different cellular stresses [158]. These two types of transcriptional controls are not mutually exclusive. Repressors or activators can regulate transcription by holoenzymes containing alternative sigma factors, thereby integrating multiple environmental signals at a particular promoter to modulate gene expression appropriately.

Alternative sigma factors are regulated by dedicated signal transduction pathways that are activated by particular stresses, conditions, or developmental programs. For example, in *E. coli* the extracytoplasmic stress factor σ^E is activated by degradation of the σ^E-specific anti-sigma factor RseA in response to conditions that interfere with the folding of outer membrane porin proteins [185, 268]. The activity of the nitrogen-responsive sigma factor, σ^N, is regulated by the signal-dependent activation of enhancer-binding proteins, which are required for transcription initiation by $E\sigma^N$ [269]. Some alternative sigma factors can respond to several different stresses and are regulated by multiple regulators, utilizing a variety of mechanisms. A classic example is σ^S whose activity is modulated in response to an assortment of stresses including osmotic shock, acid stress, cold shock, and entry into stationary phase, by regulators that affect transcription of the *rpoS* gene, translation of the *rpoS* mRNA, stability of the σ^S protein, and activity of the σ^S protein [270]. Regulation of stress responses by these signaling pathways is essentially reactionary; damage or stress occurs, is sensed, and then the cell responds.

Reactionary stress responses, such as those described above, are not always the most effective way to respond to stress. When nutrients are scarce, the cell may not have the resources to rapidly mount a response that requires the energy-consuming processes of transcription and translation. In this case, a more general alarm that activates the individual responses in tandem could effectively preload the cell with stress factors, allowing it to combat stresses should they arise. In *E. coli* the alarmone ppGpp fulfills this role, as it can activate several alternative sigma factors independently of their dedicated signaling pathways [82, 271].

ppGpp is well known as a general signal of starvation stress [272]. The level of ppGpp in the cell is determined by the relative activities of the ppGpp synthase RelA and the bifunctional synthase/degradase SpoT [273]. The best-studied cellular role of ppGpp is its involvement in balancing the protein synthetic capacity of the cell with nutrient availability. ppGpp levels rise during nutrient downshifts caused by either abrupt changes in media conditions, such as starvation for amino acids, carbon, phosphate, or nitrogen, or upon more complex growth-limiting conditions, such as entry into stationary phase [272]. Under steady-state growth conditions ppGpp levels are inversely correlated with the growth rate of the culture, lower in rich media and higher in nutrient-poor media [274].

ppGpp, in conjunction with a cofactor DksA, can regulate transcription both negatively and positively [85, 86, 275]. Unlike more conventional activators and repressors that regulate promoter activity by binding to sites on the DNA and contacting RNAP, ppGpp and DksA bind only to RNAP and modulate its activity directly [86, 276]. ppGpp and DksA reduce transcription of ribosomal RNA genes by $E\sigma^{70}$ and therefore the number of ribosomes in the cell [86, 277]. In addition to inhibiting rRNA synthesis, ppGpp and DksA activate transcription of several genes by $E\sigma^{70}$, including a subset of the genes encoding enzymes required for amino acid biosynthesis [85]. The ω subunit of RNAP, encoded by the *rpoZ* gene, also contributes to the regulation of transcription by ppGpp. RNAP lacking the ω subunit cannot respond to ppGpp in vitro [278]. ω can also affect transcriptional regulation by ppGpp under some conditions in vivo [279].

ppGpp not only regulates transcription by $E\sigma^{70}$, but also has been shown to activate transcription by the alternative holoenzymes $E\sigma^{S}$, $E\sigma^{N}$, and $E\sigma^{E}$, during entry into stationary phase, and $E\sigma^{S}$ and $E\sigma^{N}$ in response to other growth-limiting conditions [82, 87,

90, 280]. The mechanism by which ppGpp regulates alternative sigma factor activity has been investigated for σ^S and σ^N, but is still not fully understood [87-89, 91, 281, 282]. ppGpp increases both the expression and activity of σ^S in response to a variety of starvation conditions, and increases the activity, but not expression, of σ^N during entry into stationary phase [87, 90, 280]. Since ppGpp binds to core RNAP, it can potentially regulate transcription by RNAP associated with any sigma factor. However, even though the activity of the σ^S and σ^N alternative sigma factors is clearly ppGpp-dependent in vivo, this activation may be entirely indirect because ppGpp has not been found to affect transcription by either $E\sigma^S$ or $E\sigma^N$ in in vitro transcription assays [88, 91, 281, 282]. This has led to a model in which ppGpp is proposed to activate the alternative sigma factors indirectly by altering the competition among sigma factors for core RNAP such that the fraction of alternative sigma factors associated with RNAP increases [91, 271].

The extracytoplasmic stress factor, σ^E, is activated during entry into stationary phase and this activation is dependent upon ppGpp, not the dedicated cell envelope stress-sensing pathway [82]. In this paper, we further investigate the correlation between σ^E activity and ppGpp levels, and the mechanism of regulation of σ^E-dependent transcription by ppGpp. In theory, the model that ppGpp alters the competition among sigma factors for RNAP could be sufficient to explain activation of σ^E by ppGpp. However, in vitro experiments demonstrate that ppGpp and DksA can also directly activate $E\sigma^E$-dependent transcription. Therefore, we propose that the positive regulation of σ^E-dependent promoters by ppGpp and DksA has two components: direct regulation of σ^E-dependent transcription and indirect regulation by increasing the amount of $E\sigma^E$ in the cell through negative regulation of transcription of rRNA promoters. This is the first reported example

of direct activation of alternative sigma factor-dependent transcription by ppGpp and DksA.

Results

ppGpp regulates σ^E activity

Previous work demonstrated that the activity of the alternative sigma factor, σ^E, was regulated with respect to growth phase [82]. In wild-type cells growing in rich media, σ^E activity can be divided into three parts. σ^E activity is low in early exponential phase ($OD_{600} < 0.3$), then increases 4-fold during mid-exponential phase (OD_{600} from 0.4 to 1.8) and 19-fold during entry into stationary phase ($OD_{600} > 1.8$) compared to its activity in early exponential phase. The low activity of σ^E during early exponential phase is not due to recovery from stationary phase and is likely to be a function of growth in fresh media [82]. The increase in activity during entry into stationary phase required the alarmone ppGpp, suggesting that ppGpp is a positive regulator of σ^E-dependent promoters. Because these experiments were performed with cultures grown in a rich medium, LB broth, and entry into stationary phase is complex in rich medium due to both depletion of nutrients and accumulation of secondary metabolites, all of which can affect gene regulation; it is possible that signals in addition to ppGpp are required for regulation of σ^E activity. To further explore the connection between ppGpp and σ^E activity, we examined σ^E activity under several conditions in which ppGpp levels are known to increase.

To determine if σ^E activity increased during entry into stationary phase under more defined conditions caused by the depletion of a specific nutrient, σ^E activity was monitored in cultures grown in media containing limiting concentrations of glucose (0.02%) or phosphate (0.13 mM). ppGpp levels are known to increase when cells enter stationary phase as either nutrient is depleted [84, 272]. Transcription by $E\sigma^E$ was assayed by

measuring β-galactosidase activity produced from a chromosomally encoded reporter in which the σ^E-dependent *rpoH*P3 promoter directs transcription of the *lacZ* gene [68]. This reporter contains the σ^E-dependent promoter from the *rpoH* gene and has been used extensively to monitor σ^E activity under a wide range of conditions in a variety of strain backgrounds. σ^E activity increased when cell growth slowed due to glucose or phosphate depletion, indicating that additional signals accompanying entry into stationary phase in rich media are not required for the regulation of σ^E activity (Fig. B-1A and B-1B). The increase in σ^E activity following depletion of a specific nutrient is not a property unique to the *rpoH*P3 promoter. The σ^E-dependent *fkpA* promoter, previously shown to be activated by ppGpp during entry into stationary phase in rich media [82], was activated during entry into stationary phase following phosphate depletion as well (data not shown).

The cellular level of ppGpp in exponential phase varies inversely with the growth rate of a culture, such that ppGpp levels are high in nutrient-poor media supporting slow growth rates, and low in nutrient-rich media supporting high growth rates. Many genes that are subject to regulation by ppGpp also show growth rate-dependent regulation [87, 272, 277]. We monitored transcription by $E\sigma^E$ during early exponential phase in cultures grown in media that support different growth rates. σ^E activity was regulated with respect to growth rate, increasing as the growth rate decreased (Fig. B-1C).

The above experiments demonstrate that under conditions in which ppGpp levels are known to increase, σ^E activity also increases. However, each of these conditions is accompanied by physiological adaptations to changes in nutrient availability, in addition to the production of ppGpp. This raises the possibility that $E\sigma^E$ could be sensing another signal that is coincident with nutrient depletion or culture conditions. If ppGpp were

sufficient to regulate transcription by $E\sigma^E$ by itself, then σ^E activity should increase when ppGpp is made during exponential phase in rich medium. To test this hypothesis a truncated variant of the *relA* gene that constitutively produces ppGpp without associating with ribosomes was overexpressed from the plasmid pALS13 [283]. When this variant *relA* gene was overexpressed, σ^E activity increased (Fig. B-1D). σ^E activity did not increase upon overexpression of a catalytically inactive variant of the truncated RelA protein that cannot synthesize ppGpp (Fig. B-1D). These results demonstrate that additional signals accompanying nutrient depletion are not required for induction of σ^E activity by ppGpp. Because cell growth slows when ppGpp production increases, we cannot formally eliminate the possibility that σ^E activity increases due to a separate event associated with the transition to slower growth.

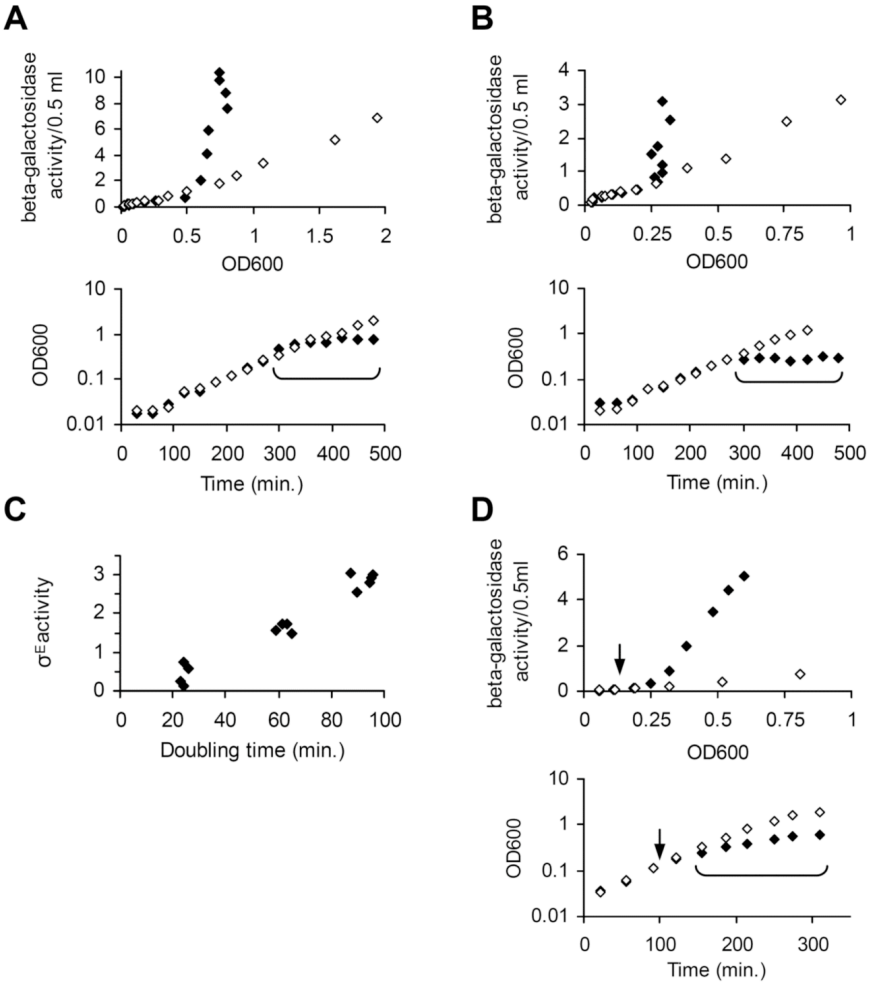


Figure B-1: σ^E activity increases under conditions in which ppGpp levels are known to increase. (A), (B), (D), differential rate plots (top) displaying the accumulation of β-galactosidase produced from the σ^E-dependent *rpoH*P3::*lacZ* reporter in SEA001 as a function of the growth of a culture (see Materials and Methods for an explanation of differential rate plots) and corresponding growth curves (bottom). β-galactosidase activity is the amount of *o*-nitrophenol formed, as measured by OD_{420}, divided by the reaction time for each 0.5 ml sample (see Materials and Methods). The samples with increased σ^E activity are noted with a bracket on the growth curves. (A-B), cultures were grown in MOPS minimal media (Teknova) with limiting concentrations of either phosphate or glucose: *A*, 0.13 mM phosphate (filled diamonds) or 1.3 mM phosphate (open diamonds), and *B*, 0.02% glucose (filled diamonds) or 0.2% glucose (open diamonds). (C) σ^E activity in strain SEA001 was measured in cultures grown in media supporting different growth rates (see Experimental Procedures for media composition). σ^E activity was determined in early exponential phase ($OD_{600} < 0.3$), before effects from increases in ppGpp levels in late exponential phase occurred. (D) σ^E activity was measured from the *rpoH*P3::*lacZ* reporter following gratuitous induction of ppGpp. IPTG was added (indicated by the arrow) to a

final concentration of 20 μM to induce active *relA*' from pALS13 in SEA2025 (filled diamonds) or the catalytically inactive form of *relA* from pALS14 in SEA2026 (open diamonds). The slight increase in σ^E activity seen in the strain with pALS14 reflects the growth phase-dependent increase of σ^E activity in mid-exponential phase [82]. In parts (A), (B), and (D), representative datasets are shown. Variation between datasets was less than 10%. In part (C), data from 4 independent experiments at each growth rate are shown.

If ppGpp is required for increasing σ^E activity, then eliminating the ability of RNAP to respond to ppGpp should eliminate the increase in sigma E activity observed in cells that can make ppGpp. Since both DksA and ω have been shown to affect the ability of ppGpp to function in positive and negative regulation in vitro and in vivo [278, 279], we constructed a Δ*dksA* Δ*rpoZ* strain and tested the $_{\text{effects}}$ on transcription by $E\sigma^E$. Growth phase-dependent regulation of σ^E activity was disrupted in the Δ*dksA* Δ*rpoZ* strain. Transcription by $E\sigma^E$ in this strain was low throughout the growth curve, increasing only slightly between exponential and stationary phase (Fig. B-2A, C).

σ^E activity was altered in the single-mutant Δ*rpoZ* strain, suggesting a role for ω in σ^E-dependent transcription. In wild-type cells, σ^E activity increased 4-fold from early exponential phase to mid-exponential phase, while in the Δ*rpoZ* strain σ^E activity only increased 2-fold at this point in the growth curve (Fig. B-2A, C). The regulator responsible for the increase in σ^E activity between early and mid-exponential phase has not been identified, but these results suggest that ω contributes to this regulation. This is one of the few phenotypes that have been detected from deletion of the *rpoZ* gene. It is possible that the effects of the *rpoZ* deletion are in part due to polar effects on the *spoT* gene immediately downstream. However, the *rpoZ* allele used in this study has only a weak effect on SpoT expression [284].

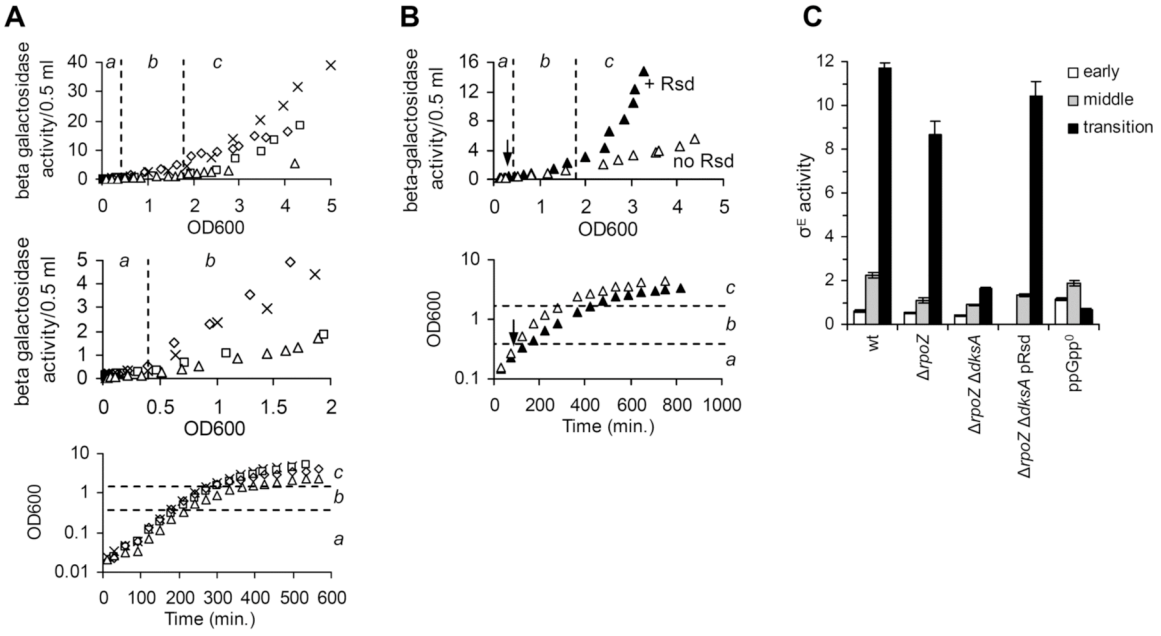


Figure B-2: Growth phase-dependent regulation of σ^E activity is shown for strains lacking proteins required for RNAP to respond to ppGpp. **(A)** Differential rate plots display the accumulation of β-galactosidase activity in 0.5 ml samples (*o*-nitrophenol/min) produced from the σ^E-dependent *rpoH*P3::*lacZ* reporter as a function of the growth of a culture for WT (crosses), Δ*dksA* (diamonds, SEA2051), Δ*rpoZ* (squares, SEA6017), and Δ*dksA*Δ*rpoZ* (triangles, SEA6028) strains. The top graph displays data for the entire growth curve, whereas the middle graph displays data only for exponential phase to show the lower activity in strains lacking *rpoZ* during mid-exponential phase. The corresponding growth curves are shown in the lower graph. Data from the Δ*dksA* mutant [82] are shown for comparison with the Δ*dksA*Δ*rpoZ* mutant. **(B)** Differential rate plot (top) and corresponding growth curve (bottom) displaying σ^E activity measured with the *rpoH*P3::*lacZ* reporter in Δ*dksA*Δ*rpoZ* strains with and without overexpression of *rsd*; pRsd + IPTG (closed triangles, SEA6145), empty vector + IPTG (open triangles, SEA6142), arrow indicates the time of IPTG addition. The portions of the plots in (A) and (B) corresponding to early exponential phase (a, $OD_{600} < 0.4$), mid-exponential phase (b, OD_{600} from 0.4 to 1.8), the transition into stationary phase and early stationary phase (c, $OD_{600} > 1.8$) are indicated on both the differential rate plots and growth curves. For *A* and *B*, representative datasets are shown and variation between datasets was less than 10%. **(C)** σ^E activity for each part of the growth curve, determined from the slope of the line on the differential rate plot, is shown. Slopes were determined using compiled data from at least two experiments, and error bars represent the standard error of the slope. No value is shown for σ^E activity in early exponential phase for the pRsd Δ*dksA*Δ*rpoZ* strain, since *rsd* overexpression was not induced until the end of this phase of growth.

ppGpp affects the activity, not production, of σ^E

How does ppGpp regulate σ^E activity? ppGpp could affect the production of σ^E, transcription by $E\sigma^E$, or both production and activity. If ppGpp activates transcription by $E\sigma^E$ by increasing the production of σ^E itself, then the level of σ^E in the cell during entry into stationary phase should increase concomitantly with the increase in $E\sigma^E$ activity, and the increase in σ^E levels should be dependent on ppGpp. To address this issue, σ^E levels were examined by Western blotting in wild-type strains and in strains unable to produce ppGpp ($ppGpp^o$) due to deletion of the *relA* and *spoT* genes. σ^E levels were similar in both the wild-type and $ppGpp^o$ strains (Fig. B-3). Furthermore, the levels of σ^E did not increase following gratuitous production of ppGpp due to overexpression of the constitutive *relA* variant from pALS13 (data not shown). Therefore the ppGpp-dependent increase in σ^E activity is not caused by an increase in the overall amount of σ^E in the cell.

To further demonstrate that ppGpp regulates transcription by $E\sigma^E$ and not the transcription of the *rpoE* gene, we constructed a strain in which the chromosomal copy of the *rpoE* operon (*rpoE*, *rseA*, *rseB*, and *rseC*) was deleted and the *rpoE* gene was expressed from a plasmid under the control of the σ^{70}-dependent pTrc promoter. The pTrc promoter is repressed by the Lac repressor, but is somewhat leaky. The uninduced level of σ^E expression was sufficient to maintain cell viability (*rpoE* is essential). σ^E activity was higher in this strain than in a wild-type strain, because the anti-sigma factor, *rseA*, was deleted as part of the *rpoE* operon. However, σ^E activity was still regulated with respect to growth phase and changes in σ^E levels cannot account for this regulation, indicating that transcriptional regulation of the *rpoE* operon promoter is not required for the growth phase-dependent increase in transcription by $E\sigma^E$ (data not shown).

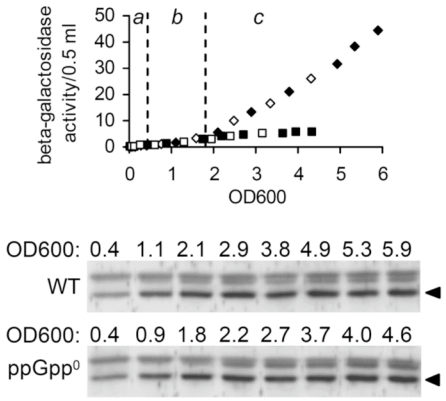


Figure B-3: ppGpp does not affect the steady state level of σ^E. Differential rate plot displaying σ^E activity measured with the *rpoH*P3::*lacZ* reporter in wild-type (diamonds, SEA001) and ppGpp^o (squares, SEA2010) strains as a function of the growth of the culture is shown. At the time points indicated by filled symbols, samples were removed and used for the Western blotting analysis to determine the steady state level of σ^E. The portions of the plots corresponding to early exponential phase (a), mid-exponential phase (b), and the transition into stationary phase and early stationary phase (c) are indicated as described for Fig. A-2. Western blots probed with anti-σ^E antibody are displayed below. Equal amounts of protein extracts were loaded in each lane. The σ^E band is indicated with an arrowhead. A cross-reacting band, which is present in strains lacking the *rpoE* gene, runs directly above the σ^E band. A representative blot is shown and similar results were obtained in three separate experiments.

Regulation of σ^E-dependent transcription by ppGpp

Since ppGpp does not affect the level of σ^E, it must alter its activity. ppGpp binds to core RNAP and could directly affect transcription by Eσ^E with or without the assistance of DksA. Alternatively, ppGpp could indirectly affect σ^E-dependent transcription. Indirect regulation could be achieved by several mechanisms. For example, ppGpp could regulate the expression of a coactivator protein or molecule that is required specifically for transcription by Eσ^E. ppGpp could alter the competition among sigma factors, including σ^E, for core RNAP, as suggested by studies on the regulation of transcription by the alternative sigma factors σ^S and σ^N by ppGpp [271].

The model that ppGpp alters the competition among sigma factors for RNAP is supported by the observation that the *rpoD*P504L and *rpoD*S506F mutations, which lower the affinity of σ^{70} for core RNAP [285-287], restore σ^{N} activity during entry into stationary phase in cells lacking ppGpp [87, 88]. These mutations were originally isolated as suppressors of growth defects of ppGpp^{o} strains on media lacking amino acids, suggesting that they also suppress defects in Eσ^{70} activity in ppGpp^{o} strains [285]. To determine whether these σ^{70} variants could suppress the defect in σ^{E} activity in strains lacking ppGpp, transcription by Eσ^{E} was monitored in ppGpp^{o} strains carrying these mutations. In the ppGpp^{o} strains with either the *rpoD*S506F or *rpoD*P504L alleles, not only was transcription by Eσ^{E} restored during entry into stationary phase, but it was also nearly constitutive throughout the growth curve. σ^{E} activity increased by early exponential phase to a level comparable to that observed during entry into stationary phase in the wild-type strain (Fig. B-4A, D and E). Similar results were obtained with the σ^{E}-dependent *fkpA* promoter in *rpoD*S506F ppGpp^{o} and *rpoD*P504L ppGpp^{o} strains (data not shown). These results indicate that ppGpp itself is not absolutely required for σ^{E} activity. The σ^{70} variants not only suppress the requirement of ppGpp for σ^{E} activity, but also nearly eliminate the growth phase-dependent regulation of transcription by Eσ^{E} throughout the growth curve.

The observation that mutations in σ^{70} that reduce its affinity for core RNAP render σ^{E} activity nearly constitutive, suggests that σ^{E} activity is higher because it can compete better against σ^{70} for binding to core RNAP, and ppGpp is no longer needed. Therefore, other mechanisms to reduce the ability of σ^{70} to bind to RNAP should also restore σ^{E} activity in a strain lacking ppGpp. Rsd is a σ^{70}-specific anti-sigma factor that binds to σ^{70} and blocks its association with core RNAP [288]. To determine if overexpression of *rsd*

could also restore σ^E activity in a strain lacking ppGpp, transcription by $E\sigma^E$ was monitored in a ppGpp^o strain carrying a plasmid with the *rsd* gene placed under the control of an IPTG-inducible promoter. Overexpression of *rsd* restored the growth phase-dependent increase in σ^E activity during entry into stationary phase (Fig. A-4B, D, and E). These results provide further evidence that ppGpp itself is not absolutely required for this activation. Rsd overproduction will also restore σ^S and σ^N activity in a ppGpp^o background, indicating that altering the competition among sigma factors for core RNAP by disabling σ^{70} is a general mechanism to restore alternative sigma factor activity in the absence of ppGpp [88, 91].

We demonstrated that the growth phase-dependent regulation of transcription by $E\sigma^E$ is disrupted in a Δ*dksA*Δ*rpoZ* strain (Fig. B-2A, C), presumably because RNAP cannot respond to ppGpp. Since overexpression of *rsd* suppressed the requirement for ppGpp, it should also suppress the defects in σ^E activity in the Δ*dksA*Δ*rpoZ* strain, unless there are additional defects unrelated to ppGpp that affect σ^E activity in this strain. Overexpression of *rsd* did restore σ^E activity during entry into stationary phase (Fig. B-2B and C), providing further evidence that at least some aspects of the Δ*dksA*Δ*rpoZ* strain phenocopy a ppGpp^o strain.

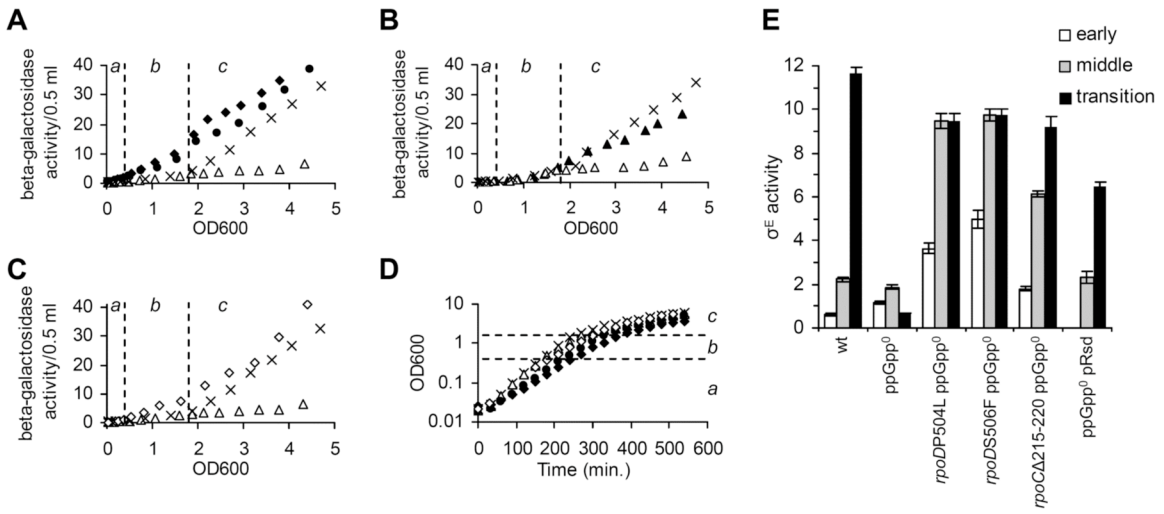


Figure A-4: Suppression of the requirement for ppGpp for the increase in σ^E activity during entry into stationary phase by **(A)** mutations in *rpoD*, **(B)** overexpression of *rsd*, and **(C)** a mutation in *rpoC*. (A-C) are differential rate plots displaying the accumulation of β-galactosidase activity in 0.5 ml samples (*o*-nitrophenol/min) produced from the σ^E-dependent *rpoH*P3::*lacZ* reporter as a function of growth and **(D)** shows growth curves for the strains. Each panel displays data for the wild-type strain (crosses, SEA001), the ppGpp^o (open triangles, SEA2010) strain, and the following strains: (A and D), ppGpp^o strain with the σ^{70} variants *rpoD*P504L (filled circles, SEA2027) and *rpoD*S506F (filled diamonds, SEA2028); (B and D), ppGpp^o with the pRsd plasmid (filled triangles, SEA2144), and (C and D) ppGpp^o strain with the β' variant *rpoC*(Δ215-233) (open diamonds, SEA2136) . In *B*, *rsd* overexpression was induced at an OD_{600}=0.2. The portions of the plots corresponding to early exponential phase (a), mid-exponential phase (b), and the transition into stationary phase and early stationary phase (c) are indicated as described for Fig. A-2. Representative datasets are shown in parts A-D and variation between datasets was less than 10%. **(E)** σ^E activity for each part of the growth curve, determined from the slope of the line on the differential rate plot, is shown. Slopes were determined using compiled data from at least three experiments, and error bars represent the standard error of the slope. For the *rpoD* mutant strains, σ^E activity increased before mid-exponential phase and remained high (see text), as such σ^E activity in early exponential phase was calculated before this increase. No value is shown for σ^E activity in early exponential phase for the pRsd ppGpp^o strain since *rsd* overexpression was not induced until the end of this phase of growth.

The expression of *rsd* itself is regulated by ppGpp, and increases during entry into stationary phase [288]. This observation, along with the results that overexpression of *rsd* can restore the growth phase-dependent regulation of σ^E activity, suggested that ppGpp

may indirectly affect σ^E activity by increasing the expression of *rsd*. If this were true, then the pattern of $E\sigma^E$ activity during the growth of cells lacking *rsd* should resemble that of cells lacking ppGpp, i.e. no activation during entry into stationary phase. However, no defects in the growth phase-dependent regulation of transcription by $E\sigma^E$ were observed in a Δ*rsd* strain (data not shown). Therefore either *rsd* is not required for ppGpp to act, or the regulation is redundant and another regulator controls σ^E activity in the absence of *rsd*.

The above results demonstrate that the requirement for ppGpp to increase σ^E activity can be complemented by conditions in the cell that decrease the ability of σ^{70} to compete for core RNAP. Many mutations that suppress the growth defects in ppGpp-deficient cells have also been mapped to the β and β' subunits of RNAP [289, 290]. Several of these mutations will also complement defects in σ^N activity in ppGpp^o strains [87]. We examined the ability of one such mutation, a deletion of amino acids 215-220 of the β' subunit, to restore $E\sigma^E$ activity in a ppGpp deficient strain [289]. This deletion in β' is not at the sigma/core interface and is unlikely to affect the affinity of σ^{70} for core RNAP. The increase in σ^E-dependent activity during entry into stationary phase was restored in the ppGpp^o strain containing the *rpoC*Δ215-220 mutation, indicating that alterations in β' can also complement the defects in σ^E activity observed when ppGpp is lacking (Fig. B-4C, D, and E). σ^E activity was also elevated during mid-exponential phase in this strain, ~3-fold compared to the wild-type strain, but not to the same extent as in the strains with mutations in *rpoD* (Fig. B-4E). Again, similar results were obtained from the *fkpA* promoter in the ppGpp^o *rpoC*Δ215-220 strain (data not shown). This mutation in β' most likely restores σ^E activity indirectly via its destabilizing effects on ppGpp-sensitive σ^{70}-dependent rRNA promoter complexes, thereby increasing the availability of core RNAP for binding to σ^E

(see Discussion). We have not formally ruled out that this mutation could also directly favor transcription initiation by $E\sigma^E$, thus compensating for the loss of activation in the absence of ppGpp.

To explore further the mechanism(s) responsible for activation of σ^E activity, direct effects of ppGpp and DksA were tested on promoter complexes formed by $E\sigma^E$. ppGpp and DksA destabilize RNAP complexes formed on all promoters examined to date [85, 86, 275, 281, 291, 292], but their effects on $E\sigma^E$ promoter complexes have not been tested previously. Complexes were formed on two promoters, *rpoH* P3, the same $E\sigma^E$-dependent promoter used in the *in vivo* experiments described above, and *rybB*, the promoter for a σ^E-dependent small RNA whose expression is also regulated by ppGpp/DksA [262]. The fraction of complexes in the presence of ppGpp, DksA, or the two together that remained at different times after addition of the competitor heparin was measured by in vitro transcription [292]. Both promoters formed relatively stable complexes with $E\sigma^E$ under these solution conditions ($t_{½}$ =72 min for *rpoH*P3 and $t_{½}$ =26 min for *rybB*; Fig. B-5A). DksA by itself had little effect on the lifetimes of competitor-resistant complexes at either promoter, while ppGpp by itself destabilized complex stability by approximately 2.5-fold (Fig. B-5A). In contrast, DksA and ppGpp together greatly decreased the lifetimes of the promoter complexes, by 6-fold for the *rybB* promoter and 15-fold for the *rpoH*P3 promoter (Fig. B-5A). Future studies will be needed to address why DksA alone had little effect on the half-lives of these promoters under these conditions; nevertheless, these results show that ppGpp and DksA can together function on promoter complexes containing $E\sigma^E$ RNAP.

Although ppGpp and DksA destabilize open complexes on all promoters yet examined, their effect on overall transcription depends on the properties of the individual promoters [85, 86, 275, 281, 291, 292]. Therefore, effects of ppGpp and DksA on transcriptional output were measured by multi-round transcription from *rpoH*P3 (representative assays are shown at the left in Fig. B-5B, and the results from multiple assays are quantified at the right in Fig. B-5B). As observed previously with certain σ^{70}-dependent amino acid biosynthesis promoters [85], ppGpp by itself exerted little or no effect (Fig. B-5B), and as also observed previously, DksA by itself activated transcription up to ~2-fold (Fig. A-5B). When 200 μM ppGpp was also included in the reactions, transcription increased slightly more (up ~3-fold) and was observed at lower concentrations of DksA (Fig. A-5B). Similar effects were obtained on transcription from the *rybB* promoter, although the magnitude of the activation was smaller. In preliminary experiments DksA alone (2 μM) increased transcription from P*rybB* approximately 1.5-fold and DksA/ppGpp (2 μM/100 μM, respectively) increased transcription from P*rybB* 1.8-fold, while ppGpp alone had no effect (data not shown). We conclude that ppGpp and DksA can directly activate transcription by $E\sigma^{E}$, and, as with $E\sigma^{70}$-dependent promoters [85], different promoters are activated to different extents.

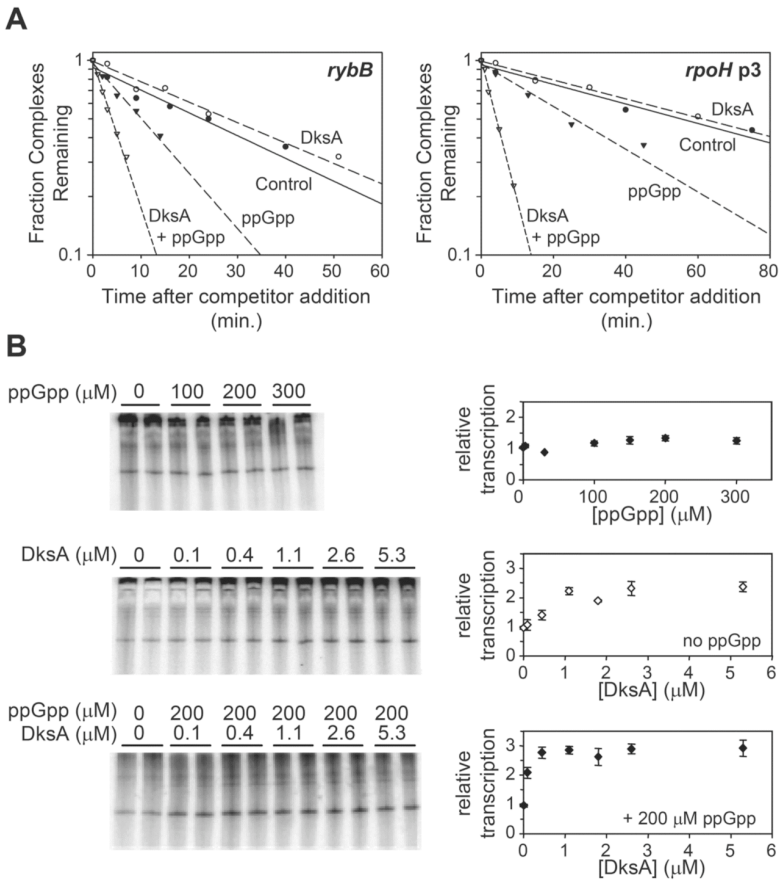


Figure B-5: Effects of ppGpp and DksA on $E\sigma^E$-dependent transcription. **(A)** DksA and ppGpp together, as well as ppGpp alone, increase the RNAP-promoter open complex decay rate in vitro at the $E\sigma^E$-specific *rpoH*P3 and *rybB* promoters. Lifetimes of the competitor-resistant open complex were measured using a transcription-based assay under the following conditions: 2 μM His-DksA, 100 μM ppGpp, 10 nM $E\sigma^E$ and 1 nM supercoiled plasmid template (pSEB014 or pSEB015) with 10 μg/ml heparin as a competitor (see Materials and Methods for additional details). Representative decay curves are shown; the absolute rates differed over three assays but the ratios with and without DksA and ppGpp were reproducible (≤ 15% variation). **(B)** DksA and ppGpp together, as well as DksA alone, increase $E\sigma^E$-dependent transcription in multi-round transcription assays from the *rpoH*P3 promoter. To initiate transcription $E\sigma^E$ was added to a final concentration of 20 nM to reaction mixes containing 2.5 nM supercoiled plasmid template carrying the *rpoH*P3 promoter (pSEB015), nucleotides, and the indicated amounts of ppGpp and/or His-DksA (see Materials and Methods for additional details). Representative gels are shown on the left and quantitation of the data on the right. The fold increase in transcription represents the amount of transcript in reactions with ppGpp and/or DksA relative to that in a reaction without ppGpp and DksA. The data were compiled from a minimum of three experiments, average values are shown, and error bars represent the standard deviation.

Discussion

The data presented here are consistent with the model that ppGpp works as a general alarm system to redistribute RNAP from promoters for genes required for rapid growth to promoters for genes that are important for survival during stress, shifting the spectrum of genes transcribed by $E\sigma^{70}$ and activating genes transcribed by alternative sigma factors [82, 271, 293, 294]. Our data also support the model that ppGpp, along with its cofactor DksA, changes the distribution of RNAP holoenzymes on promoters, both through direct effects on promoter complexes sensitive to ppGpp and DksA and also through indirect effects achieved primarily by reducing transcription of rRNA genes [85, 91, 282, 290, 292]. In this manner, ppGpp and DksA provide a general mechanism to coordinately activate alternative sigma factors.

ppGpp and the regulation of σ^E expression

In addition to coordinately regulating alternative sigma factors by altering the distribution of RNAP holoenzymes, ppGpp and DksA exert specific effects on individual alternative sigma factors. ppGpp is thought to increase the production of σ^S, but not of σ^N [87, 90, 280]. We found that the production of σ^E is not dependent on ppGpp. Although one of the two promoters that directs transcription of the *rpoE* operon, *rpoE*P2, is dependent on σ^E [70], not all promoters transcribed by $E\sigma^E$ are necessarily subject to control by ppGpp/DksA. Only specific σ^{70}-dependent promoters have the kinetic characteristics that make them sensitive to ppGpp/DksA, and this is likely to be true of σ^E-dependent promoters as well (see below). Alternatively, since the *rpoE* gene is transcribed from two major promoters, one dependent on σ^{70} and the other dependent on σ^E, increased expression from the σ^{70}-dependent *rpoE*P1 promoter might compensate for decreased

expression from *rpoE*P2 in the absence of ppGpp. Conversely, if transcription from *rpoE*P2 increases in the presence of ppGpp, transcription from *rpoE*P1 may decrease, thereby maintaining a constant amount of σ^E in the cell.

Direct regulation of σ^E-dependent transcription by ppGpp

ppGpp and DksA can have positive, negative or no effects on overall transcription, even though these factors reduce the stability of open complexes on all promoters tested to date, including the two σ^E-dependent promoters tested here. This observation can be explained by the fact that ppGpp/DksA affect transcription from only a subset of promoters, those with specific kinetic characteristics. Studies with *rrnB*P1 promoter variants strongly suggest that ppGpp/DksA inhibit transcription by further destabilizing promoter complexes that are intrinsically short-lived. Decay of the open complex formed on these promoters is rate-limiting for transcription initiation [295]. In contrast, promoters that form stable open complexes are not inhibited by ppGpp/DksA, even though the open complexes are destabilized, because dissociation of the open complex is not rate-limiting for transcription initiation [295]. Consistent with this model, the open complexes formed by $E\sigma^E$ are relatively long-lived and ppGpp/DksA do not inhibit transcription from these promoters. It was somewhat surprising that effects of ppGpp alone but not DksA alone were observed on the lifetime of the open complex, while effects of DksA alone but not ppGpp alone were observed in multi-round transcription assays. However, we emphasize that ppGpp/DksA together had parallel effects in the two assays, and this is the condition most relevant to cells.

The mechanism of positive control of transcription by ppGpp/DksA is much less well understood than negative control. Experiments with the σ^{70}-dependent P*argI* promoter

suggest that ppGpp/DksA activate transcription by increasing the isomerization rate from the closed to open complex [85]. A model based on the results with P*argI* and other amino acid biosynthetic promoters has been proposed explaining how DksA/ppGpp could reduce the free energy of a transition state on the pathway to open complex formation, and this could result in positive effects on transcription [85]. The kinetics of transcription initiation for $E\sigma^E$ have not yet been studied in detail. In fact these studies present the first measurements of open complex stability for $E\sigma^E$ on any promoter. Studies on the mechanics of transcription initiation at $E\sigma^E$-dependent promoters will be needed to elucidate how pGpp/DksA acts on these complexes.

Indirect regulation of alternative sigma factor transcription by ppGpp

$E\sigma^E$ activity increased 5-fold during entry into stationary phase in the in vivo reporter assays with the *rpoH*P3 promoter. This increase was greater than that observed by ppGpp and DksA in vitro. Although the solution conditions may not adequately mimic the conditions in vivo, a likely explanation for the disparity between the magnitude of the observed effects in vivo and in vitro is that a component of the activation of σ^E-dependent promoters results from indirect effects of ppGpp and DksA. Previous models proposed to explain how ppGpp positively regulated transcription by $E\sigma^S$ and $E\sigma^N$ invoked alteration of the competition among sigma factors for RNAP [91, 271, 282, 285, 290].

Several hypotheses can explain how ppGpp alters the competition among sigma factors for RNAP. The observation that both mutations in *rpoD*, which reduce the affinity for core RNAP, and overexpression of the σ^{70}-inhibitor *rsd* bypass the requirement of ppGpp for $E\sigma^E$, $E\sigma^S$, and $E\sigma^N$ activity in vivo suggest that ppGpp could affect the association of sigma factors with RNAP favoring alternative sigma factors [87, 88, 91,

285, 286]. However, to our knowledge it has not been demonstrated that ppGpp and/or DksA directly alters the affinity of any sigma factor for core RNAP. The model also does not explain suppression of the defects in σ^E activity by the *rpoC* variant. The *rpoC* mutation, a deletion of amino acids 215-220, is not at the sigma/core interface and therefore is less likely than the *rpoD* mutants to alter the affinity of σ^{70} for core RNAP.

An alternative explanation is that ppGpp alters the competition by increasing the amount of free RNAP available to bind to all sigma factors via its effects on transcription of ribosomal RNAs. In rapidly growing cells, up to 70% of the RNAs transcribed in the cell are stable rRNAs encoded in long operons [296]. When ppGpp levels increase, transcription of these operons decreases and the core RNAP that was actively engaged in transcription will be released upon transcription termination [291, 296]. This release of core RNAP will increase the size of the free pool of the enzyme available to bind all sigma factors. This model is consistent with phenotypes of the *rpoC* mutation and with additional properties of the *rpoD* mutations, which appear to functionally mimic the effect of ppGpp/DksA on transcription initiation from the rRNA promoters [88, 282, 289, 292]. RNAP containing the β'Δ215-220 deletion reduced transcription from the *rrnB*P1 promoters, and the open complexes formed by the variant holoenzyme on this promoter are extremely unstable [289, 292]. The σ^{70} variants, in addition to having a lower affinity for core RNAP than WT σ^{70}, also further destabilize competitor-resistant open complexes formed on the *rrnB*P1 promoter in vitro [88, 282].

The combination of direct and indirect effects of ppGpp on gene expression provide a powerful means of transcriptional regulation. On a global level, ppGpp coordinately alters the activity of individual sigma factors redistributing RNA polymerase

among the sigma factors in the cell. This general regulatory mechanism works in conjunction with the ppGpp-independent signaling pathways that are specific for each alternative sigma factor and determine the overall amounts of the different alternative holoenzymes available to respond to changing ppGpp levels. For example, the overall amount of free σ^E available to interact with core RNAP is determined primarily by the amount and proteolytic stability of the anti-sigma factor RseA [75, 82]. Finally, the unique thermodynamic and kinetic properties of individual promoters ultimately determine the extent to which the direct effects of ppGpp on transcription and the changing amounts of alternative holoenzymes modulate gene expression.

Materials and Methods

Strains and growth conditions

Strains used in this work are listed in Table B-1. Mutant alleles were moved into the appropriate strains using P1 transductions according to standard techniques [177]. The *rpoD* and *rpoC* mutant alleles with tightly linked *zgh*-3075::Tn*10* (*rpoD* alleles) or *thi39*::Tn*10* (*rpoC* allele) insertions [285, 289] were transferred into strain SEA2010 by P1 transduction. Transductants were selected on tetracycline and the presence of the correct mutation was verified by sequencing. Experiments with the *rpoD*, *rpoC*, and *dksA rpoZ* mutant strains were performed with at least two independent transductants to ensure that the results were not affected by spontaneous suppressor mutations. All ppGpp^o strains were verified as being unable to grow on minimal media lacking amino acids, and the *rpoD* and *rpoC* suppressor strains were verified to have reverted this auxotrophy. Strain SEA2023 was made by a targeted disruption of the *lacYA* genes in the Φλ*rpoH*P3::*lacZ* reporter of SEA001. The genes were deleted according to the procedure of Datsenko and Wanner and the drug marker removed by FLP recombinase [169]. Strain SEA2043 was made by transformation of SEA001 with pLC245 expressing the *rpoE* gene [48] followed by P1 transduction of the Δ*rpoE-rseC::kan* allele. Strains were grown in Luria Bertani broth at 30°C with aeration unless otherwise noted.

Plasmid constructions

Plasmids used in this study are listed in Table B-1. The plasmid pRsd was constructed by PCR amplifying the *rsd* gene from the chromosome and cloning it into the expression plasmid pTrc99a. Plasmids used as transcription templates are derivatives of pRLG770 [207], containing the σ^E-dependent *rpoH*P3 (pSEB015) or *rybB* (pSEB014)

promoters. The *rpoH*P3 promoter was amplified with the primers 5'-GGGCCGGAATTCGCCTTGATGTTACCCGAGAG-3' and 5'-GGGCCAGGTGGAGACCCCCAGTCACGACGTTGTAAA-3' from plasmid p2*rpoH*P3, which has the isolated σ^E-dependent *rpoH*P3 promoter cloned into the PstI-HindIII sites of pUC19. The *rpoH*P3 promoter sequence, from the -35 region to the +1 for transcription, is identical to that in the chromosomal *lacZ* fusion (λ*rpoH*P3::*lacZ*) used to study σ^E activity [68]. The resulting PCR product was digested with EcoRI and BsaI then cloned into the EcoRI and HindIII of pRLG770 to make pSEB015. The *rybB* promoter region and *rybB* gene, including the native transcription terminator, were amplified from genomic DNA by PCR with the primers: 5'-GGGCGGGAATTCGTTGTTCGGCGCAATGAT-3' and 5'-GGGCCAAGCTTGTTGAGAGGGTTGCAGGGTA-3' and cloned into the EcoRI and HindIII of pRLG770 to make pSEB014.

Western Blotting

Whole-cell extracts were prepared as described [90]. Briefly, cells were lysed in protein sample buffer, proteins were precipitated with acetone, and resuspended in 2% SDS. Protein concentrations were determined using the BCA Protein Assay (Pierce). 10 μg of total protein from each sample were loaded onto 12% polyacrylamide-SDS gels and transferred to Hybond^{TM}-P, PVDF membrane (GE-Healthcare). Bands containing σ^E were detected by probing the blots with rabbit polyclonal antibodies raised against σ^E (gift from CA Gross) then with alkaline phosphatase conjugated goat anti-rabbit IgG (Vector Laboratories, Inc. Burlingame, CA). The secondary antibody was visualized with the ECF

reagent from GE Healthcare according to the manufacturer's instructions. Blots were scanned using Typhoon 8600 Imager in fluorescence mode.

β-galactosidase assays

Overnight cultures were diluted to an OD_{600} of 0.02 and grown with shaking in a gyratory water bath at 30 °C. Samples (0.5 ml) were collected throughout the growth curve. The β-galactosidase activity of each sample was measured by the standard assay [177] and is expressed as the OD_{420} of the reaction mixture divided by the reaction time (*o*-nitrophenol/min). The data are presented as differential rate plots in which β-galactosidase activity in each 0.5 ml sample is plotted versus the optical density (OD_{600}) of the sample. β-galactosidase activity (*o*-nitrophenol/min) per 0.5 ml sample is plotted, rather than standard Miller Units (*o*-nitrophenol/min/OD_{600}), therefore the slope of the curve at each time point indicates the change in β-galactosidase activity with increased cell number. The plots illustrate how σ^E activity changes throughout the growth curve, in recovery from stationary phase, exponential phase, and re-entry into stationary phase. A complete explanation of differential rate plots is presented in [82]. Experiments were repeated a minimum of three times with independent cultures.

For measurements of σ^E activity as a function of growth rate, cultures were grown in MOPS minimal media (Teknova) with 0.4% glycerol, MOPS minimal media (Teknova) with 0.2% glucose, EZ rich media (Teknova) and LB at 37 °C. σ^E activity was determined in early exponential phase to avoid interference from any additional regulation due to changing ppGpp levels during entry into stationary phase. Activity was determined by the slope of the line on a differential rate plot of β-galactosidase activity in 0.5 ml of culture (as described above) vs. OD_{600}. Experiments were repeated a minimum of three times to

ensure reproducibility. σ^E activity in early exponential phase reflects growth in fresh media at a low optical density. σ^E activity is the same whether it is measured in cultures directly following dilution of a saturated overnight or following repeated dilution of cultures that had reached exponential phase (OD_{600} ~ 0.3) to ensure that the cells are in so-called steady state growth ([82], and data not shown).

Protein Purification

N-terminally His-tagged σ^E was purified from strain BL21(DE3) *slyD*::*kan* p*LysS* pPER76 as previously described [189]. Briefly, cells were grown at 25°C to an OD_{600} of 0.5 at which point IPTG was added to induce protein production. Following 1.5 hours of induction, cells were harvested by centrifugation, resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 2.5 mM β-mercaptoethanol, 1 mM PMSF). Cells were lysed by sonication and the lysate was cleared by centrifugation. The supernatant containing soluble His-σ^E was loaded onto a NiNTA column. Bound proteins were eluted with a stepwise gradient of 20, 60, 100, and 200 mM imidazole in column buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2.5 mM β –mercaptoethanol). Fractions containing σ^E were pooled and dialyzed into 20 mM Tris-HCl pH 8.0, 50 mM NaCl, 2.5 mM β-mercaptoethanol.

His-tagged DksA was purified as described in [86] or with minor modifications. Briefly, cells were grown at 37°C to an OD_{600} of 0.4 at which point IPTG was added to induce protein production. Following three hours of induction, cells were harvested by centrifugation, resuspended in buffer 1 (50 mM TrisHCl pH 8.0, 300 mM NaCl, 5 mM imidazole, 5 mM β-mercaptoethanol, 0.1 mM PMSF) and lysed by sonication. The lysate was cleared by centrifugation and the supernatant loaded onto a NiNta column. Bound

proteins were eluted with a stepwise gradient of buffer 1 with 75, 150, and 300 mM imidazole. Fractions containing His-DksA were combined and dialyzed into buffer 2 (20 mM Tris-HCl pH8.0, 100 mM NaCl, 0.1 mM DTT) then loaded onto a Hi-Trap^{TM} Q FF column (GE Healthcare). Bound proteins were eluted with a stepwise gradient of buffer 2 with 200, 300, and 400 mM NaCl. Fractions containing His-DksA were pooled and dialyzed into Buffer 2 with 20% glycerol.

RNAP-Promoter Complex Decay Assays

Lifetime of the competitor-resistant open complex was measured by single-round in vitro transcriptional assays as described in [291] with the exception of the use of Eσ^E (reconstituted at 30 °C from 1:2 ratio of native core RNAP: His-σ^E). Briefly, 10 nM Eσ^E and 1 nM supercoiled plasmid DNA in transcription buffer (40 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT, 0.1 mg/ml BSA), and ppGpp and DksA (as indicated in the figures) were incubated for 10 min. at 30 °C. Heparin was added to a final concentration of 10 µg/ml, and aliquots were removed to tubes containing NTPs (500 µM ATP, 200 µM GTP, 200 µM CTP, 10 µM UTP, and 1.0 µCi [α^{32}-P]UTP) at various times after heparin addition. Transcription reactions were stopped after 10 min. with an equal volume of urea-based gel loading buffer. Transcripts were separated on a 6% polyacrylamide gel containing 7 M urea and then visualized and quantitated by phosphorimaging using ImageQuant software (Molecular Dynamics).

Multi-round transcription assays

Eσ^E holoenzyme was formed by incubating 200 nM core RNAP (Epicentre) with 800 nM His-σ^E for 10 minutes at 30°C. Multi-round transcription reactions were initiated by addition of Eσ^E, to a final concentration of 80 nM σ^E and 20 nM core RNAP, to

prewarmed (30 °C) transcription mix containing 2.5 nM supercoiled plasmid template, 5% glycerol, transcription buffer, 500 μM ATP, 200 μM CTP, 200 μM GTP, 10 μM UTP, 2.5 μCi [α^{32}-P]UTP, and the appropriate concentrations of ppGpp and/or His-DksA. After 10 minutes at 30 °C, reactions were stopped by the addition of stop solution (80% formamide, 20 mM EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue). Samples were electrophoresed on 6% polyacrylamide gels containing 7.5 M urea, visualized by phosphorimaging, and quantified using ImageQuant software (Molecular Dynamics).

TABLE B-1: Strains and Plasmids

Strain/Plasmid	Genotype	Source, Reference, P1 Donor Strain
Strains		
RLG7505	BL21(DE3) p RLG7067	[86]
SEA001	MG1655 Φλ*rpoH*P3::*lacZ* Δ*lacX74*	[82]
SEA2010	SEA001 Δ*relA*251::*kan*, Δ*spoT*207::*cam*	[82]
SEA2023	SEA001 Φλ*rpoH*P3::*lacZ ΔlacYA*	this work
SEA2025	SEA2023 pALS13	this work
SEA2026	SEA2023 pALS14	this work
SEA2027	SEA2010 *rpoD*(P504L) *zgh-3075*::Tn*10*	this work, P1 donor *rpoD*5 [285]
SEA2028	SEA2010 *rpoD*(S506F) *zgh-3075*::Tn*10*	this work, P1 donor *rpoD*11 [285]
SEA2043	SEA001 Δ*rpoE*::*kan* pLC245	this work
SEA2051	SEA001 Δ*dksA*::*tet*	[82]
SEA2103	SEA001 Δ*rsd*::*tet*	this work
SEA2136	SEA2010 *rpoC*(Δ215-220) *thi39*::Tn*10*	this work, P1 donor RLG3381 [289]
SEA2144	SEA001 pRsd	this work
SEA5036	BL21(DE3) *slyD*::*kan* p*LysS* pPER76	this work, P1 donor BB101 [297]
SEA6017	SEA001 *rpoZ*::*cam* (Δ*spoS3*::*cam*)	this work, P1 donor CF2790 [284]
SEA6028	SEA001 *dksA*::*tet rpoZ*::*cam* (Δ*spoS3*::*cam*)	this work, P1 donors SEA2027 and CF270 [284]
SEA6145	SEA6028 pRsd	this work
SEA6142	SEA6028 pTrc99a	this work
Plasmids		
pALS13	P*tac* truncated *relA*, active protein, Ap^R	[283]
pALS14	P*tac* truncated *relA*, inactive protein, Ap^R	[283]
pRLG7067	pET28a-His-*dksA*, Kan^R	[86]
pLC245	*rpoE* in pTrc99a, Ap^R	[48]
pPER76	*rpoE* in T7 expression vector pET15b, Kan^R	[70]
pRLG770	General transcription vector, Ap^R	[207]
pRsd	*rsd* in pTrc99a, Ap^R	this work
pSEB014	*rybB* promoter and gene in pRLG770, Ap^R	this work
pSEB015	isolated *rpoH*P3 promoter in pRLG770, Ap^R	this work
pTrc99a	vector, pBR322 ori, Ap^R	Pharmacia

Acknowledgements

We thank Mike Cashel, Carol Gross, Jim Hernandez, and Karen Wassarman for sharing strains, Steve Rutherford for preparing the σ^E used in some experiments, numerous discussions, and other contributions, and Pui-Wen Cheung for technical assistance. This work was supported by grant MCB-0347302 from the National Science Foundation to S.E.A., grant R37-GM37048 from the National Institutes of General Medical Sciences to R.L.G., an NIH Graduate Research Fellowship to A.B.B., and a National Science Foundation Graduate Research Fellowship to S.E.B.

Appendix C

Regulated Proteolysis: Control of the *Escherichia coli* σ^E-dependent Cell Envelope Stress

Abstract

Over the past decade, regulatory proteolysis has emerged as a paradigm for transmembrane signal transduction in all organisms, from bacteria to man. These conserved proteolytic pathways share a common design that involves the sequential proteolysis of a membrane-bound regulatory protein by two proteases. Proteolysis releases the regulator, which is inactive in its membrane-bound form, into the cytoplasm where it performs its cellular function. One of the best-characterized examples of signal transduction via regulatory proteolysis is the pathway governing the σ^E-dependent cell envelope stress response in *Escherichia coli*. In unstressed cells, σ^E is sequestered at the membrane by the transmembrane anti-sigma factor, RseA. Stresses that compromise the cell envelope and interfere with the proper folding of outer membrane porins (OMPs) activate the proteolytic pathway. The C-terminal residues of unfolded OMPs bind to the inner membrane protease, DegS, to initiate the proteolytic cascade. DegS removes the periplasmic domain of RseA creating a substrate for the next protease in the pathway, RseP. RseP cleaves RseA in the periplasmic region in a process called regulatory intramembrane proteolysis. The remaining fragment of RseA is released into the cytoplasm and fully degraded by the ATP-dependent protease, ClpXP, with the assistance of the adaptor protein, SspB, thereby freeing σ^E to reprogram gene expression. A growing body of evidence indicates that the overall proteolytic framework that governs the σ^E response is used to regulate similar anti-sigma factor/sigma factor pairs throughout the bacterial world and has been adapted to

recognize a wide variety of signals and control systems as diverse as envelope stress responses, sporulation, virulence, and iron-siderophore uptake. In this chapter, we review the extensive physiological, biochemical, and structural studies on the σ^E system that provide remarkable insights into the mechanistic underpinnings of this regulated proteolytic signal transduction pathway. These studies reveal design principles that are applicable to related proteases and regulatory proteolytic pathways in all domains of life.

*Submitted for publication as: Sarah E. Barchinger[1] and Sarah E. Ades "Regulated Proteolysis: Control of the *Escherichia coli* σ^E-dependent Cell Envelope Stress" General and Regulatory Proteolysis in Microorganisms

Introduction

Proteolysis serves a dual role in biological systems as both a cellular housekeeper and an orchestrator of regulatory systems. As housekeepers, proteases rid the cell of damaged proteins in a relatively non-specific manner, helping to preserve the efficiency of cellular physiology. In contrast, when acting as regulators, proteases degrade or process only select substrates in response to distinct signals, yielding a defined change in the activity of their targets. Unlike regulatory systems in which protein activity is reversibly modulated by ligand binding, proteolysis provides a rapid and irreversible change in the activity of target proteins. Regulatory proteolysis rivals post-translational modifications in the diversity of cellular pathways it controls, ranging from transcriptional regulation to modulation of enzyme activity.

One of the most widely distributed and highly conserved regulatory proteolytic systems solves a biological problem common to all cells: how a signal generated on one side of a membrane can be communicated across the membrane to elicit the required response. This regulatory paradigm governs transmembrane signaling responses ranging from the cell envelope stress response of *E. coli*, the subject of this chapter, to the SREBP (sterol regulatory element binding proteins) pathway of humans. In its simplest form, the overall design of the proteolytic system consists of two proteases that sequentially cleave a transmembrane regulatory protein on either side of the membrane, resulting in release of the regulatory domain with a distinct biological function (Fig. C-1) [298]. These systems are most commonly found in compartmental membranes in eukaryotes and the cytoplasmic membrane of prokaryotes. The signal is generally sensed on the lumenal or extracytoplasmic side of the membrane and the regulatory protein is released into the

cytoplasm [298-300]. The first protease cleaves the target protein in its lumenal domain only after receiving the inducing signal, removing most of this domain. The second protease then cleaves the remaining fragment of the protein in the membrane-spanning region in a process termed regulated intramembrane proteolysis, or RIP [298, 299]. Intramembrane cleavage releases the regulatory domain from the membrane so that it can fulfill its biological mission [299]. Not only is the overall framework of this proteolytic pathway conserved, but the proteases are also evolutionarily related. The initiating proteases are often serine or aspartate proteases [298, 299]. Several conserved families of proteases have been identified that perform intramembrane cleavage: the S2P metalloproteases, rhomboid serine proteases, and pre-senilin aspartate proteases [299]. Only S2P and rhomboid proteases have been found thus far in prokaryotes.

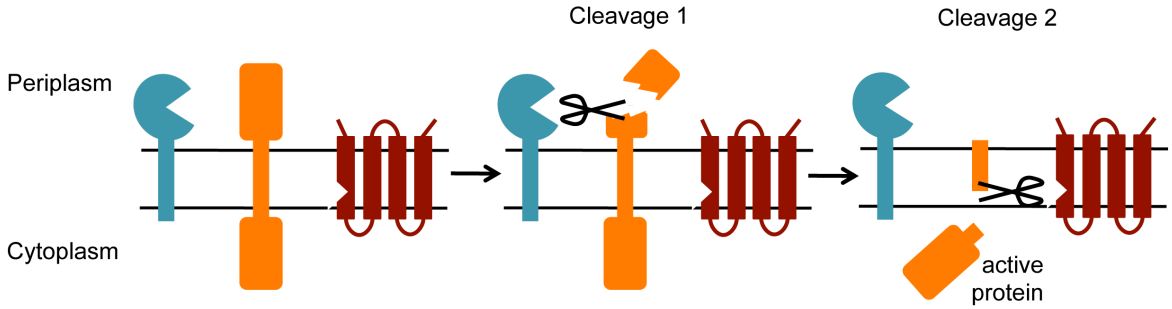


Figure C-1: Overview of transmembrane signal transduction by regulated intramembrane proteolysis. The first cleavage is signal-dependent and removes the lumenal domain of the substrate (orange). The RIP protease (red) then cuts the substrate in the membrane spanning region only after the first cleavage event and releases the biologically active domain of the substrate into the cytoplasm.

Among the first examples of regulatory proteolysis involving RIP were the SREBP pathway that controls lipid metabolism in animals and the sporulation system of *Bacillus subtilis* [298]. SREBP is a transcription factor that activates the expression of genes

required for lipid synthesis and uptake. It is synthesized as a membrane-bound precursor that is localized to the endoplasmic reticulum when sterols are present [301]. The two proteases, Site-1 protease (S1P) and Site-2 protease (S2P), are also integral membrane proteins localized in the Golgi apparatus, and therefore unable to access SREBP. When sterols are depleted, SREBP translocates to the Golgi, where it is cleaved first by S1P, a subtilisin-like serine protease [302]. This cleavage creates a substrate for S2P, a zinc metalloprotease, that cleaves SREBP in the first transmembrane segment, releasing the amino terminal transcription factor domain into the cytoplasm where it rapidly translocates into the nucleus [302].

In *B. subtilis*, an analogous proteolytic system regulates the activity of the transcription factor σ^K, which controls genes required for forespore development. σ^K is synthesized in an inactive, membrane-anchored form, similar to SREBP [303]. At the appropriate time during sporulation, pro-σ^K is cleaved within the membrane-spanning region by SpoIVFB, an ortholog of mammalian S2P, releasing σ^K to direct transcription [303]. Just as S2P can only cleave SREBP after S1P acts, cleavage of pro-σ^K is reliant on an upstream proteolytic event. Two proteins, SpoIVFA and BofA, bind to SpoIVFB and prevent it from cleaving pro-σ^K [304, 305]. Signals from the developing forespore lead to production of the serine proteases, SpoVB and CtpB, which cleave SpoIVFA, alleviating inhibition of SpoIVFB [306, 307]. The overall similarities of the design of these two proteolytic systems, the polytopic membrane-bound metalloprotease S2Ps, and their signal-dependent cleavage of a membrane-bound substrate to release a biologically active regulatory protein led to the realization that this framework is a conserved solution for

transmembrane signaling. Diverse proteolytic regulatory systems with this overall design have since been found in all domains of life.

In this chapter, we describe one of the best-studied systems that utilizes RIP, the proteolytic cascade that governs the σ^E-dependent cell envelope stress response in *E. coli* (Fig. C-2). The key players in the σ^E system are the transcription factor, σ^E, its regulators, RseA and RseB, and the proteases DegS, RseP, and ClpXP. The response is activated by stresses that affect the integrity of the outer compartment of the bacterium, the cell envelope [53, 308]. The proteolytic cascade serves to communicate this information across the inner membrane to the transcription factor, σ^E, in the cytoplasm. The inner membrane proteases, DegS and RseP, correspond to S1P and S2P, respectively [309, 310]. Their target is another inner membrane protein, RseA (for Regulator of SigmaE), which is an anti-sigma factor that binds tightly to σ^E sequestering it at the membrane [53, 75]. Signals that induce the pathway activate DegS to initiate the proteolytic cascade. The sequential action of DegS followed by RseP releases the cytoplasmic domain of RseA still bound to σ^E from the membrane [309, 310]. The adaptor protein SspB binds near the new C-terminus of RseA, and delivers it to the cytoplasmic protease, ClpXP. ClpXP then degrades the remaining portion of RseA, freeing σ^E to bind to RNA polymerase [311]. Extensive biochemical, structural, and physiological studies on this system have provided a remarkably detailed understanding of the proteolytic pathway, revealing fundamental principles of regulatory proteolysis that are applicable to all such systems.

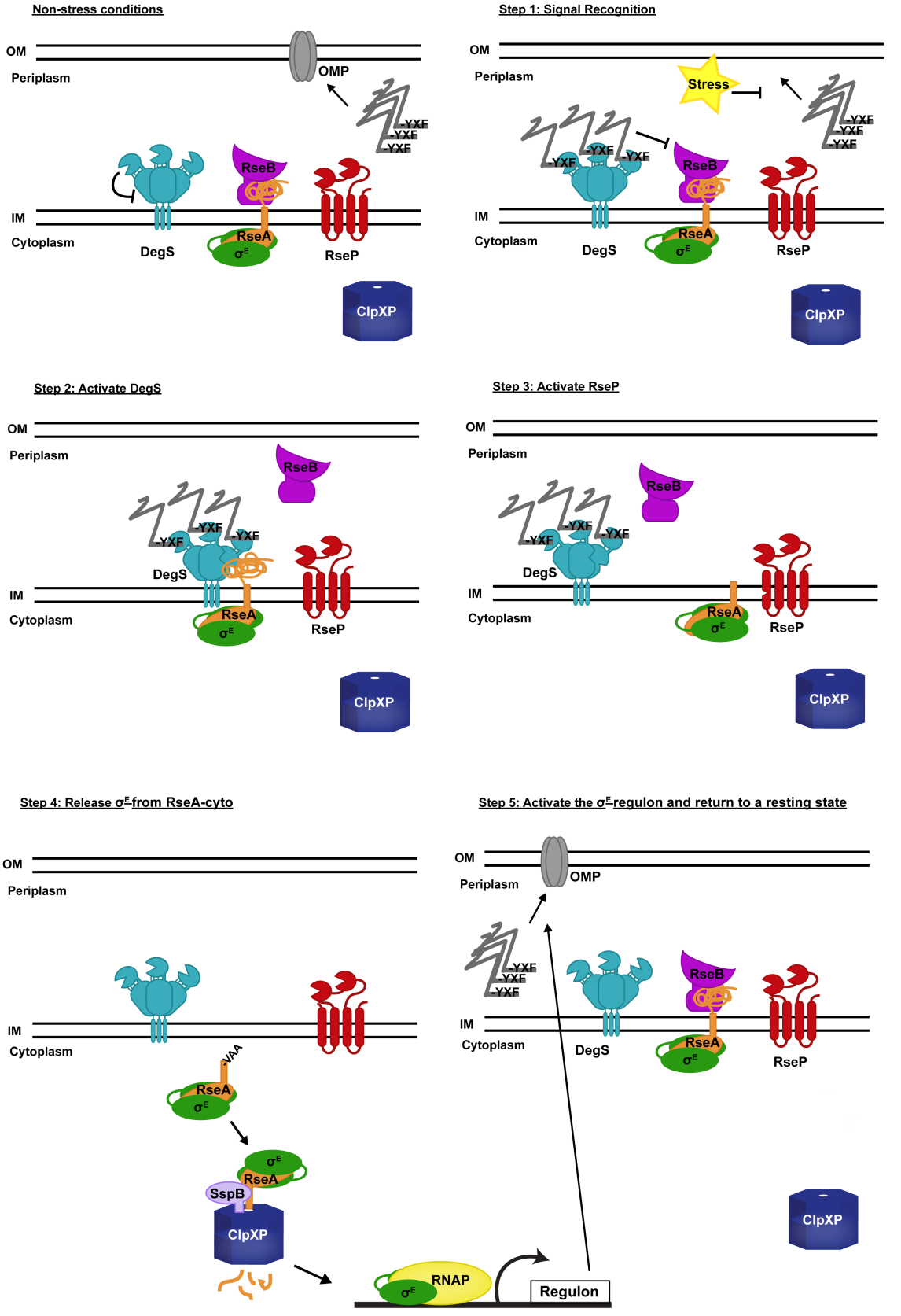

Non-stress conditions
OM
Periplasm
OMP
-YXF
IM
Cytoplasm
DegS
RseB
RseA
σE
RseP
ClpXP
Step 1: Signal Recognition
Stress
Step 2: Activate DegS
Step 3: Activate RseP
Step 4: Release σE from RseA-cyto
-VAA
SspB
Step 5: Activate the σE regulon and return to a resting state
RNAP
Regulon

Figure C-2: Regulation of the σ^E-dependent envelope stress response. The different steps of the proteolytic cascade regulating σ^E are outlined.

Cell envelope stress and σ^E

The cell envelope is the hallmark of Gram-negative bacteria. It is composed of the inner membrane, outer membrane, and periplasmic space between the membranes that contains a thin layer of peptidoglycan. This compartment not only is critical for the structural integrity of the cell, but also serves as the interface by which the bacterium interacts with its surroundings. The cell envelope is not a static structure, but is actively remodeled in response to changes in the environment [7]. Cell envelope stress responses have evolved to maintain the integrity of this important compartment, to assist in the elaboration of complex cell envelope structures such as pili and fimbrae, and to protect the envelope from damaging stresses [5, 23]. The σ^E-dependent envelope stress response is one of the key pathways that monitors the state of the cell envelope [61, 71].

In bacteria, promoter recognition is conferred by the sigma subunit of RNA polymerase. The primary sigma subunit in the cell is σ^{70}, while σ^E is one of an array of alternative sigma factors that displace σ^{70} and direct RNA polymerase to the promoters of genes in their regulons in response to specific inducing signals and/or stresses [39, 158]. As such, activation of σ^E rapidly reprograms transcription to focus on genes that allow the cell to cope with cell envelope stress. σ^E was first identified not for its role in extracytoplasmic stress, but for its role in transcribing the gene encoding another alternative sigma factor, the cytoplasmic heat shock factor σ^{32}, at high temperatures [67]. A connection between σ^E and the cell envelope was found in genetic screens demonstrating that σ^E activity increased following overproduction of outer membrane

porins, inactivation of genes encoding periplasmic chaperones and proteases, and deletion of genes involved in lipopolysaccharide biosynthesis [68, 69, 312, 313]. These findings, along with work on the CpxAR two-component system, were among the first pieces of evidence indicating that Gram-negative bacteria use distinct stress response pathways to combat damage in the cell envelope and the cytoplasm [9, 314, 315].

Compartmentalization of stress responses in *E. coli* was strikingly demonstrated for the σ^E system by the observation that overproduction of the outer membrane porin, OmpX, with an intact signal sequence activated σ^E, but not σ^{32} (we note that post-transcriptional regulation of σ^{32} prevents its activation under these conditions despite increased transcription by σ^E), while accumulation of OmpX in the cytoplasm due to disruption of its signal sequence resulted in activation of σ^{32}, but not σ^E [72].

σ^E systems have now been identified in at least 112 sequenced bacterial genomes and, where investigated, share a number of basic properties [40]. The gene encoding σ^E, *rpoE*, is essential in *E. coli*, *Yersinia* spp., and probably *Vibrio cholerae* [61-63]. In other bacteria, such as *S. enterica* serovar Typhimurium and *Bordetella bronchiseptica*, the σ^E system is important for interactions with the host immune system during infection [65, 97]. Despite differences in the responses to specific stress conditions across species, in all bacteria where the σ^E system has been studied in any detail, it has been found to be involved in cell envelope-associated processes [5, 6, 39, 40].

The major group of conserved genes in the σ^E regulon consists of a series of proteins including chaperones and proteases that are central to the synthesis, assembly, and maintenance of outer membrane porins (OMPs) and LPS [48]. In addition to the proteins that serve to fold or degrade misfolded OMPs, σ^E regulates the expression of several

sRNAs that target mRNAs encoding OMPs for degradation [77, 78, 262]. Unfolded OMPs, which can be toxic when they accumulate in the cell, serve as a barometer for the overall state of the cell envelope [53]. Their proper folding and assembly depends on lipoprotein and LPS biosynthesis, as well as chaperones that prevent aggregation and escort unfolded OMPs across the periplasm, and the Bam complex that assembles OMPs in the outer membrane [7, 76, 316]. Therefore, disruption of any component of the cell envelope that hinders OMP folding activates the σ^E response, which, in turn, increases the levels of proteins important for both OMP and LPS synthesis, while decreasing the load on the envelope by reducing de novo synthesis of OMPs via the sRNA regulators. In addition to the aforementioned regulon members, σ^E transcribes its own gene in most bacteria in which it is found [46, 62, 69, 70, 98, 183]. Activation of σ^E, therefore, results in an autoregulatory loop ensuring that σ^E continues to be made as long as the inducing stress remains.

Regulation of σ^E activity

σ^E activity is controlled by two proteins, RseA and RseB, that are encoded in the same operon as the gene encoding σ^E [56, 57]. This operon structure is widely conserved amongst σ^E orthologs [40], suggesting that the regulatory pathway is conserved as well. RseA is the central player in the regulatory system that controls σ^E activity. It is a single-pass transmembrane protein located in the inner membrane. The cytoplasmic domain is a σ^E-specific anti-sigma factor [56, 189]. Early work on the σ^E system found that deletion of *rseA* resulted in constitutively elevated σ^E activity and rendered σ^E insensitive to signals in the cell envelope [57]. These data provided strong evidence that RseA forms the critical

link between events in the cell envelope and σ^E in the cytoplasm. In addition, overexpression of the cytoplasmic domain of RseA in a Δ*rseA* mutant greatly reduced σ^E activity and did not restore the response to envelope stress, demonstrating that the inducing signal is generated in the periplasm [56]. The co-crystal structure of the cytoplasmic domain of RseA bound to σ^E revealed the molecular basis of RseA's anti-sigma factor activity (Fig. C-3) [189]. RseA forms a compact helical structure that is sandwiched between the two conserved domains of σ^E that are responsible for promoter recognition and binding to core RNA polymerase [33, 189]. Biochemical studies complement the structural data, and show that RseA is a strong competitive inhibitor of RNAP for binding to σ^E. In fact, the K_d of the RseA:σ^E complex is estimated to be < 10 pM, compared to ~ 1 nM for σ^E binding to RNA polymerase [317, 318].

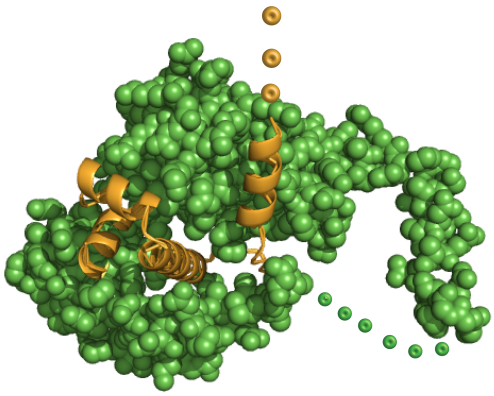


Figure C-3: σ^E bound to the cytoplasmic domain of RseA. Ribbon representation of the cytoplasmic domain of RseA (orange) embedded between conserved regions 2 and 4 of σ^E (green), shown in space-filling mode (1OR7 [189], generated using PyMol [190]). The surfaces used by σ^E to contact RNA polymerase are buried in the interface of the complex. Green dots represent the linker between regions 2 and 4 of σ^E, which was not ordered in the crystal structure. Orange dots represent the residues of RseA that lead to the transmembrane domain in the cytoplasmic membrane.

RseB is the second key regulator of σ^E. It is a soluble periplasmic protein and binds to the periplasmic domain of RseA [56, 57]. RseB does not regulate σ^E in the absence of RseA [56, 57]. Deletion of *rseB* results in a modest two- to three-fold increase in σ^E activity, and σ^E activity is still induced in response to envelope stress in cells lacking *rseB*, suggesting that RseB fine-tunes the response [57, 319]. In contrast, in vitro experiments suggested that RseB plays a greater role in regulating the pathway, because RseB completely shielded RseA from DegS-dependent cleavage in a reconstituted purified system, even in the presence of inducing signals [129]. Recent work indicates that RseB must be inactivated along with and independently of DegS activation, before the proteolytic cascade can begin [320]. Therefore, as described below, it is now thought that RseB plays a major role in maintaining the uninduced state of the system.

The proteolytic cascade

The discovery that the signal transduction pathway is controlled by proteolysis was uncovered through a series of experiments establishing that RseA is an unstable protein whose half-life in the cell is correlated with σ^E activity [74, 81]. The half-life of RseA decreased under conditions of envelope stress, elevated temperature and over-expression of outer membrane proteins, when σ^E activity was high [74, 81]. The half-life increased in strains lacking the regulator OmpR, when σ^E activity was low ([215] and S.E. Ades, unpublished observations). RseA was degraded completely, presumably to completion, since no fragments were observed by western blotting with antibodies raised against the periplasmic or cytoplasmic domains. Once the stress was removed, the system reset, *i.e.* the stability of RseA returned to that in the absence of stress [81]. A survey of strains

lacking periplasmic and inner membrane proteases revealed that RseA was stable, and σ^E was no longer inducible in a strain in which the gene encoding the inner membrane serine protease, DegS, was inactivated [74]. Genetic studies demonstrated that DegS is encoded by an essential gene whose function is to degrade RseA so that sufficient σ^E will be available to support viability [215]. Strains lacking DegS accumulated suppressor mutations that could also suppress the requirement of σ^E for viability, providing further evidence of the close connection between DegS and σ^E [215].

Participation of a second protease in the pathway was found in studies on the RseP protease, originally called YaeL. RseP was first identified as an S2P zinc metalloprotease ortholog and was shown to be an integral inner membrane protein [216]. Like DegS, RseP is encoded by an essential gene [309]. Genetic studies identified *rpoE* as a multicopy suppressor that allowed *E. coli* to grow in the absence of *rseP*, establishing a genetic connection between RseP and the σ^E pathway [310]. Both the Ito and Gross groups demonstrated that DegS and RseP worked in tandem to cleave RseA in a scheme analogous to the SREBP system [309, 310]. DegS cuts first and releases most of the periplasmic domain of RseA (Fig. C-2) [309]. The remaining fragment of RseA (residues 1-148) stays in the inner membrane and retains anti-sigma factor activity [129]. This fragment is a substrate for RseP, which cuts RseA in the transmembrane region, releasing the cytoplasmic domain of RseA (residues 1-108), still bound to σ^E, into the cytoplasm (Fig. C-2) [311, 321-323]. The proteases that degrade the periplasmic and inner membrane fragments of RseA released upon cleavage by DegS and RseP, respectively, have yet to be identified.

The final step in the proteolytic pathway that completes the degradation of RseA (Fig. B-2) was found not from studies of the σ^E pathway, but from a proteomic analysis of substrates of the cytoplasmic protease ClpXP [184, 311]. The cytoplasmic domain of RseA was among the proteins trapped in the cavity of a catalytically inactive ClpP variant. RseP cleavage exposes recognition signals for ClpXP at the C terminus of the soluble $RseA^{1-108}$ fragment released from the membrane [184, 324]. In vitro and in vivo experiments verified that ClpXP degraded $RseA^{1-108}$, but not σ^E, thereby freeing σ^E to bind RNAP and transcribe the genes in its regulon to combat cell envelope stress [311].

Thus, three proteases are required to fully degrade RseA and initiate the stress response. DegS is the sole protease to sense the inducing signal and each cleavage event generates a substrate for the next protease in the proteolytic cascade [317]. As a result, σ^E is both released from RseA and prevented from re-binding to RseA, so that it is free to bind RNA polymerase and transcribe the genes in its regulon. DegS cleavage of RseA is the rate-limiting step in the full degradation of RseA, which specifically tunes the system to the folding state of OMPs [317]. In the sections below, we outline in detail the structural, biochemical, and physiological details of the proteolytic pathway that controls σ^E activity.

Activating σ^E via regulated proteolysis

Step 1: Signal Recognition

The proteolytic pathway that activates σ^E has many built-in checkpoints to ensure that σ^E is properly regulated, only activated when necessary, and only to the extent required. The system is held in the "off-state" with minimal signal-independent proteolysis

by multiple inhibitory interactions that control the activity of DegS and RseP [319, 320, 325-327]. These inhibitory interactions are alleviated as the proteolytic cascade progresses, resulting in complete degradation of RseA.

Two independent signal recognition events are currently known to be required to initiate the proteolytic cascade: 1) DegS, which exists primarily in an inactive conformation, must be activated, and 2) RseB, which protects RseA from proteolysis, must be inactivated (Fig. C-2, top row). We are just beginning to understand how RseB activity is modulated. In contrast, structural and biochemical studies have yielded a wealth of information about how DegS activity is regulated.

DegS is a member of the HtrA family of serine proteases in which the protease domain is followed by one or more peptide-binding PDZ domains [328]. DegS is anchored in the inner membrane by a single transmembrane helix and the majority of the protein, including the active site and its single PDZ domain, project from the membrane into the periplasm [329, 330]. DegS lacking the transmembrane domain can be expressed as a soluble protein that retains the same properties as the intact protein in a purified in vitro system [327]. However, DegS is inactive in vivo without the transmembrane region, suggesting that it must be localized to the inner membrane near RseA to function in the cell [215].

A survey of peptides that bind to the DegS PDZ domain revealed that DegS preferentially bound to peptides with the C-terminal sequence, YXF. In vitro experiments demonstrated that binding of these peptides converted DegS from a proteolytically inactive state into an active state [327]. These findings provided the critical piece of information to explain how envelope stress is sensed by the σ^E pathway. YXF tripeptides are found at the

extreme C terminus of many of the major OMPs in *E. coli* [327]. These C-terminal residues are buried between beta strands in the fully folded OMP beta-barrel structure, and are inaccessible in the properly folded protein [331]. However, when OMP folding is disrupted, the residues are exposed.

OMP folding and insertion into the outer membrane is a complex process [76, 332]. OMPs are translocated as unfolded polypeptides from the cytoplasm to the periplasm via the Sec machinery. Following secretion, periplasmic chaperones bind to the unfolded OMPs to prevent their aggregation in the periplasm. The chaperones then deliver the OMPs to the Bam complex in the outer membrane, which assists in folding and assembly of the properly folded OMPs in the membrane [76, 316]. If any of these steps are disrupted or if the system is overwhelmed, folding intermediates accumulate with exposed C termini. The majority of the known inducers of the σ^E pathway have the potential to disrupt the proper folding of OMPs and include conditions that stress both the cytoplasm and cell envelope, such as heat and addition of ethanol, as well as those that specifically stress the cell envelope, such as deletion of periplasmic folding catalysts and chaperones, deletion of genes required for proper elaboration of LPS, and overexpression of outer membrane proteins [56, 68, 70, 312]. Most of the folding catalysts, chaperones, and members of the Bam complex are encoded by genes in the σ^E regulon [48, 333]. As such, problems associated with OMP maturation, due to overload or incapacitation of the folding and assembly pathway, trigger increased expression via σ^E of the very proteins needed to restore the flux of OMPs to the outer membrane.

In the second event required to induce the σ^E response, RseB must be inactivated. It was originally proposed that unfolded proteins in the periplasm competed with RseA for

binding to RseB and titrated RseB away from RseA when they accumulated [55]. However, this model is not well supported by in vivo or in vitro experiments [319]. Recent work suggests that, in addition to activating DegS, unassembled OMPs antagonize inhibition by RseB through different parts of the C-terminal peptide [320]. Residues between 10 and 20 amino acids preceding the YXF-COOH motif were found to be required for full activation of the response in vivo in the presence of RseB, but not in its absence. In addition, sequences in this region from different OMPs activated the response to slightly different extents, again only in the presence of RseB. These data led to the model that residues upstream of the YXF motif are specifically required not to activate DegS, but to antagonize RseB [320]. However, the peptides did not alleviate inhibition in vitro, indicating that another component is required to antagonize RseB [320]. Therefore, the mechanism that relieves inhibition by RseB remains unclear. Based on structural homology between RseB and the lipid-binding domains of LolA, LolB, and LppX [129, 320, 334], it has been proposed that a lipid, free lipoprotein, or LPS that has not been correctly delivered to the outer membrane is the second signal required for activation of the response [320], although this model has yet to be tested experimentally. If the inducing signal for RseB proves to be a lipophilic molecule, then DegS and RseB integrate distinct signals from the cell envelope to control the stress response.

The crystal structure of RseB bound to the periplasmic domain of RseA ($RseA_{peri}$) revealed how RseB protects RseA from proteolysis by DegS (Fig. B-4) [335]. RseB consists of two domains, a smaller C-terminal domain and a larger N-terminal domain, which has homology to lipoprotein-binding proteins [334-336]. $RseA_{peri}$ is largely unstructured with the exception of two regions encompassing residues 132-151 and 169-

190 [335]. The structural observations are consistent with previous experiments indicating that $RseA_{peri}$ assumed a molten globule-like conformation [327]. In biochemical studies, $RseA^{169-190}$ was found to be necessary and sufficient for binding to RseB [129]. In the $RseA_{peri}$:RseB co-crystal structure, these residues forms a helical structure that binds to the smaller domain of RseB (Fig. B-4). The other structured region of RseA in the complex, $RseA^{132-151}$, includes the site where RseA is cleaved by DegS [335]. These residues bind in the cleft between the two domains of RseB, largely burying the cleavage site, suggesting that RseB prevents cleavage of RseA by blocking access to the sessile bond (Fig. C-4). RseB also protects RseA from proteolysis by RseP in a process that is less well understood and involves the PDZ domains of RseP [319].

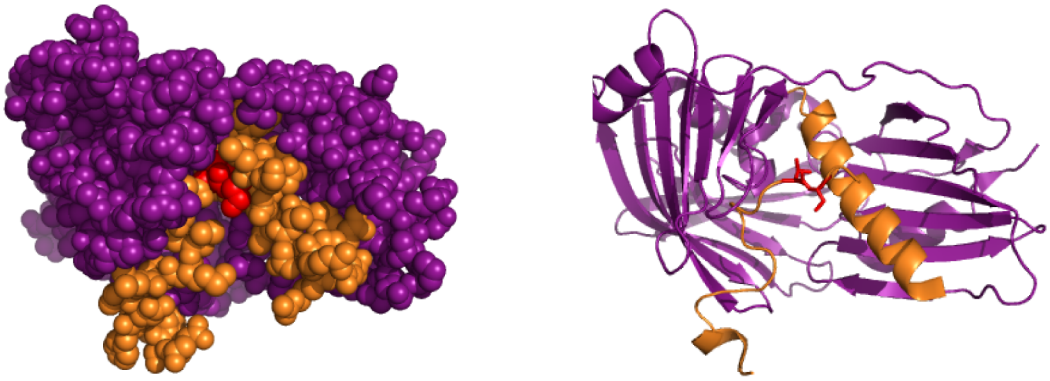


Figure C-4: RseB bound to the periplasmic domain of RseA. Space filling and ribbon representations of RseB in (purple) bound the periplasmic domain of RseA (orange) are shown (3M4W [335], generated using PyMOL [190]). The peptide bond cleaved by DegS, Val148-Ser149 (side chains shown in red), is almost completely occluded. Two regions of RseA were ordered in the crystal structure, and bind in the cleft between the two domains of RseB.

Step 2: Activate DegS

DegS is the gate-keeper of the σ^E response (Fig. C-2, middle left). Once it cleaves RseA, all the subsequent cleavage events occur in a signal-independent manner and with kinetics that are faster than the initial cleavage by DegS [317]. Therefore, the level of DegS activity, which is set by the amount and identity of the unfolded porins, determines how much σ^E is released from RseA and the extent of the response. Structural and biochemical studies have provided amazingly detailed views of DegS and how it functions.

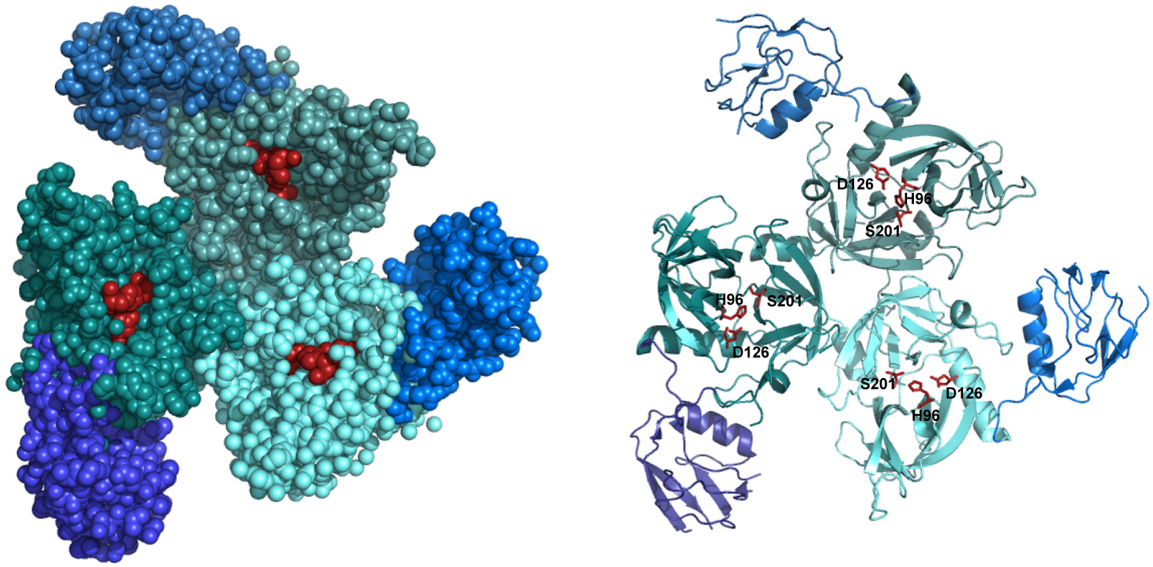


Figure C-5: DegS trimer. Space-filling and ribbon representations of the DegS trimer are shown viewed from the top (1SOT [329], generated using PyMOL [190]). The cytoplasmic membrane is below the molecule. The active site residues, H96, D126, and S201 (side chains shown in red) are accessible to the periplasm. The PDZ domains that are arranged around the peripmeter of the trimer are shown in blue. The protease domains of each trimer are in different shades of green.

Given the destructive nature of proteases, including DegS, their activity must be controlled to prevent rampant degradation of cellular proteins. Like other serine proteases, DegS is held in an inactive state until a specific activating event occurs [300]. In structures of DegS, the protease domains form a funnel-shaped trimer with the PDZ domains

decorating the edges (Fig. C-5). Although the active sites are fully exposed on the inner surfaces of the trimer facing the periplasm, DegS has extremely low activity in the absence of inducing peptide [327, 329, 330]. The crystal structures of unliganded DegS provide a ready explanation for this apparent dichotomy. The active site Ser-His-Asp triad is not appropriately aligned for catalysis, and the oxyanion hole is not in the proper conformation to form the requisite hydrogen bonds needed to stabilize the tetrahedral intermediate of the peptide cleavage reaction (Fig. C-6) [329, 330]. Therefore, although the active sites are exposed, they can do no damage in this inactive conformation. By contrast, in structures of the active form of the enzyme bound to inducing peptides, the catalytic triad is realigned in the appropriate position for catalysis (Fig. C-6) [329, 337]. The oxyanion hole is also properly formed due to rotation of His198 (Fig. C-6), which repositions the backbone amide so that it can form a hydrogen bond with the carbonyl group of the sessile peptide bond [329, 330, 337, 338].

Role of the PDZ domains: As with many other members of the HtrA family of proteases, the PDZ domain of DegS regulates its proteolytic activity [328]. In a purified system, DegS cleaved RseA very slowly in the absence of peptide [339]. Peptides that bound to the PDZ domain dramatically increased the rate of cleavage, with the strongest inducer, the YYF tripeptide, increasing the rate of cleavage by nearly 1000-fold [219]. In contrast, DegS lacking the PDZ domain was only 4-5 fold less active that peptide-bound DegS (~200-fold more active than unliganded DegS) and was no longer sensitive to inducing peptides [338-340]. In vivo, DegSΔPDZ had a 12-fold higher basal level of activity than wild-type DegS in strains lacking *rseB* (*rseB* was deleted to separate DegS activity from inhibition of cleavage by RseB) [320]. These data demonstrate clearly that

the unliganded PDZ domains of DegS act as negative regulators of the protease domains and peptide binding alleviates this inhibition.

Comparisons of the liganded, unliganded, active, and inactive structures of DegS have provided information as to how the PDZ domain regulates DegS activity and how peptide binding leads to structural changes required for proteolysis. The interface between the PDZ and protease domains appears to be very flexible, particularly in the peptide-bound forms of DegS [339]. Therefore, detailed biochemical experiments with variants of DegS with specific site-directed mutations have been critical in determining which of the contacts seen to change amongst the various DegS structures are actually important for regulated proteolysis.

In the unliganded state, the PDZ domain stabilizes the inactive conformation of DegS [329, 337, 339, 340]. The PDZ domain is anchored to the protease domain through a series of interactions that include three salt bridges that span the two domains: D122-R256, R178-E317/D320, and K243-E324 (Fig. C-6) [329, 337]. Two of the three salt bridges are no longer seen in structures of peptide-bound DegS (Fig. C-6), and the PDZ domains appear to be more mobile relative to the protease domain than in the unliganded enzyme, suggesting that in the active enzyme, the PDZ domains are no longer tightly associated with the protease domains [329, 337, 340]. Disruption of the salt bridges by site-directed mutagenesis increased DegS activity, consistent with a role for the PDZ domain in stabilizing the inactive form of the enzyme [320, 339]. In addition to holding the enzyme in an inactive conformation, two of the amino acids that form salt bridges with the PDZ domain, Arg178 and Asp122, also participate in interactions that stabilize the active site when peptide is bound [329, 338, 339]. Arg178 in particular appears to be a key residue in

the transition. In the inactive enzyme, Arg 178 forms a salt bridge with a pair of amino acids in the PDZ domain (Fig. B-6) [329, 330]. In the liganded form of the enzyme, it rotates 90° and makes a different set of hydrogen bonds that link it to the functional conformation of the oxyanion hole [329, 337-339]. Asp122 is part of a hydrogen-bonding network that includes the peptide backbone adjacent to the catalytic site and stabilizes the active conformation in the liganded enzyme [329, 339]. Other amino acids also change conformation in the liganded structure and help to promote the active conformation. These residues include, among others, His198, mentioned above, and Tyr162, which also moves allowing the formation of hydrogen bonds that stabilize the active conformation of the peptide backbone around His198 and the oxyanion hole [329].

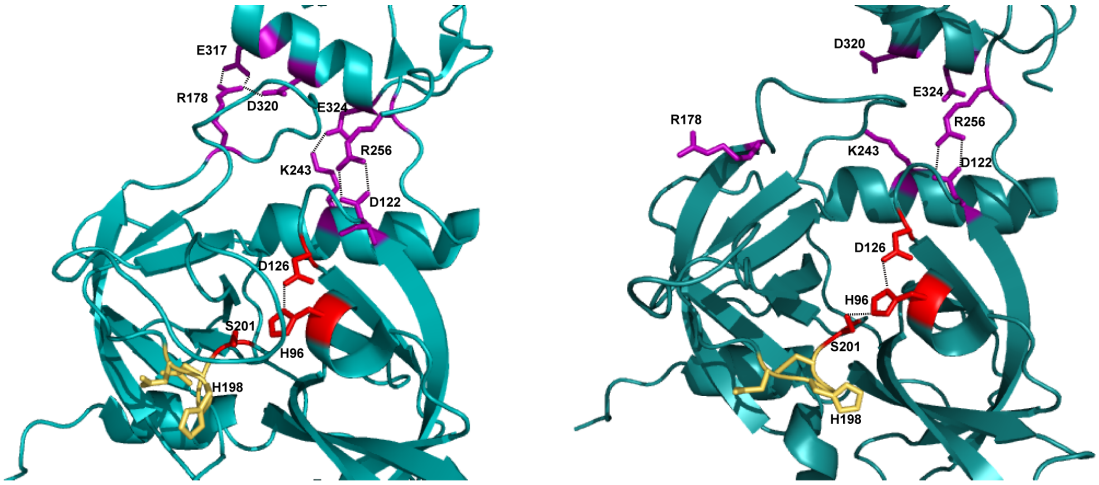


Figure C-6. Free and peptide-bound DegS monomer. *Left*, Ribbon representation of the inactive, peptide-free monomer of DegS (1TE0 [330], generated using PyMOL [190]). *Right*, Ribbon representation of the active, peptide-bound monomer of DegS (1SOZ [329], generated using PyMOL [190]). Side chains of the amino acids forming salt bridges between the PDZ domain (top) and protease domain (bottom) are shown in purple and salt bridges are indicated by dashed lines. Residues participating in the formation of the oxyanion hole (197-201) are shown in yellow. Side chains in active site catalytic triad (His96, Asp126, Ser201) are shown in red. Hydrogen bonds between side chains in the active site are shown by dashed lines.

Mechanism of activation: Two models have been proposed to explain how peptide binding leads to enzyme activation. The first model, called the scaffolding or peptide-activation model, proposes that the penultimate residue of the activating peptide contacts the protease domain and directly participates in a network of interactions that serve to remodel the active site [329, 337]. This model is based on crystallographic evidence showing that this amino acid interacts with the L3 loop in the protease domain, reorienting the stem of the loop through a series of interactions that ultimately stabilize the active conformation of the catalytic site [329, 337]. Since DegS can be activated by peptides with several different amino acids of chemically diverse nature at the penultimate position, proponents of this model propose that each peptide forms a slightly different set of interactions with the L3 loop to accommodate the different amino acids [337]. Each of these interactions ultimately alters the conformation of the L3 loop in a way that leads to stabilization of the active enzyme. As such, the L3 loop acts as a sensor of peptide binding via contacts with the inducing peptide bound to the PDZ domain. The PDZ domains in the scaffolding model not only stabilize the inactive state, but they are also required for the transition to the active state because they position the peptide to interact with the L3 loop [328, 329, 337].

Several observations counter the scaffolding model. The finding that the active form of the enzyme could be obtained by deleting the PDZ domain altogether suggests that the PDZ domain is not needed to form or stabilize the active conformation [338, 339]. Additionally, peptides that vary only at the penultimate residue activated DegS to nearly the same extent at saturating concentrations [219, 339]. More variation would be expected because different amino acids at the penultimate position must form different contacts with

the L3 loop that have different energies [219, 339]. Substitutions at other positions in the inducing peptides actually caused larger changes in activation than those at the penultimate positions [219, 339, 340]. Finally, in the structures of peptide-bound DegS of different crystal forms and with different peptides, the orientation of the PDZ domains varied substantially even within the same trimer, and no contacts were seen between the peptide and the protease domain in many of the structures [340].

The second model, called the relief of inhibition model, proposes that the PDZ domain holds the protease domain in the inactive conformation [339]. Peptide binding leads to an allosteric rearrangement that relieves the inhibitory interactions and stabilizes the active conformation of the protease domain. However, the PDZ domain itself is not required to stabilize the active enzyme. An extensive series of biochemical and structural studies have provided considerable support for the relief of inhibition model [219, 338-340]. The main caveat proposed against this model is that it is based on structural and biochemical studies with mutant enzymes, and these mutations may introduce changes that are not representative of the wild-type enzyme [328]. Nevertheless, given the thorough and careful experiments supporting it, the relief of inhibition model provides a compelling model for DegS activity and is outlined below. This model is backed by detailed biochemical analyses of peptide activation and the kinetics of proteolysis by DegS using an optimized in vitro degradation assay with purified components and is coupled with crystallographic studies of many DegS variants [219, 338-340]. At this point, there are over 20 different structures of wild-type DegS and DegS variants that provide many views of the protein.

Positive Cooperativity: A wide range of peptides ending in the YXF motif can activate DegS, which is likely to ensure that the system can monitor the folding of different OMPs in the cell [327]. However, not all peptides are equivalent activators of DegS. Peptides differ in the maximal extent of activation (V_{max}) and the concentration required for half-maximal activation (K_{act}) [219, 339]. In addition, the V_{max} and concentration dependency of activation are not correlated, so that a peptide that activates DegS to lesser extent than another peptide may do so at a lower concentration [219, 339]. Despite the differences in the kinetic parameters, all the peptides exhibit positively cooperative activation of DegS [219, 339]. Positive cooperativity indicates that binding of a peptide to one DegS in the trimer facilitates binding of peptides to the remaining monomers resulting a sensitive and rapid switch from the inactive to the active form of the enzyme. In addition to exhibiting positive cooperativity with respect to peptide binding, DegS also exhibits positive cooperativity in substrate degradation [338, 339]. When the concentration of RseA was varied in the presence of saturating peptide, the Michaelis-Menton plot was sigmoidal with a Hill constant >1 [338, 339]. Therefore, RseA facilitates its own degradation. Positive cooperativity in RseA degradation was also seen with the DegSΔPDZ variant, indicating that allosteric regulation of DegS is not confined solely to the PDZ domain, but is also an inherent property of the protease domain [338].

Allosteric regulation- The MWC model: The presence of interactions that stabilize the inactive form of an enzyme, combined with positively cooperative ligand binding, are hallmarks of the Monod Wyman Changot (MWC) model of allostery (Fig. C-7) [341]. This model has been used to explain allosteric regulation of diverse proteins ranging from enzymes such as aspartate decarbamylase to hemoglobin to G-protein coupled receptors

[342]. The central tenet of this model is that the protein exists in two conformations, tense (inactive) and relaxed (active), that are in a dynamic equilibrium [341, 342]. The tense state is more stable, and predominates in the absence of ligand. Ligands bind preferentially to the relaxed form of the protein, thereby shifting the equilibrium toward the relaxed state (Fig. B-7). Data from peptide activation experiments could be fit to the MWC model, and the model explained the variations in activating potential of different peptides [219, 339, 340]. Because peptides can bind to both tense and relaxed DegS, the difference in the affinity of a given peptide for each state determines the overall amount of DegS in the active form (Fig. C-7). The stronger the preference for the relaxed state compared to the tense state, the stronger the cooperativity and the greater the extent of activation.

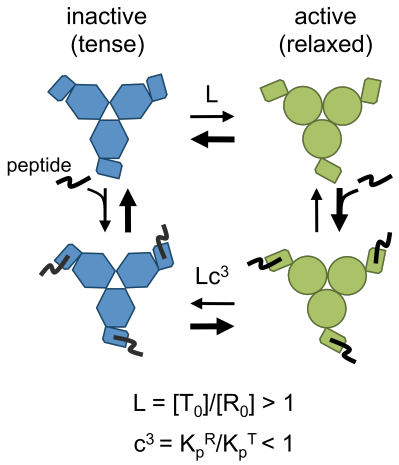


Figure C-7: MWC model for allosteric regulation of DegS by activating peptides. The cartoon depicts regulation of DegS by YxF peptides according to the MWC model of allostery [341]. The tense form (blue hexagons) of DegS is not proteolytically active and binds to inducing peptides with a lower affinity (K_T^P) than the proteolytically active relaxed form (green circles) of the enzyme (K_R^P). Without bound peptide, the ratio of the tense to relaxed forms of the enzyme (L) is greater than one, and the tense form of the enzyme predominates. The ratio of the two forms of DegS with bound peptide is given by Lc^3, where c is the ratio of the affinities of tense and relaxed DegS for peptide. Because c is less than one, Lc^3 is less than one and the relaxed active form predominates. The direction that is favored in the equilibria among the different states in shown with thick arrows. Only the fully peptide-bound forms of DegS are shown for simplicity.

For the MWC model to be an accurate description of DegS, mutations that lower the energy barrier between the tense and relaxed states, shifting the equilibrium toward the relaxed state, should increase activity of the unliganded enzyme and reduce cooperativity. These mutations could either destabilize the tense state or stabilize the relaxed state. For DegS, key candidates for interactions that stabilize the tense form are the salt bridges formed between the PDZ domain and protease domain in unliganded DegS. Disruption of each of these salt bridges by mutation increased the basal rate of proteolysis significantly, although not to the level of the fully liganded enzyme indicating that other interactions remain that stabilize the tense state [219, 339, 340]. These mutations also reduced the cooperativity of both peptide activation and RseA degradation, providing solid evidence in support of the MWC model [219, 339, 340].

Although reorientation of the PDZ domains is critical for the allosteric activation of DegS, the protease domains themselves exhibit allostery. Cleavage of RseA by DegSΔPDZ is cooperative, suggesting that the tense state is still significantly populated and not all of the enzyme is in the relaxed form [338]. Further evidence that the inactive state is populated even in the absence of the PDZ domain comes from experiments with the H198P variant of DegS. This mutation eliminated nearly all of the cooperativity of RseA degradation by DegSΔPDZ and increased the activity of DegSΔPDZ, as it does in the full-length enzyme [339]. Crystal structures of both H198P DegSΔPDZ and full-length H198P DegS revealed that the proline makes packing interactions that should stabilize the active conformation of the oxyanion hole [338, 339]. These stabilizing interactions would shift the equilibrium to favor the active, relaxed form. Indeed, a fit of the data to the MWC model predicted that 1% of the wild-type DegSΔPDZ enzymes are in the active

conformation in the absence of substrate, compared 87% of the H198P DegSΔPDZ variants [339]. One confounding piece of data is that the catalytic site is properly formed in crystal structures of DegSΔPDZ [338, 339]. However, it is likely that the active form is trapped by crystal packing conditions, given the biochemical data [338]. The cooperative substrate activation inherent in the protease domain may reflect an evolutionarily early form of allosteric regulation, before the PDZ domains were acquired, that provided a mechanism to reduce protease activity in the absence of substrate.

Taken together, what do these data and models mean? What benefit can be gained by having an allosteric system control the envelope stress response? Several answers to these questions have been proposed [219, 338, 339]. DegS is essential in *E. coli* because it must initiate degradation of RseA to release sufficient σ^E to maintain cell viability. Because DegS equilibrates between inactive and active forms, a small number of DegS enzymes will assume the active conformation, even in the absence of inducing peptides. This small population of active enzyme, combined with cooperative binding by RseA itself, may ensure sufficient basal level cleavage of RseA to maintain viability [219, 338, 339]. The basal level of uninduced degradation may also be supported by a low level of peptide-induced degradation due to stochastic fluctuations in the OMP folding pathways that expose a small number of inducing peptides at any given time. In terms of activating the response, positive cooperativity ensures that during stress, the pathway can be rapidly activated over a narrow concentration range. Activation is also reversible so that once the unfolded OMPs have been cleared, DegS will quickly return to the inactive state. Since different peptides activate the enzyme to different extents, it is tantalizing to envision that the response is tuned to monitor the folding of different OMPs.

Step 3: Activate RseP

Release of the periplasmic domain of RseA by DegS generates a substrate for RseP, the next enzyme in the proteolytic cascade (Fig. C-2, middle right). RseP is a zinc metalloprotease and belongs to the S2P group of RIP proteases that are found in a wide range of organisms [299, 343]. RseP is an inner membrane protein with four transmembrane (TM) segments [216]. The active site is formed by the HEXXH motif in TM1 and the LDG sequence in TM3, which are conserved in S2P family members [216]. TM3 is also thought to be critical for substrate binding by RseP [344]. Two circularly permuted PDZ domains are located in the periplasmic domain between TM2 and TM3 [321, 345]. Because they are polytopic membrane proteins, the S2P family of proteases has been far more difficult to characterize biochemically and structurally than proteases such as DegS, which can be readily expressed as soluble active enzymes.

A major question that is relative to all families of intramembrane proteases is how peptide bond cleavage takes place in the lipid environment of the membrane. Proteolysis is thought to occur through nucleophilic attack by a water molecule that is bound to the active site zinc and activated for peptide bond hydrolysis by the glutamate of the HEXXH motif [343]. Therefore, water must be able to access the active site of the enzyme. The structure of the catalytic core of a S2P family member from *Methanocaldococcus jannaschii* has been solved in the presence of detergents and provides a high-resolution view of how intramembrane peptide bond cleavage is likely to be achieved [346]. This S2P has six transmembrane segments, and the active site is positioned so that it lies within the plane of the membrane. The zinc ion is coordinated by the two histidine residues of the HEXXH motif in TM2 and the aspartate of the LDG motif in TM4, as predicted from

biochemical and genetic data [216, 323, 346]. The glutamate residue is also properly aligned for catalysis. A narrow channel lined with hydrophilic amino acids connects the active site to the cytosolic side of the membrane providing a way for water molecules to access the active site.

Although no structure is yet available for the catalytic domain of RseP, biochemical experiments provide some structural insights. The environment of the active site of RseP was analyzed by determining the accessibility of a membrane-impermeable alkylating reagent to cysteine residues engineered in the active site of the enzyme [347]. When the protein was in the native state in membrane vesicles, the cysteines were not modified. Accessibility increased when the protein denaturant, guanidine HCl, was present, and full modification of the cysteines was seen only when the membrane vesicles were solubilized with detergent and guanidine. Increased accessibility of the cysteines in the presence of guanidine indicates that the active site lies in a proteinaceous structure that can be accessed by the denaturant and at least partially unfolds [347]. These results, combined with cleavage site studies with RseA and model substrates, suggest that the active site of RseP is sequestered in a folded protein structure from the extramembrane environment and probably from the membrane lipids as well [347, 348].

Substrate selectivity: RseP appears to have a relatively broad substrate specificity compared to DegS, whose only known substrate is RseA. For example, RseP could cleave transmembrane segments TM1 and TM5 from LacY and the signal sequence from beta-lactamase in model substrates that contain no sequences related to RseA [348]. Experiments varying the sequence of target TM segments, including the TM of RseA, indicated that the major requirement for binding to and efficient cleavage by RseP was the

presence of helix-destabilizing residues within the transmembrane region of substrates, as opposed to a sequence-specific recognition motif [344, 347]. Destabilizing residues in the transmembrane helix may make the peptide backbone more accessible to RseP for hydrolysis compared to the peptide backbone of a stable helix that is fully engaged in hydrogen bonds. The ability of RseP to cleave substrates other than RseA and the relatively low sequence specificity suggest that RseP plays a role in the cell beyond the envelope stress response. Indeed, recent work indicates that RseP is responsible for degrading signal peptides once they have been cleaved from secreted proteins by signal protein peptidase [349]. A connection between the envelope stress response and degradation of signal peptides has not been established and this new activity may be an independent function of RseP.

Regulation of RseP: Similar to DegS, RseP cleaves full-length RseA very slowly [321-323, 345]. In contrast, the N-terminal fragment of RseA generated by DegS cleavage is rapidly cleaved by RseP [322, 323, 326]. This second cleavage event by RseP happens around three-fold faster than the initial cleavage by DegS, such that the signal-sensitive step is the rate-limiting step in the signaling pathway [317]. To further insulate the system from uninduced degradation of RseA by RseP, RseP activity is blocked by series of inhibitory interactions involving the PDZ domains of RseP, a glutamine-rich region in the periplasmic domain of RseA, RseB, and DegS [319, 325, 326]. The mechanism by which these different factors act to restrain RseP is not fully understood, especially compared to the wealth of information about DegS. RseB and the Gln-rich regions of RseA, but not DegS, protect RseA from cleavage by RseP only when the PDZ domains of RseP are intact [319]. These data suggest that either RseB or RseA interact with the RseP PDZ domains.

Structural studies indicate that the second PDZ domain of RseP can bind to the C-terminal amino acid of $RseA^{1-148}$, although binding is too weak to be detected with biochemical assays [321]. While the mechanism of inhibition is not clear, the framework for how some of the inhibitory interactions are relieved is evident. The RseB binding sites and the Gln-rich regions of RseA are on the C-terminal side of the DegS cleavage site, so they will be removed when DegS acts [327]. How inhibition by DegS and the PDZ domains is alleviated is not as easily explained and is not yet known. DegS may sequester RseA from RseP or directly interact with RseP in an inhibitory manner.

Role of the PDZ domains: The role of the PDZ domains in RseP remains an intriguing puzzle. Originally, RseP was predicted to have a single PDZ domain, but it was later shown through sequence alignments and crystallography to have two circularly permuted PDZ domains [321, 325, 326, 343, 345]. The RsePΔPDZ deletion used in the early studies was thought to have excised the single PDZ domain [325, 326], although in actuality the deletion removed part of each of the two circularly permuted PDZ domains. Nevertheless, the deletion disrupted the native PDZ domains and resulted in DegS-independent cleavage of RseA that was no longer inhibited by RseB [319, 325, 326, 345]. These results support a model in which the PDZ domains block RseP function. Additional evidence supporting a regulatory function for the PDZ domains came from genetic studies isolating mutations in RseP that increase the basal level of σ^E activity in vivo [345]. Most of these mutations fell in the predicted peptide-binding regions of both PDZ domains, with the strongest mutations in the N-terminal PDZ domain, suggesting that it plays a critical role in regulating RseP. The mutations, which are predicted to disrupt ligand binding, did not increase the intrinsic proteolytic activity of RseP, but instead increased the basal level

of σ^E activity by alleviating the requirement for prior cleavage by DegS. Interestingly, variants of RseP lacking either PDZ domain did not degrade full-length RseA, although they could still cleave $RseA^{1-140}$, which lacks most of its periplasmic domain. Therefore, it appears that either PDZ domain can regulate RseP.

How do the RseP PDZ domains block protease activity? One straightforward hypothesis is that one (or both?) of the PDZ domains bind to RseA to keep RseP in the "off" state. Although PDZ domains often bind to the C-termini of proteins, it is not likely that the C-terminus of intact RseA is the ligand. RseP does not degrade RseA variants that retain most of the periplasmic domain, but have different C termini [326]. If a specific binding interaction between RseA and the PDZ domains were required for inhibition, then changes to the C terminus should have abrogated the interaction and led to DegS-independent cleavage. Other experiments suggested that the C-terminal valine of $RseA^{1-148}$ was required to activate cleavage [321]. Mutation of Val148 in $RseA^{1-148}$ to dissimilar amino acids reduced cleavage by RseP [321]. This model is attractive, since similar results have been found for the intramembrane protease, γ-secretase, of animals, suggesting that this mechanism is evolutionarily conserved [350]. However, RseP will cleave a variety of model substrates and RseA fragments with different C-terminal amino acids, indicating that peptide binding to the PDZ domain is not a prerequisite for activity, or that the specificity of binding is quite broad. [326, 348] It is possible that RseA is a unique substrate for RseP and interacts with the protein somewhat differently than other substrates. Future experiments will surely clarify the role of the PDZ domains and illuminate the mechanistic underpinnings of the proteolytic activity of RseP and its regulation.

Step 4: Releasing σ^E from RseA-cyto

After RseP cleaves RseA, the remaining fragment of RseA (residues 1-108) is released from the membrane with σ^E still tightly bound [321]. The final step in the proteolytic cascade completes the degradation of RseA, releasing σ^E to bind to RNA polymerase (Fig. B-2, bottom left). Because the interaction between σ^E and the cytoplasmic domain of RseA is extremely stable and the dissociation rate is extremely slow, proteolysis is the predominant mechanism to free σ^E [317]. The fragment of RseA remaining after RseP cleavage, $RseA^{1-108}$, contains the σ^E binding domain (residues 1-66), followed by residues that are not required for σ^E binding, but target the protein for degradation. $RseA^{1-108}$ terminates with the sequence VAA, which is a recognition sequence for ClpX, the ATP-dependent unfoldase of the ClpXP protease [184, 311]. Upstream of the ClpX binding site is a binding site for the adaptor protein, SspB, which facilitates proteolysis by ClpXP [324]. Both in vitro and in vivo experiments demonstrated that ClpXP degraded $RseA^{1-108}$ rapidly, and that SspB increased the degradation rate even further [311]. Although ClpXP is the major protease to degrade $RseA^{1-108}$, other cytoplasmic proteases can also degrade it to release σ^E [311, 317]. Thus the final step in the degradation of RseA is relatively non-specific, a marked contrast to the initial cleavage events that are wholly dependent on DegS and RseP. Presumably the redundancy in the final cytoplasmic degradation step is important to ensure that σ^E is released, it will be free to direct transcription rather than rebinding to the cytoplasmic domain of RseA.

Step 5: Activation of the σ^E regulon and return to a resting state

Once σ^E is released from RseA, it binds to core RNA polymerase and transcribes the genes in its regulon. Included among these genes are chaperones and proteases that help refold or degrade aberrantly folded OMPs, and sRNAs that target OMP mRNAs for degradation [48, 77, 78, 262]. Together, these regulon members serve to simultaneously restore the OMP folding pathway and prevent continued load on the system from newly synthesized proteins (Fig. C-2, bottom right). As a result, the overall concentration of unfolded OMPs with exposed C-termini decreases and DegS returns to the inactive state, effectively shutting off the proteolytic pathway and the rapid degradation of RseA. Because the σ^E:RseA complex is extremely stable, σ^E will be bound by RseA and the response will rapidly return to basal levels.

A common system for regulation of membrane-localized sigma/anti-sigma modules

σ^E belongs to a large group of sigma factors, the group 4 or extracytoplasmic function (ECF) sigma factors that is widely distributed throughout the bacterial world [33, 40]. Many of these sigma factors are regulated by membrane-bound anti-sigma factors [40], and the regulatory proteolytic scheme used in regulation of σ^E is emerging as a paradigm for the signal transduction pathways governing these systems. The proteolytic pathway controlling the σ^E-dependent stress response of *E. coli* is the best characterized of these signaling systems, especially at the structural and biochemical levels. Studies on other systems are shedding light on the themes and variations associated with the regulatory proteolytic pathways.

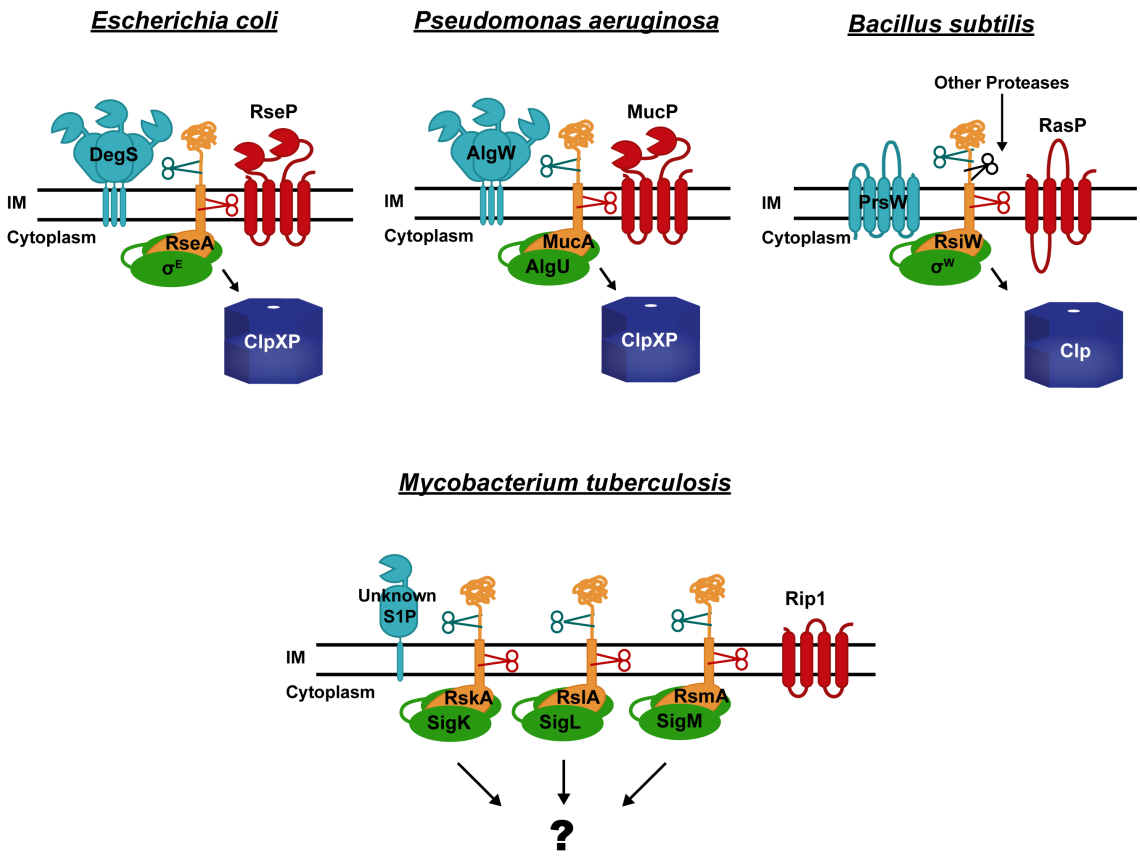


Figure C-8: Sequential proteolysis of transmembrane anti-sigma factors as a regulatory paradigm. The overall design of the regulated proteolytic pathways controlling σ^E in *E. coli*, AlgU in *P. aeruginosa*, σ^W in *B. subtilis*, and SigK, L, and M in *M. tuberculosis* is shown. The proteases on the left perform the first cleavage (blue scissors). The S2P proteases are on the right and perform the intramembrane cleavage step (red scissors).

Regulation of *Pseudomonas aeruginosa* AlgU

In *P. aeruginosa*, the proteolytic pathway governing AlgU is analogous to the σ^E pathway in *E. coli*. Work in this system provides instructive comparisons with the *E. coli* system [109]. AlgU, the σ^E ortholog in *P. aeruginosa*, mediates a cell envelope stress response and transcribes genes that control the expression of the exopolysaccharide alginate [102]. The regulatory pathway controlling AlgU activity is very similar that in *E.*

coli (Fig. C-8). AlgU activity is inhibited by the RseA-like anti-sigma factor MucA [59]. MucA is degraded by a proteolytic cascade that is initiated by AlgW, a DegS homologue [109, 110, 217]. Following AlgW-dependent cleavage, the S2P, MucP, cleaves MucA in its transmembrane domain [110]. The cytoplasmic domain of MucA bound to AlgU is released into the cytoplasm following MucP cleavage, where it is degraded by ClpP, freeing AlgU [351]. MucB, a homologue of RseB, binds to the periplasmic domain of MucA and blocks cleavage by AlgW [109].

Cleavage by AlgW is the most extensively investigated part of the pathway, and there are interesting similarities and differences compared to DegS in *E. coli*. The pathway is initiated in a manner similar to the *E. coli* system. AlgW is inactive until a protein with a C-terminal WVF sequence binds to its PDZ domain [109, 110]. In vitro, a WVF peptide activates AlgW with positive cooperativity [109]. In contrast to *E. coli,* in which a large percentage of OMPs have the activating YxF motif, the only known native protein to end in WVF is the periplasmic protein, MucE. Very little is known about MucE, other than it is an inducer of AlgU. The C termini of two of the major porins in *P. aeruginosa* were not inducers of AlgW, whereas inducing sequences were found at the end of two phosphate/pyrophosphate specific OMPs [110]. The *P. aeruginosa* system therefore, may not be tuned to sense overall OMP folding. Regardless of the source of the activating peptide, the outcome of peptide binding to AlgW's PDZ domain is the same as for DegS, activation of AlgW to initiate the proteolytic cascade.

The PDZ domain of AlgW negatively regulates its activity, similar to that of DegS, but it also appears to be a positive regulator [110, 129]. Deletion of the PDZ domain increased AlgW activity. However, the increase was far less than that seen for DegSΔPDZ,

suggesting that the PDZ domain is required for full activity of AlgW [109]. In addition to regulation via the PDZ domain, a second inhibitory interaction was found for AlgW that is not seen in DegS. AlgW has an extended LA loop in its protease domain and truncations of the loop increased activity [109]. In the crystal structure of the HtrA protease, DegP, from *Thermotoga maritima*, the LA loop blocks access to the active site [352]. Therefore, additional regulators may be needed to alter the conformation of the LA loop to fully activate AlgW, adding another level of regulation to AlgW not found for DegS.

Regulation of *B. subtilis* σ^W

In *B. subtilis*, the ECF sigma factor σ^W is activated by antimicrobial peptides, cell-wall active antibiotics, and alkali shock [353, 354]. It is regulated by the anti-sigma factor RsiW [214]. As with the σ^E system, release of σ^W is controlled by a proteolytic cascade that starts with the inner membrane protease, PrsW, followed by the S2P, RasP, and concluded by ClpP in the cytoplasm (Fig. B-8) [214, 355-357]. The overall layout of the pathway is the same as for the σ^E system, although the first protease to act is not one of the three DegS orthologs in *B. subtilis* [214]. Instead the multi-pass inner membrane protein, PrsW, is thought to perform the first cleavage [355, 356]. PrsW, is not a member of any of the canonical families of proteases and does not have any PDZ domains, although PrsW orthologs are wide-spread [355, 356]. PrsW-dependent cleavage of RsiW occurs following one of the inducing stresses, although the nature of the signal is not known. It has been proposed that degradation of RsiW involves two proteolytic modules [358]. First PrsW cleaves RsiW in the periplasmic domain. Then, other periplasmic proteases trim the C terminus of RsiW down to the point where it becomes a substrate for RasP. RasP next

cleaves RsiW in the transmembrane region, releasing the anti-sigma factor domain with the sequence AAA at its C-terminus. This sequence targets the protein to the Clp protease system, similar to the σ^E and AlgU systems [357].

Regulation of virulence in *Mycobacterium tuberculosis*

In *M tuberculosis*, Rip1, an S2P protease, has been shown to be a major virulence factor [359]. Its role in virulence appears to be attributable, in a large part, to its regulation of three anti-sigma/sigma factor pairs (Fig. C-8) [360]. Deletion of *rip1*, and notably not the other two S2P proteases in the *M. tuberculosis* genome, led to the accumulation of C-terminally truncated anti-sigma factors, similar to those seen for RseA in *E. coli* lacking RseP, MucA in *P. aeruginosa* lacking MucP, and RsiW in *B. subtilis* lacking RasP [110, 217, 360]. Although S2P proteases are known to have other substrates, in addition to anti-sigma factors, Rip1 is the first S2P shown to cleave multiple anti-sigma factors [360]. The initiating protease(s) and inducing signal(s) have not been identified yet for these proteolytic pathways. It will be of interest to learn whether the three systems all use the same initiating protease, such that a single regulatory proteolytic module has been adapted to control three different sigma/anti-sigma factor systems, or if the initiating proteases are unique and integrate distinct inducing signals.

Perspectives

The overall design of this proteolytic system has been adapted to regulate alternative sigma factors with transmembrane anti-sigma factors in different bacteria that have different cellular roles. In keeping with its role as the signal sensor, the first protease

in the pathway is the most variable component of the system, while the other proteases are more highly conserved. Therefore, the overall proteolytic module appears to have evolved to sense different signals by varying the initiating protease. Even when the initiating protease is conserved, for example DegS and AlgW, subtle differences in nature of peptide binding specificity and allosteric activation serve to tailor the protease to the specific needs of the bacterium [109, 327].

While the initiating proteases appear to be variable, the downstream proteases are more highly conserved. The second cleavage is performed by a RIP protease, most often an S2P family member, and a Clp protease degrades the remaining anti-sigma factor domain. In all cases examined thus far, it appears that only the first cleavage is signal-dependent and each cleavage generates a substrate for the next protease in the pathway [184, 214, 217, 322, 323, 351, 357, 360]. As more information is learned from other related regulatory proteolytic cascades, it will be of interest to determine whether the somewhat intricate inhibitory interactions seen in the σ^E system are a hallmark of the overall regulatory design or if they are adaptations specific to the *E. coli* system. Many of these transmembrane anti-sigma factor/ECF sigma systems have an ortholog of RseB [40], suggesting modulation of the proteolytic cascade by RseB-like proteins is a conserved and important part of the regulatory pathways.

Why is such a complex hierarchial proteolytic cascade needed to control the activity of individual alternative sigma factor, like σ^E? From a design perspective, the proteolytic cascade provides a fast response to an inducing signal [317]. No step in the pathway is dependent on the synthesis of the next component, a much slower process, or on additional outside inputs. Once the signaling pathway is triggered, other signals are not

required, ensuring that σ^E is rapidly released to promote cell survival. The numerous inhibitory interactions throughout the system prevent proteolysis in the absence of an inducing signal, yet allow the system to be poised to proceed as soon as a signal occurs. As a result, regulatory proteolysis provides a solution to the transmembrane signaling problem that not only generates and on-off switch, but that can also finely tune biological activity to the strength of the inducing signal.

References

1. Zasloff, M., *Antimicrobial peptides of multicellular organisms.* Nature, 2002. **415**(6870): p. 389-95.

2. Shai, Y., *Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides.* Biochim Biophys Acta, 1999. **1462**(1-2): p. 55-70.

3. Schaible, U.E. and S.H. Kaufmann, *A nutritive view on the host-pathogen interplay.* Trends Microbiol, 2005. **13**(8): p. 373-80.

4. Dahlgren, C., A. Karlsson, and J. Bylund, *Measurement of respiratory burst products generated by professional phagocytes.* Methods Mol Biol, 2007. **412**: p. 349-63.

5. Rowley, G., et al., *Pushing the envelope: extracytoplasmic stress responses in bacterial pathogens.* Nat Rev Microbiol, 2006. **4**(5): p. 383-94.

6. Raivio, T.L., *Envelope stress responses and Gram-negative bacterial pathogenesis.* Mol Microbiol, 2005. **56**(5): p. 1119-28.

7. Silhavy, T.J., D. Kahne, and S. Walker, *The bacterial cell envelope.* Cold Spring Harb Perspect Biol, 2010. **2**(5): p. a000414.

8. Osborn, M.J. and L. Rothfield, *Cell shape determination in Escherichia coli.* Curr Opin Microbiol, 2007. **10**(6): p. 606-10.

9. Ruiz, N. and T.J. Silhavy, *Sensing external stress: watchdogs of the Escherichia coli cell envelope.* Curr Opin Microbiol, 2005. **8**(2): p. 122-6.

10. Bos, M.P., V. Robert, and J. Tommassen, *Biogenesis of the gram-negative bacterial outer membrane.* Annu Rev Microbiol, 2007. **61**: p. 191-214.

11. Kabanov, D.S. and I.R. Prokhorenko, *Structural analysis of lipopolysaccharides from Gram-negative bacteria.* Biochemistry (Mosc), 2010. **75**(4): p. 383-404.

12. Preston, A. and D. Maskell, *The molecular genetics and role in infection of lipopolysaccharide biosynthesis in the Bordetellae.* J Endotoxin Res, 2001. **7**(4): p. 251-61.

13. Trent, M.S., *Biosynthesis, transport, and modification of lipid A.* Biochem Cell Biol, 2004. **82**(1): p. 71-86.

14. Poxton, I.R., *Antibodies to lipopolysaccharide.* J Immunol Methods, 1995. **186**(1): p. 1-15.

15. Kovacs-Simon, A., R.W. Titball, and S.L. Michell, *Lipoproteins of bacterial pathogens.* Infect Immun, 2011. **79**(2): p. 548-61.

16. Pages, J.M., C.E. James, and M. Winterhalter, *The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria.* Nat Rev Microbiol, 2008. **6**(12): p. 893-903.

17. Hayes, C.S., S.K. Aoki, and D.A. Low, *Bacterial contact-dependent delivery systems.* Annu Rev Genet, 2010. **44**: p. 71-90.

18. Cornelis, G.R., *The type III secretion injectisome.* Nat Rev Microbiol, 2006. **4**(11): p. 811-25.

19. Macnab, R.M., *How bacteria assemble flagella.* Annu Rev Microbiol, 2003. **57**: p. 77-100.

20. Ernst, R.K., T. Guina, and S.I. Miller, *Salmonella typhimurium outer membrane remodeling: role in resistance to host innate immunity.* Microbes Infect, 2001. **3**(14-15): p. 1327-34.

21. Wells, T.J., et al., *Autotransporter proteins: novel targets at the bacterial cell surface.* FEMS Microbiol Lett, 2007. **274**(2): p. 163-72.

22. Kline, K.A., et al., *Bacterial adhesins in host-microbe interactions.* Cell Host Microbe, 2009. **5**(6): p. 580-92.

23. MacRitchie, D.M., et al., *Two-component signaling and gram negative envelope stress response systems.* Adv Exp Med Biol, 2008. **631**: p. 80-110.

24. Darwin, A.J., *The phage-shock-protein response.* Mol Microbiol, 2005. **57**(3): p. 621-8.

25. Joly, N., et al., *Managing membrane stress: the phage shock protein (Psp) response, from molecular mechanisms to physiology.* FEMS Microbiol Rev, 2010. **34**(5): p. 797-827.

26. Cotter, P.A. and A.M. Jones, *Phosphorelay control of virulence gene expression in Bordetella.* Trends Microbiol, 2003. **11**(8): p. 367-73.

27. Bock, A. and R. Gross, *The BvgAS two-component system of Bordetella spp.: a versatile modulator of virulence gene expression.* Int J Med Microbiol, 2001. **291**(2): p. 119-30.

28. Majdalani, N. and S. Gottesman, *The Rcs phosphorelay: a complex signal transduction system.* Annu Rev Microbiol, 2005. **59**: p. 379-405.

29. Darst, S.A., *Bacterial RNA polymerase.* Curr Opin Struct Biol, 2001. **11**(2): p. 155-62.

30. Mooney, R.A., S.A. Darst, and R. Landick, *Sigma and RNA polymerase: an on-again, off-again relationship?* Mol Cell, 2005. **20**(3): p. 335-45.

31. Gross, C.A., et al., *The functional and regulatory roles of sigma factors in transcription.* Cold Spring Harb Symp Quant Biol, 1998. **63**: p. 141-55.

32. Kazmierczak, M.J., M. Wiedmann, and K.J. Boor, *Alternative sigma factors and their roles in bacterial virulence.* Microbiol Mol Biol Rev, 2005. **69**(4): p. 527-43.

33. Paget, M.S. and J.D. Helmann, *The sigma70 family of sigma factors.* Genome Biol, 2003. **4**(1): p. 203.

34. Shingler, V., *Signal sensory systems that impact sigma -dependent transcription.* FEMS Microbiol Rev, 2011. **35**(3): p. 425-40.

35. Lonetto, M., M. Gribskov, and C.A. Gross, *The sigma 70 family: sequence conservation and evolutionary relationships.* J Bacteriol, 1992. **174**(12): p. 3843-9.

36. Campbell, E.A., et al., *Structure of the bacterial RNA polymerase promoter specificity sigma subunit.* Mol Cell, 2002. **9**(3): p. 527-39.

37. Murakami, K.S., et al., *Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex.* Science, 2002. **296**(5571): p. 1285-90.

38. Murakami, K.S., S. Masuda, and S.A. Darst, *Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution.* Science, 2002. **296**(5571): p. 1280-4.

39. Helmann, J.D., *The extracytoplasmic function (ECF) sigma factors.* Adv Microb Physiol, 2002. **46**: p. 47-110.

40. Staron, A., et al., *The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family.* Mol Microbiol, 2009. **74**(3): p. 557-81.

41. Glaeser, J., et al., *Singlet oxygen stress in microorganisms.* Adv Microb Physiol, 2011. **58**: p. 141-73.

42. Ziegelhoffer, E.C. and T.J. Donohue, *Bacterial responses to photo-oxidative stress.* Nat Rev Microbiol, 2009. **7**(12): p. 856-63.

43. Hughes, K.T. and K. Mathee, *The anti-sigma factors.* Annu Rev Microbiol, 1998. **52**: p. 231-86.

44. Ades, S.E., *Regulation by destruction: design of the sigmaE envelope stress response.* Curr Opin Microbiol, 2008. **11**(6): p. 535-40.

45. Heinrich, J. and T. Wiegert, *Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors.* Res Microbiol, 2009. **160**(9): p. 696-703.

46. Miticka, H., et al., *Transcriptional analysis of the rpoE gene encoding extracytoplasmic stress response sigma factor sigmaE in Salmonella enterica serovar Typhimurium.* FEMS Microbiol Lett, 2003. **226**(2): p. 307-14.

47. Missiakas, D. and S. Raina, *The extracytoplasmic function sigma factors: role and regulation.* Mol Microbiol, 1998. **28**(6): p. 1059-66.

48. Rhodius, V.A., et al., *Conserved and variable functions of the sigmaE stress response in related genomes.* PLoS Biol, 2006. **4**(1): p. e2.

49. Keith, L.M. and C.L. Bender, *AlgT (sigma22) controls alginate production and tolerance to environmental stress in Pseudomonas syringae.* J Bacteriol, 1999. **181**(23): p. 7176-84.

50. Flannagan, R.S. and M.A. Valvano, *Burkholderia cenocepacia requires RpoE for growth under stress conditions and delay of phagolysosomal fusion in macrophages.* Microbiology, 2008. **154**(Pt 2): p. 643-53.

51. Raivio, T.L. and T.J. Silhavy, *Periplasmic stress and ECF sigma factors.* Annu Rev Microbiol, 2001. **55**: p. 591-624.

52. Crouch, M.L., et al., *The alternative sigma factor sigma is required for resistance of Salmonella enterica serovar Typhimurium to anti-microbial peptides.* Mol Microbiol, 2005. **56**(3): p. 789-99.

53. Alba, B.M. and C.A. Gross, *Regulation of the Escherichia coli sigma-dependent envelope stress response.* Mol Microbiol, 2004. **52**(3): p. 613-9.

54. Ahuja, N., et al., *Analyzing the interaction of RseA and RseB, the two negative regulators of the sigmaE envelope stress response, using a combined bioinformatic and experimental strategy.* J Biol Chem, 2009. **284**(8): p. 5403-13.

55. Collinet, B., et al., *RseB binding to the periplasmic domain of RseA modulates the RseA:sigmaE interaction in the cytoplasm and the availability of sigmaE.RNA polymerase.* J Biol Chem, 2000. **275**(43): p. 33898-904.

56. Missiakas, D., et al., *Modulation of the Escherichia coli sigmaE (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins.* Mol Microbiol, 1997. **24**(2): p. 355-71.

57. De Las Penas, A., L. Connolly, and C.A. Gross, *The sigmaE-mediated response to extracytoplasmic stress in Escherichia coli is transduced by RseA and RseB, two negative regulators of sigmaE.* Mol Microbiol, 1997. **24**(2): p. 373-85.

58. Mathee, K., C.J. McPherson, and D.E. Ohman, *Posttranslational control of the algT (algU)-encoded sigma22 for expression of the alginate regulon in Pseudomonas aeruginosa and localization of its antagonist proteins MucA and MucB (AlgN).* J Bacteriol, 1997. **179**(11): p. 3711-20.

59. Schurr, M.J., et al., *Control of AlgU, a member of the sigma E-like family of stress sigma factors, by the negative regulators MucA and MucB and Pseudomonas aeruginosa conversion to mucoidy in cystic fibrosis.* J Bacteriol, 1996. **178**(16): p. 4997-5004.

60. Bosse, J.T., et al., *Regulation of pga operon expression and biofilm formation in Actinobacillus pleuropneumoniae by sigmaE and H-NS.* J Bacteriol, 2010. **192**(9): p. 2414-23.

61. De Las Penas, A., L. Connolly, and C.A. Gross, *SigmaE is an essential sigma factor in Escherichia coli.* J Bacteriol, 1997. **179**(21): p. 6862-4.

62. Heusipp, G., M.A. Schmidt, and V.L. Miller, *Identification of rpoE and nadB as host responsive elements of Yersinia enterocolitica.* FEMS Microbiol Lett, 2003. **226**(2): p. 291-8.

63. Davis, B.M. and M.K. Waldor, *High-throughput sequencing reveals suppressors of Vibrio cholerae rpoE mutations: one fewer porin is enough.* Nucleic Acids Res, 2009. **37**(17): p. 5757-67.

64. Yu, H., et al., *Virulence properties of Pseudomonas aeruginosa lacking the extreme-stress sigma factor AlgU (sigmaE).* Infect Immun, 1996. **64**(7): p. 2774-81.

65. Testerman, T.L., et al., *The alternative sigma factor sigmaE controls antioxidant defences required for Salmonella virulence and stationary-phase survival.* Mol Microbiol, 2002. **43**(3): p. 771-82.

66. Bashyam, M.D. and S.E. Hasnain, *The extracytoplasmic function sigma factors: role in bacterial pathogenesis.* Infect Genet Evol, 2004. **4**(4): p. 301-8.

67. Erickson, J.W. and C.A. Gross, *Identification of the sigma E subunit of Escherichia coli RNA polymerase: a second alternate sigma factor involved in high-temperature gene expression.* Genes Dev, 1989. **3**(9): p. 1462-71.

68. Mecsas, J., et al., *The activity of sigma E, an Escherichia coli heat-inducible sigma-factor, is modulated by expression of outer membrane proteins.* Genes Dev, 1993. **7**(12B): p. 2618-28.

69. Raina, S., D. Missiakas, and C. Georgopoulos, *The rpoE gene encoding the sigma E (sigma 24) heat shock sigma factor of Escherichia coli.* EMBO J, 1995. **14**(5): p. 1043-55.

70. Rouviere, P.E., et al., *rpoE, the gene encoding the second heat-shock sigma factor, sigma E, in Escherichia coli.* EMBO J, 1995. **14**(5): p. 1032-42.

71. Hayden, J.D. and S.E. Ades, *The Extracytoplasmic Stress Factor, sigma, Is Required to Maintain Cell Envelope Integrity in Escherichia coli.* PLoS ONE, 2008. **3**(2): p. e1573.

72. Mecsas, J., et al., *Identification and characterization of an outer membrane protein, OmpX, in Escherichia coli that is homologous to a family of outer membrane proteins including Ail of Yersinia enterocolitica.* J Bacteriol, 1995. **177**(3): p. 799-804.

73. Walsh, N.P., et al., *OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain.* Cell, 2003. **113**(1): p. 61-71.

74. Ades, S.E., et al., *The Escherichia coli sigma(E)-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-sigma factor.* Genes Dev, 1999. **13**(18): p. 2449-61.

75. Ades, S.E., *Control of the alternative sigma factor sigmaE in Escherichia coli.* Curr Opin Microbiol, 2004. **7**(2): p. 157-62.

76. Ruiz, N., D. Kahne, and T.J. Silhavy, *Advances in understanding bacterial outer-membrane biogenesis.* Nat Rev Microbiol, 2006. **4**(1): p. 57-66.

77. Vogel, J. and K. Papenfort, *Small non-coding RNAs and the bacterial outer membrane.* Curr Opin Microbiol, 2006. **9**(6): p. 605-11.

78. Johansen, J., et al., *Conserved small non-coding RNAs that belong to the sigmaE regulon: role in down-regulation of outer membrane proteins.* J Mol Biol, 2006. **364**(1): p. 1-8.

79. Valentin-Hansen, P., J. Johansen, and A.A. Rasmussen, *Small RNAs controlling outer membrane porins.* Curr Opin Microbiol, 2007. **10**(2): p. 152-5.

80. Gogol, E.B., et al., *Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon.* Proc Natl Acad Sci U S A, 2011. **108**(31): p. 12875-80.

81. Ades, S.E., I.L. Grigorova, and C.A. Gross, *Regulation of the alternative sigma factor sigma(E) during initiation, adaptation, and shutoff of the extracytoplasmic heat shock response in Escherichia coli.* J Bacteriol, 2003. **185**(8): p. 2512-9.

82. Costanzo, A. and S.E. Ades, *Growth phase-dependent regulation of the extracytoplasmic stress factor, sigmaE, by guanosine 3',5'-bispyrophosphate (ppGpp).* J Bacteriol, 2006. **188**(13): p. 4627-34.

83. Chatterji, D. and A.K. Ojha, *Revisiting the stringent response, ppGpp and starvation signaling.* Curr Opin Microbiol, 2001. **4**(2): p. 160-5.

84. Spira, B., N. Silberstein, and E. Yagil, *Guanosine 3',5'-bispyrophosphate (ppGpp) synthesis in cells of Escherichia coli starved for Pi.* J Bacteriol, 1995. **177**(14): p. 4053-8.

85. Paul, B.J., M.B. Berkmen, and R.L. Gourse, *DksA potentiates direct activation of amino acid promoters by ppGpp.* Proc Natl Acad Sci U S A, 2005. **102**(22): p. 7823-8.

86. Paul, B.J., et al., *DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP.* Cell, 2004. **118**(3): p. 311-22.

87. Sze, C.C. and V. Shingler, *The alarmone (p)ppGpp mediates physiological-responsive control at the sigma 54-dependent Po promoter.* Mol Microbiol, 1999. **31**(4): p. 1217-28.

88. Laurie, A.D., et al., *The role of the alarmone (p)ppGpp in sigma N competition for core RNA polymerase.* J Biol Chem, 2003. **278**(3): p. 1494-503.

89. Lange, R., D. Fischer, and R. Hengge-Aronis, *Identification of transcriptional start sites and the role of ppGpp in the expression of rpoS, the structural gene for the sigma S subunit of RNA polymerase in Escherichia coli.* J Bacteriol, 1995. **177**(16): p. 4676-80.

90. Gentry, D.R., et al., *Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp.* J Bacteriol, 1993. **175**(24): p. 7982-9.

91. Jishage, M., et al., *Regulation of sigma factor competition by the alarmone ppGpp.* Genes Dev, 2002. **16**(10): p. 1260-70.

92. Costanzo, A., et al., *ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor sigmaE in Escherichia coli by both direct and indirect mechanisms.* Mol Microbiol, 2008. **67**(3): p. 619-32.

93. Skovierova, H., et al., *Characterization of the sigmaE-dependent rpoEp3 promoter of Salmonella enterica serovar Typhimurium.* FEMS Microbiol Lett, 2006. **261**(1): p. 53-9.

94. Homerova, D., et al., *Characterization of the micA gene encoding a small regulatory sigmaE-dependent RNA in Salmonella enterica serovar Typhimurium.* Folia Microbiol (Praha), 2011. **56**(1): p. 59-65.

95. Muller, C., et al., *Acid stress activation of the sigma(E) stress response in Salmonella enterica serovar Typhimurium.* Mol Microbiol, 2009. **71**(5): p. 1228-38.

96. Kenyon, W.J., et al., *The starvation-stress response of Salmonella enterica serovar Typhimurium requires sigma(E)-, but not CpxR-regulated extracytoplasmic functions.* Microbiology, 2002. **148**(Pt 1): p. 113-22.

97. Humphreys, S., et al., *The alternative sigma factor, sigmaE, is critically important for the virulence of Salmonella typhimurium.* Infect Immun, 1999. **67**(4): p. 1560-8.

98. Kovacikova, G. and K. Skorupski, *The alternative sigma factor sigma(E) plays an important role in intestinal survival and virulence in Vibrio cholerae.* Infect Immun, 2002. **70**(10): p. 5355-62.

99. Mathur, J., B.M. Davis, and M.K. Waldor, *Antimicrobial peptides activate the Vibrio cholerae sigmaE regulon through an OmpU-dependent signalling pathway.* Mol Microbiol, 2007. **63**(3): p. 848-58.

100. Ding, Y., B.M. Davis, and M.K. Waldor, *Hfq is essential for Vibrio cholerae virulence and downregulates sigma expression.* Mol Microbiol, 2004. **53**(1): p. 345-54.

101. Brown, R.N. and P.A. Gulig, *Roles of RseB, sigmaE, and DegP in virulence and phase variation of colony morphotype of Vibrio vulnificus.* Infect Immun, 2009. **77**(9): p. 3768-81.

102. Martin, D.W., B.W. Holloway, and V. Deretic, *Characterization of a locus determining the mucoid status of Pseudomonas aeruginosa: AlgU shows sequence similarities with a Bacillus sigma factor.* J Bacteriol, 1993. **175**(4): p. 1153-64.

103. Martin, D.W., et al., *Mechanism of conversion to mucoidy in Pseudomonas aeruginosa infecting cystic fibrosis patients.* Proc Natl Acad Sci U S A, 1993. **90**(18): p. 8377-81.

104. Deretic, V., et al., *Conversion of Pseudomonas aeruginosa to mucoidy in cystic fibrosis: environmental stress and regulation of bacterial virulence by alternative sigma factors.* J Bacteriol, 1994. **176**(10): p. 2773-80.

105. Hershberger, C.D., et al., *The algT (algU) gene of Pseudomonas aeruginosa, a key regulator involved in alginate biosynthesis, encodes an alternative sigma factor (sigma E).* Proc Natl Acad Sci U S A, 1995. **92**(17): p. 7941-5.

106. Martin, D.W., et al., *Differentiation of Pseudomonas aeruginosa into the alginate-producing form: inactivation of mucB causes conversion to mucoidy.* Mol Microbiol, 1993. **9**(3): p. 497-506.

107. Gill, J.F., V. Deretic, and A.M. Chakrabarty, *Alginate production by the mucoid Pseudomonas aeruginosa associated with cystic fibrosis.* Microbiol Sci, 1987. **4**(10): p. 296-9.

108. Yu, H., M.J. Schurr, and V. Deretic, *Functional equivalence of Escherichia coli sigma E and Pseudomonas aeruginosa AlgU: E. coli rpoE restores mucoidy and reduces sensitivity to reactive oxygen intermediates in algU mutants of P. aeruginosa.* J Bacteriol, 1995. **177**(11): p. 3259-68.

109. Cezairliyan, B.O. and R.T. Sauer, *Control of Pseudomonas aeruginosa AlgW protease cleavage of MucA by peptide signals and MucB.* Mol Microbiol, 2009. **72**(2): p. 368-79.

110. Qiu, D., et al., *Regulated proteolysis controls mucoid conversion in Pseudomonas aeruginosa.* Proc Natl Acad Sci U S A, 2007. **104**(19): p. 8107-12.

111. Deretic, V., et al., *Conversion to mucoidy in Pseudomonas aeruginosa.* Biotechnology (N Y), 1993. **11**(10): p. 1133-6.

112. Hogardt, M. and J. Heesemann, *Adaptation of Pseudomonas aeruginosa during persistence in the cystic fibrosis lung.* Int J Med Microbiol, 2010. **300**(8): p. 557-62.

113. Boucher, J.C., M.J. Schurr, and V. Deretic, *Dual regulation of mucoidy in Pseudomonas aeruginosa and sigma factor antagonism.* Mol Microbiol, 2000. **36**(2): p. 341-51.

114. Jones, A.K., et al., *Activation of the Pseudomonas aeruginosa AlgU regulon through mucA mutation inhibits cyclic AMP/Vfr signaling.* J Bacteriol, 2010. **192**(21): p. 5709-17.

115. Silo-Suh, L., et al., *A simple alfalfa seedling infection model for Pseudomonas aeruginosa strains associated with cystic fibrosis shows AlgT (sigma-22) and RhlR contribute to pathogenesis.* Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15699-704.

116. Firoved, A.M., J.C. Boucher, and V. Deretic, *Global genomic analysis of AlgU (sigma(E))-dependent promoters (sigmulon) in Pseudomonas aeruginosa and implications for inflammatory processes in cystic fibrosis.* J Bacteriol, 2002. **184**(4): p. 1057-64.

117. Wu, W., et al., *MucA-mediated coordination of type III secretion and alginate synthesis in Pseudomonas aeruginosa.* J Bacteriol, 2004. **186**(22): p. 7575-85.

118. Damron, F.H., et al., *Lipotoxin F of Pseudomonas aeruginosa is an AlgU-dependent and alginate-independent outer membrane protein involved in resistance to oxidative stress and adhesion to A549 human lung epithelia.* Microbiology, 2009. **155**(Pt 4): p. 1028-38.

119. Potvin, E., F. Sanschagrin, and R.C. Levesque, *Sigma factors in Pseudomonas aeruginosa.* FEMS Microbiol Rev, 2008. **32**(1): p. 38-55.

120. Schenk, A., et al., *The algT gene of Pseudomonas syringae pv. glycinea and new insights into the transcriptional organization of the algT-muc gene cluster.* J Bacteriol, 2006. **188**(23): p. 8013-21.

121. Schenk, A., H. Weingart, and M.S. Ullrich, *The alternative sigma factor AlgT, but not alginate synthesis, promotes in planta multiplication of Pseudomonas syringae pv. glycinea.* Microbiology, 2008. **154**(Pt 2): p. 413-21.

122. Schnider-Keel, U., et al., *The sigma factor AlgU (AlgT) controls exopolysaccharide production and tolerance towards desiccation and osmotic stress in the biocontrol agent Pseudomonas fluorescens CHA0.* Appl Environ Microbiol, 2001. **67**(12): p. 5683-93.

123. Korbsrisate, S., et al., *The Burkholderia pseudomallei RpoE (AlgU) operon is involved in environmental stress tolerance and biofilm formation.* FEMS Microbiol Lett, 2005. **252**(2): p. 243-9.

124. Vanaporn, M., et al., *The rpoE operon regulates heat stress response in Burkholderia pseudomallei.* FEMS Microbiol Lett, 2008. **284**(2): p. 191-6.

125. Craig, J.E., A. Nobbs, and N.J. High, *The extracytoplasmic sigma factor, final sigma(E), is required for intracellular survival of nontypeable Haemophilus influenzae in J774 macrophages.* Infect Immun, 2002. **70**(2): p. 708-15.

126. Govan, J.R. and V. Deretic, *Microbial pathogenesis in cystic fibrosis: mucoid Pseudomonas aeruginosa and Burkholderia cepacia.* Microbiol Rev, 1996. **60**(3): p. 539-74.

127. Bordes, P., et al., *Insights into the extracytoplasmic stress response of Xanthomonas campestris pv. campestris: role and regulation of {sigma}E-dependent activity.* J Bacteriol, 2011. **193**(1): p. 246-64.

128. Cheng, C.Y., et al., *Characterization and transcriptional analysis of an ECF sigma factor from Xanthomonas campestris pv. campestris.* FEMS Microbiol Lett, 2008. **289**(2): p. 250-7.

129. Cezairliyan, B.O. and R.T. Sauer, *Inhibition of regulated proteolysis by RseB.* Proc Natl Acad Sci U S A, 2007. **104**(10): p. 3771-6.

130. da Silva Neto, J.F., et al., *The single extracytoplasmic-function sigma factor of Xylella fastidiosa is involved in the heat shock response and presents an unusual regulatory mechanism.* J Bacteriol, 2007. **189**(2): p. 551-60.

131. Heungens, K., C.E. Cowles, and H. Goodrich-Blair, *Identification of Xenorhabdus nematophila genes required for mutualistic colonization of Steinernema carpocapsae nematodes.* Mol Microbiol, 2002. **45**(5): p. 1337-53.

132. Nunez, C., et al., *Role of Azotobacter vinelandii mucA and mucC gene products in alginate production.* J Bacteriol, 2000. **182**(23): p. 6550-6.

133. Moreno, S., et al., *Role of alternative sigma factor algU in encystment of Azotobacter vinelandii.* J Bacteriol, 1998. **180**(10): p. 2766-9.

134. Sheehan, B.J., et al., *Identification of Actinobacillus pleuropneumoniae genes important for survival during infection in its natural host.* Infect Immun, 2003. **71**(7): p. 3960-70.

135. Arico, B., et al., *Evolutionary relationships in the genus Bordetella.* Mol Microbiol, 1987. **1**(3): p. 301-8.

136. Mattoo, S. and J.D. Cherry, *Molecular pathogenesis, epidemiology, and clinical manifestations of respiratory infections due to Bordetella pertussis and other Bordetella subspecies.* Clin Microbiol Rev, 2005. **18**(2): p. 326-82.

137. Mazumder, S.A. and K.O. Cleveland, *Bordetella bronchiseptica bacteremia in a patient with AIDS.* South Med J, 2010. **103**(9): p. 934-5.

138. Musser, J.M., et al., *Clonal diversity and host distribution in Bordetella bronchiseptica.* J Bacteriol, 1987. **169**(6): p. 2793-803.

139. Parkhill, J., et al., *Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella parapertussis and Bordetella bronchiseptica.* Nat Genet, 2003. **35**(1): p. 32-40.

140. Diavatopoulos, D.A., et al., *Bordetella pertussis, the causative agent of whooping cough, evolved from a distinct, human-associated lineage of B. bronchiseptica.* PLoS Pathog, 2005. **1**(4): p. e45.

141. Cummings, C.A., et al., *Species- and strain-specific control of a complex, flexible regulon by Bordetella BvgAS.* J Bacteriol, 2006. **188**(5): p. 1775-85.

142. Williams, C.L. and P.A. Cotter, *Autoregulation is essential for precise temporal and steady-state regulation by the Bordetella BvgAS phosphorelay.* J Bacteriol, 2007. **189**(5): p. 1974-82.

143. Mattoo, S., et al., *Mechanisms of Bordetella pathogenesis.* Front Biosci, 2001. **6**: p. E168-86.

144. Kirimanjeswara, G.S., et al., *The complex mechanism of antibody-mediated clearance of Bordetella from the lungs requires TLR4.* J Immunol, 2005. **175**(11): p. 7504-11.

145. Irie, Y. and M.H. Yuk, *In vivo colonization profile study of Bordetella bronchiseptica in the nasal cavity.* FEMS Microbiol Lett, 2007. **275**(2): p. 191-8.

146. Brockmeier, S.L., *Early colonization of the rat upper respiratory tract by temperature modulated Bordetella bronchiseptica.* FEMS Microbiol Lett, 1999. **174**(2): p. 225-9.

147. Mattoo, S., J.F. Miller, and P.A. Cotter, *Role of Bordetella bronchiseptica fimbriae in tracheal colonization and development of a humoral immune response.* Infect Immun, 2000. **68**(4): p. 2024-33.

148. Burns, V.C., et al., *Role of Bordetella O antigen in respiratory tract infection.* Infect Immun, 2003. **71**(1): p. 86-94.

149. Harvill, E.T., et al., *Probing the function of Bordetella bronchiseptica adenylate cyclase toxin by manipulating host immunity.* Infect Immun, 1999. **67**(3): p. 1493-500.

150. Mann, P.B., M.J. Kennett, and E.T. Harvill, *Toll-like receptor 4 is critical to innate host defense in a murine model of bordetellosis.* J Infect Dis, 2004. **189**(5): p. 833-6.

151. Mann, P.B., et al., *Comparative toll-like receptor 4-mediated innate host defense to Bordetella infection.* Infect Immun, 2005. **73**(12): p. 8144-52.

152. Fung, K.S., et al., *Pertussis--a re-emerging infection?* J Infect, 2004. **48**(2): p. 145-8.

153. *Pertussis outbreak in California.* Hum Vaccin, 2010. **6**(9).

154. Jongerius, I., S. Ram, and S. Rooijakkers, *Bacterial complement escape.* Adv Exp Med Biol, 2009. **666**: p. 32-48.

155. Mathur, J. and M.K. Waldor, *The Vibrio cholerae ToxR-regulated porin OmpU confers resistance to antimicrobial peptides.* Infect Immun, 2004. **72**(6): p. 3577-83.

156. Goodnow, R.A., *Biology of Bordetella bronchiseptica.* Microbiol Rev, 1980. **44**(4): p. 722-38.

157. Madan Babu, M., S.A. Teichmann, and L. Aravind, *Evolutionary dynamics of prokaryotic transcriptional regulatory networks.* J Mol Biol, 2006. **358**(2): p. 614-33.

158. Gruber, T.M. and C.A. Gross, *Multiple sigma subunits and the partitioning of bacterial transcription space.* Annu Rev Microbiol, 2003. **57**: p. 441-66.

159. Rowen, D.W. and V. Deretic, *Membrane-to-cytosol redistribution of ECF sigma factor AlgU and conversion to mucoidy in Pseudomonas aeruginosa isolates from cystic fibrosis patients.* Mol Microbiol, 2000. **36**(2): p. 314-27.

160. Bianchi, A.A. and F. Baneyx, *Hyperosmotic shock induces the sigma32 and sigmaE stress regulons of Escherichia coli.* Mol Microbiol, 1999. **34**(5): p. 1029-38.

161. Tomoyasu, T., et al., *Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the Escherichia coli cytosol.* Mol Microbiol, 2001. **40**(2): p. 397-413.

162. Mann, P.B., et al., *Toll-like receptor 4-dependent early elicited tumor necrosis factor alpha expression is critical for innate host defense against Bordetella bronchiseptica.* Infect Immun, 2004. **72**(11): p. 6650-8.

163. Yuk, M.H., E.T. Harvill, and J.F. Miller, *The BvgAS virulence control system regulates type III secretion in Bordetella bronchiseptica.* Mol Microbiol, 1998. **28**(5): p. 945-59.

164. Gerlach, G., et al., *Evolutionary trends in the genus Bordetella.* Microbes Infect, 2001. **3**(1): p. 61-72.

165. Porter, J.F., R. Parton, and A.C. Wardlaw, *Growth and survival of Bordetella bronchiseptica in natural waters and in buffered saline without added nutrients.* Appl Environ Microbiol, 1991. **57**(4): p. 1202-6.

166. Park, S.D., et al., *Corynebacterium glutamicum sigmaE is involved in responses to cell surface stresses and its activity is controlled by the anti-sigma factor CseE.* Microbiology, 2008. **154**(Pt 3): p. 915-23.

167. Cotter, P.A. and J.F. Miller, *BvgAS-mediated signal transduction: analysis of phase-locked regulatory mutants of Bordetella bronchiseptica in a rabbit model.* Infect Immun, 1994. **62**(8): p. 3381-90.

168. Stainer, D.W. and M.J. Scholte, *A simple chemically defined medium for the production of phase I Bordetella pertussis.* J Gen Microbiol, 1970. **63**(2): p. 211-20.

169. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.

170. Weingart, C.L., et al., *Fluorescent labels influence phagocytosis of Bordetella pertussis by human neutrophils.* Infect Immun, 1999. **67**(8): p. 4264-7.

171. Stibitz, S., et al., *The vir locus and phase-variation in Bordetella pertussis.* Tokai J Exp Clin Med, 1988. **13 Suppl**: p. 223-6.

172. Preston, A., et al., *Genetic basis for lipopolysaccharide O-antigen biosynthesis in bordetellae.* Infect Immun, 1999. **67**(8): p. 3763-7.

173. Mann, P., et al., *Use of a genetically defined double mutant strain of Bordetella bronchiseptica lacking adenylate cyclase and type III secretion as a live vaccine.* Infect Immun, 2007. **75**(7): p. 3665-72.

174. Schaeffer, L.M., et al., *Bordetella pertussis lipopolysaccharide resists the bactericidal effects of pulmonary surfactant protein A.* J Immunol, 2004. **173**(3): p. 1959-65.

175. Buboltz, A.M., et al., *Role of the type III secretion system in a hypervirulent lineage of Bordetella bronchiseptica.* Infect Immun, 2009. **77**(9): p. 3969-77.

176. Stibitz, S. and N.H. Carbonetti, *Hfr mapping of mutations in Bordetella pertussis that define a genetic locus involved in virulence gene regulation.* J Bacteriol, 1994. **176**(23): p. 7260-6.

177. Miller, J.H., *Experiments in molecular genetics.* 1972.

178. Goebel, E.M., et al., *O-antigen Protects Bordetella parapertussis from Complement.* Infect Immun, 2008.

179. Rodriguez, M.E., et al., *Fc receptor-mediated immunity against Bordetella pertussis.* J Immunol, 2001. **167**(11): p. 6545-51.

180. Rodriguez, M.E., W.L. Van der Pol, and J.G. Van de Winkel, *Flow cytometry-based phagocytosis assay for sensitive detection of opsonic activity of pneumococcal capsular polysaccharide antibodies in human sera.* J Immunol Methods, 2001. **252**(1-2): p. 33-44.

181. Harvill, E.T., et al., *Multiple roles for Bordetella lipopolysaccharide molecules during respiratory tract infection.* Infect Immun, 2000. **68**(12): p. 6720-8.

182. Nicholson, T.L., *Construction and validation of a first-generation Bordetella bronchiseptica long-oligonucleotide microarray by transcriptional profiling the Bvg regulon.* BMC Genomics, 2007. **8**: p. 220.

183. Schurr, M.J., et al., *Multiple promoters and induction by heat shock of the gene encoding the alternative sigma factor AlgU (sigma E) which controls mucoidy in cystic fibrosis isolates of Pseudomonas aeruginosa.* J Bacteriol, 1995. **177**(19): p. 5670-9.

184. Flynn, J.M., et al., *Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals.* Mol Cell, 2003. **11**(3): p. 671-83.

185. Alba, B.M., et al., *DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigma(E)-dependent extracytoplasmic stress response.* Genes Dev, 2002. **16**(16): p. 2156-68.

186. Arnold, K., et al., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.

187. Kiefer, F., et al., *The SWISS-MODEL Repository and associated resources.* Nucleic Acids Res, 2009. **37**(Database issue): p. D387-92.

188. Peitsch, M., *Protein modeling by E-mail.* Bio/Technology, 1995. **13**: p. 658-660.

189. Campbell, E.A., et al., *Crystal structure of Escherichia coli sigmaE with the cytoplasmic domain of its anti-sigma RseA.* Mol Cell, 2003. **11**(4): p. 1067-78.

190. Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.3r1*, 2010.

191. Inatsuka, C.S., et al., *Pertactin is required for Bordetella species to resist neutrophil-mediated clearance.* Infect Immun, 2010. **78**(7): p. 2901-9.

192. Zaat, S.A., et al., *Induction of the nodA promoter of Rhizobium leguminosarum Sym plasmid pRL1JI by plant flavanones and flavones.* J Bacteriol, 1987. **169**(1): p. 198-204.

193. Forde, C.B., et al., *Bordetella bronchiseptica-mediated cytotoxicity to macrophages is dependent on bvg-regulated factors, including pertactin.* Infect Immun, 1999. **67**(11): p. 5972-8.

194. Helmann, J.D., *Anti-sigma factors.* Curr Opin Microbiol, 1999. **2**(2): p. 135-41.

195. Kovach, M.E., et al., *Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes.* Gene, 1995. **166**(1): p. 175-6.

196. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* J Bacteriol, 1995. **177**(14): p. 4121-30.

197. Meltzer, M., et al., *Structure, function and regulation of the conserved serine proteases DegP and DegS of Escherichia coli.* Res Microbiol, 2009. **160**(9): p. 660-6.

198. Shrivastava, R. and J.F. Miller, *Virulence factor secretion and translocation by Bordetella species.* Curr Opin Microbiol, 2009. **12**(1): p. 88-93.

199. Maclellan, S.R., W. Eiamphungporn, and J.D. Helmann, *ROMA: an in vitro approach to defining target genes for transcription regulators.* Methods, 2009. **47**(1): p. 73-7.

200. Cao, M., et al., *Defining the Bacillus subtilis sigma(W) regulon: a comparative analysis of promoter consensus search, run-off transcription/macroarray analysis*

(ROMA), and transcriptional profiling approaches. J Mol Biol, 2002. **316**(3): p. 443-57.

201. Eiamphungporn, W. and J.D. Helmann, *The Bacillus subtilis sigma(M) regulon and its contribution to cell envelope stress responses.* Mol Microbiol, 2008. **67**(4): p. 830-48.

202. Armstrong, S.K., et al., *Bordetella pertussis major outer membrane porin protein forms small, anion-selective channels in lipid bilayer membranes.* J Bacteriol, 1986. **166**(1): p. 212-6.

203. Finn, T.M., Z. Li, and E. Kocsis, *Identification of a Bordetella pertussis bvg-regulated porin-like protein.* J Bacteriol, 1995. **177**(3): p. 805-9.

204. Kocsis, E., et al., *Orientation of porin channels in the outer membrane of Bordetella pertussis.* Mol Microbiol, 1993. **9**(3): p. 469-76.

205. Delcour, A.H., *Outer membrane permeability and antibiotic resistance.* Biochim Biophys Acta, 2009. **1794**(5): p. 808-16.

206. Preston, A., et al., *Bordetella bronchiseptica PagP is a Bvg-regulated lipid A palmitoyl transferase that is required for persistent colonization of the mouse respiratory tract.* Mol Microbiol, 2003. **48**(3): p. 725-36.

207. Ross, W., et al., *E.coli Fis protein activates ribosomal RNA transcription in vitro and in vivo.* EMBO J, 1990. **9**(11): p. 3733-42.

208. Buboltz, A.M., et al., *Replacement of adenylate cyclase toxin in a lineage of Bordetella bronchiseptica.* J Bacteriol, 2008. **190**(15): p. 5502-11.

209. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.

210. Nicholson, T.L., et al., *Microarray and functional analysis of growth phase-dependent gene regulation in Bordetella bronchiseptica.* Infect Immun, 2009. **77**(10): p. 4221-31.

211. Sojda, J., 3rd, et al., *A rhizobial homolog of IHF stimulates transcription of dctA in Rhizobium leguminosarum but not in Sinorhizobium meliloti.* Gene, 1999. **238**(2): p. 489-500.

212. Goodrich, J.A., M.L. Schwartz, and W.R. McClure, *Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for Escherichia coli integration host factor (IHF).* Nucleic Acids Res, 1990. **18**(17): p. 4993-5000.

213. Barchinger, S.E., et al., *sigE facilitates the adaptation of Bordetella bronchiseptica to stress conditions and lethal infection in mice*, 2011, Penn State University.

214. Schobel, S., et al., *The Bacillus subtilis sigmaW anti-sigma factor RsiW is degraded by intramembrane proteolysis through YluC.* Mol Microbiol, 2004. **52**(4): p. 1091-105.

215. Alba, B.M., et al., *degS (hhoB) is an essential Escherichia coli gene whose indispensable function is to provide sigma (E) activity.* Mol Microbiol, 2001. **40**(6): p. 1323-33.

216. Kanehara, K., Y. Akiyama, and K. Ito, *Characterization of the yaeL gene product and its S2P-protease motifs in Escherichia coli.* Gene, 2001. **281**(1-2): p. 71-9.

217. Wood, L.F., A.J. Leech, and D.E. Ohman, *Cell wall-inhibitory antibiotics activate the alginate biosynthesis operon in Pseudomonas aeruginosa: Roles of sigma (AlgT) and the AlgW and Prc proteases.* Mol Microbiol, 2006. **62**(2): p. 412-26.

218. Mo, E., et al., *Single, double and triple mutants of Salmonella enterica serovar Typhimurium degP (htrA), degQ (hhoA) and degS (hhoB) have diverse phenotypes on exposure to elevated temperature and their growth in vivo is attenuated to different extents.* Microb Pathog, 2006. **41**(4-5): p. 174-82.

219. Sohn, J. and R.T. Sauer, *OMP peptides modulate the activity of DegS protease by differential binding to active and inactive conformations.* Mol Cell, 2009. **33**(1): p. 64-74.

220. Dalebroux, Z.D., et al., *ppGpp conjures bacterial virulence.* Microbiol Mol Biol Rev, 2010. **74**(2): p. 171-99.

221. Potrykus, K. and M. Cashel, *(p)ppGpp: still magical?* Annu Rev Microbiol, 2008. **62**: p. 35-51.

222. Coote, J.G., *Environmental sensing mechanisms in Bordetella.* Adv Microb Physiol, 2001. **44**: p. 141-81.

223. Shimoni, Z., et al., *Fatal pneumonia due to Bordetella bronchiseptica.* Isr Med Assoc J, 2000. **2**(5): p. 402-3.

224. Letcher, J., E. Weisenberg, and A. Jonas, *Bordetella bronchiseptica pneumonia in a koala.* J Am Vet Med Assoc, 1993. **202**(6): p. 985-7.

225. Banemann, A. and R. Gross, *Phase variation affects long-term survival of Bordetella bronchiseptica in professional phagocytes.* Infect Immun, 1997. **65**(8): p. 3469-73.

226. Udekwu, K.I. and E.G. Wagner, *Sigma E controls biogenesis of the antisense RNA MicA.* Nucleic Acids Res, 2007. **35**(4): p. 1279-88.

227. Burgos, J.M., N.D. King-Lyons, and T.D. Connell, *Expression of BfrH, a putative siderophore receptor of Bordetella bronchiseptica, is regulated by iron, Fur1, and the extracellular function sigma factor EcfI.* Infect Immun, 2010. **78**(3): p. 1147-62.

228. Pradel, E. and C. Locht, *Expression of the putative siderophore receptor gene bfrZ is controlled by the extracytoplasmic-function sigma factor BupI in Bordetella bronchiseptica.* J Bacteriol, 2001. **183**(9): p. 2910-7.

229. Vanderpool, C.K. and S.K. Armstrong, *Heme-responsive transcriptional activation of Bordetella bhu genes.* J Bacteriol, 2003. **185**(3): p. 909-17.

230. Mattoo, S., et al., *Regulation of type III secretion in Bordetella.* Mol Microbiol, 2004. **52**(4): p. 1201-14.

231. Burger, M., et al., *Temperature regulation of protease in Pseudomonas fluorescens LS107d2 by an ECF sigma factor and a transmembrane activator.* Microbiology, 2000. **146 Pt 12**: p. 3149-55.

232. Sayed, D. and M. Abdellatif, *MicroRNAs in development and disease.* Physiol Rev, 2011. **91**(3): p. 827-87.

233. Li, L. and Y. Liu, *Diverse small non-coding RNAs in RNA interference pathways.* Methods Mol Biol, 2011. **764**: p. 169-82.

234. Papenfort, K. and J. Vogel, *Regulatory RNA in bacterial pathogens.* Cell Host Microbe, 2010. **8**(1): p. 116-27.

235. Storz, G., J. Vogel, and K.M. Wassarman, *Regulation by small RNAs in bacteria: expanding frontiers.* Mol Cell, 2011. **43**(6): p. 880-91.

236. Waters, L.S. and G. Storz, *Regulatory RNAs in bacteria.* Cell, 2009. **136**(4): p. 615-28.

237. Gottesman, S. and G. Storz, *Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations.* Cold Spring Harb Perspect Biol, 2010.

238. Breaker, R.R., *Prospects for riboswitch discovery and analysis.* Mol Cell, 2011. **43**(6): p. 867-79.

239. Lee, E.J. and E.A. Groisman, *An antisense RNA that governs the expression kinetics of a multifunctional virulence gene.* Mol Microbiol, 2010. **76**(4): p. 1020-33.

240. Babitzke, P. and T. Romeo, *CsrB sRNA family: sequestration of RNA-binding regulatory proteins.* Curr Opin Microbiol, 2007. **10**(2): p. 156-63.

241. Repoila, F., N. Majdalani, and S. Gottesman, *Small non-coding RNAs, co-ordinators of adaptation processes in Escherichia coli: the RpoS paradigm.* Mol Microbiol, 2003. **48**(4): p. 855-61.

242. Gottesman, S., *The small RNA regulators of Escherichia coli: roles and mechanisms*.* Annu Rev Microbiol, 2004. **58**: p. 303-28.

243. Svenningsen, S.L., K.C. Tu, and B.L. Bassler, *Gene dosage compensation calibrates four regulatory RNAs to control Vibrio cholerae quorum sensing.* EMBO J, 2009. **28**(4): p. 429-39.

244. Tu, K.C., et al., *Negative feedback loops involving small regulatory RNAs precisely control the Vibrio harveyi quorum-sensing response.* Mol Cell, 2010. **37**(4): p. 567-79.

245. Richards, G.R. and C.K. Vanderpool, *Molecular call and response: the physiology of bacterial small RNAs.* Biochim Biophys Acta, 2011. **1809**(10): p. 525-31.

246. Gorke, B. and J. Vogel, *Noncoding RNA control of the making and breaking of sugars.* Genes Dev, 2008. **22**(21): p. 2914-25.

247. Vanderpool, C.K., *Physiological consequences of small RNA-mediated regulation of glucose-phosphate stress.* Curr Opin Microbiol, 2007. **10**(2): p. 146-51.

248. Vanderpool, C.K. and S. Gottesman, *The novel transcription factor SgrR coordinates the response to glucose-phosphate stress.* J Bacteriol, 2007. **189**(6): p. 2238-48.

249. Chao, Y. and J. Vogel, *The role of Hfq in bacterial pathogens.* Curr Opin Microbiol, 2010. **13**(1): p. 24-33.

250. Moon, K. and S. Gottesman, *A PhoQ/P-regulated small RNA regulates sensitivity of Escherichia coli to antimicrobial peptides.* Mol Microbiol, 2009. **74**(6): p. 1314-30.

251. Valentin-Hansen, P., M. Eriksen, and C. Udesen, *The bacterial Sm-like protein Hfq: a key player in RNA transactions.* Mol Microbiol, 2004. **51**(6): p. 1525-33.

252. Vogel, J. and B.F. Luisi, *Hfq and its constellation of RNA.* Nat Rev Microbiol, 2011. **9**(8): p. 578-89.

253. Sonnleitner, E., et al., *Reduced virulence of a hfq mutant of Pseudomonas aeruginosa O1.* Microb Pathog, 2003. **35**(5): p. 217-28.

254. Sittka, A., et al., *The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium.* Mol Microbiol, 2007. **63**(1): p. 193-217.

255. Shakhnovich, E.A., B.M. Davis, and M.K. Waldor, *Hfq negatively regulates type III secretion in EHEC and several other pathogens.* Mol Microbiol, 2009. **74**(2): p. 347-63.

256. Papenfort, K. and J. Vogel, *Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level.* Res Microbiol, 2009. **160**(4): p. 278-87.

257. Hot, D., et al., *Detection of small RNAs in Bordetella pertussis and identification of a novel repeated genetic element.* BMC Genomics, 2011. **12**: p. 207.

258. Livny, J., et al., *High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs.* PLoS ONE, 2008. **3**(9): p. e3197.

259. Havecker, E.R., *Detection of small RNAs and microRNAs using deep sequencing technology.* Methods Mol Biol, 2011. **732**: p. 55-68.

260. Georg, J. and W.R. Hess, *cis-antisense RNA, another level of gene regulation in bacteria.* Microbiol Mol Biol Rev, 2011. **75**(2): p. 286-300.

261. Tjaden, B., *TargetRNA: a tool for predicting targets of small RNA action in bacteria.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W109-13.

262. Thompson, K.M., V.A. Rhodius, and S. Gottesman, *SigmaE regulates and is regulated by a small RNA in Escherichia coli.* J Bacteriol, 2007. **189**(11): p. 4243-56.

263. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.

264. Goujon, M., et al., *A new bioinformatics analysis tools framework at EMBL-EBI.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W695-9.

265. Goecks, J., A. Nekrutenko, and J. Taylor, *Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences.* Genome Biol, 2010. **11**(8): p. R86.

266. Blankenberg, D., et al., *Galaxy: a web-based genome analysis tool for experimentalists.* Curr Protoc Mol Biol, 2010. **Chapter 19**: p. Unit 19 10 1-21.

267. Giardine, B., et al., *Galaxy: a platform for interactive large-scale genome analysis.* Genome Res, 2005. **15**(10): p. 1451-5.

268. Kanehara, K., K. Ito, and Y. Akiyama, *YaeL (EcfE) activates the sigma(E) pathway of stress response through a site-2 cleavage of anti-sigma(E), RseA.* Genes Dev, 2002. **16**(16): p. 2147-55.

269. Shingler, V., *Signal sensing by sigma 54-dependent regulators: derepression as a control mechanism.* Mol Microbiol, 1996. **19**(3): p. 409-16.

270. Hengge-Aronis, R., *Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase.* Microbiol Mol Biol Rev, 2002. **66**(3): p. 373-95, table of contents.

271. Magnusson, L.U., A. Farewell, and T. Nystrom, *ppGpp: a global regulator in Escherichia coli.* Trends Microbiol, 2005. **13**(5): p. 236-42.

272. Cashel, M., Gentry, D.R., Hernandez, V.J., and Vinella, D. , *The Stringent Response*, in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, R.C.I. Frederick C. Neidhardt, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, H.E. Umbarger, Editor 1996, ASM Press: Washington, D.C. p. 1458-1496.

273. Xiao, H., et al., *Residual guanosine 3',5'-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations.* J Biol Chem, 1991. **266**(9): p. 5980-90.

274. Ryals, J., R. Little, and H. Bremer, *Control of rRNA and tRNA syntheses in Escherichia coli by guanosine tetraphosphate.* J Bacteriol, 1982. **151**(3): p. 1261-8.

275. Mallik, P., et al., *DksA is required for growth phase-dependent regulation, growth rate-dependent control, and stringent control of fis expression in Escherichia coli.* J Bacteriol, 2006. **188**(16): p. 5775-82.

276. Perederina, A., et al., *Regulation through the secondary channel--structural framework for ppGpp-DksA synergism during transcription.* Cell, 2004. **118**(3): p. 297-309.

277. Paul, B.J., et al., *rRNA transcription in Escherichia coli.* Annu Rev Genet, 2004. **38**: p. 749-70.

278. Vrentas, C.E., et al., *Response of RNA polymerase to ppGpp: requirement for the omega subunit and relief of this requirement by DksA.* Genes Dev, 2005. **19**(19): p. 2378-87.

279. Rutherford, S.T., et al., *Effects of DksA, GreA, and GreB on transcription initiation: insights into the mechanisms of factors that bind in the secondary channel of RNA polymerase.* J Mol Biol, 2007. **366**(4): p. 1243-57.

280. Kvint, K., A. Farewell, and T. Nystrom, *RpoS-dependent promoters require guanosine tetraphosphate for induction even in the presence of high levels of sigma(s).* J Biol Chem, 2000. **275**(20): p. 14795-8.

281. Bernardo, L.M., et al., *The guanosine tetraphosphate (ppGpp) alarmone, DksA and promoter affinity for RNA polymerase in regulation of sigma-dependent transcription.* Mol Microbiol, 2006. **60**(3): p. 749-64.

282. Szalewska-Palasz, A., et al., *Properties of RNA polymerase bypass mutants: implications for the role of ppGpp and its co-factor DksA in controlling transcription dependent on sigma54.* J Biol Chem, 2007. **282**(25): p. 18046-56.

283. Svitil, A.L., M. Cashel, and J.W. Zyskind, *Guanosine tetraphosphate inhibits protein synthesis in vivo. A possible protective mechanism for starvation stress in Escherichia coli.* J Biol Chem, 1993. **268**(4): p. 2307-11.

284. Gentry, D., et al., *The omega subunit of Escherichia coli K-12 RNA polymerase is not required for stringent RNA control in vivo.* J Bacteriol, 1991. **173**(12): p. 3901-3.

285. Hernandez, V.J. and M. Cashel, *Changes in conserved region 3 of Escherichia coli sigma 70 mediate ppGpp-dependent functions in vivo.* J Mol Biol, 1995. **252**(5): p. 536-49.

286. Sharp, M.M., et al., *The interface of sigma with core RNA polymerase is extensive, conserved, and functionally specialized.* Genes Dev, 1999. **13**(22): p. 3015-26.

287. Gruber, T.M., et al., *Binding of the initiation factor sigma(70) to core RNA polymerase is a multistep process.* Mol Cell, 2001. **8**(1): p. 21-31.

288. Jishage, M. and A. Ishihama, *Transcriptional organization and in vivo role of the Escherichia coli rsd gene, encoding the regulator of RNA polymerase sigma D.* J Bacteriol, 1999. **181**(12): p. 3768-76.

289. Bartlett, M.S., et al., *RNA polymerase mutants that destabilize RNA polymerase-promoter complexes alter NTP-sensing by rrn P1 promoters.* J Mol Biol, 1998. **279**(2): p. 331-45.

290. Zhou, Y.N. and D.J. Jin, *The rpoB mutants destabilizing initiation complexes at stringently controlled promoters behave like "stringent" RNA polymerases in Escherichia coli.* Proc Natl Acad Sci U S A, 1998. **95**(6): p. 2908-13.

291. Barker, M.M., T. Gaal, and R.L. Gourse, *Mechanism of regulation of transcription initiation by ppGpp. II. Models for positive control based on properties of RNAP mutants and competition for RNAP.* J Mol Biol, 2001. **305**(4): p. 689-702.

292. Barker, M.M., et al., *Mechanism of regulation of transcription initiation by ppGpp. I. Effects of ppGpp on transcription initiation in vivo and in vitro.* J Mol Biol, 2001. **305**(4): p. 673-88.

293. Nystrom, T., *Growth versus maintenance: a trade-off dictated by RNA polymerase availability and sigma factor competition?* Mol Microbiol, 2004. **54**(4): p. 855-62.

294. Gourse, R.L., W. Ross, and S.T. Rutherford, *General pathway for turning on promoters transcribed by RNA polymerases containing alternative sigma factors.* J Bacteriol, 2006. **188**(13): p. 4589-91.

295. Haugen, S.P., et al., *rRNA promoter regulation by nonoptimal binding of sigma region 1.2: an additional recognition element for RNA polymerase.* Cell, 2006. **125**(6): p. 1069-82.

296. Bremer, H.a.D., P.P., *Modulation of chemical composition and other parameters of the cell by growth rate*, in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, R.C.I. Frederick C. Neidhardt, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, H.E. Umbarger, Editor 1996, ASM Press: Washington, D.C. p. 1553-1569.

297. Chivers, P.T. and R.T. Sauer, *NikR is a ribbon-helix-helix DNA-binding protein.* Protein Sci, 1999. **8**(11): p. 2494-500.

298. Brown, M.S., et al., *Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans.* Cell, 2000. **100**(4): p. 391-8.

299. Chen, G. and X. Zhang, *New insights into S2P signaling cascades: regulation, variation, and conservation.* Protein Sci, 2010. **19**(11): p. 2015-30.

300. Ehrmann, M. and T. Clausen, *Proteolysis as a regulatory mechanism.* Annu Rev Genet, 2004. **38**: p. 709-24.

301. Brown, M.S. and J.L. Goldstein, *The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor.* Cell, 1997. **89**(3): p. 331-40.

302. Brown, M.S. and J.L. Goldstein, *A proteolytic pathway that controls the cholesterol content of membranes, cells, and blood.* Proc Natl Acad Sci U S A, 1999. **96**(20): p. 11041-8.

303. Rudner, D.Z., P. Fawcett, and R. Losick, *A family of membrane-embedded metalloproteases involved in regulated proteolysis of membrane-associated transcription factors.* Proc Natl Acad Sci U S A, 1999. **96**(26): p. 14765-70.

304. Cutting, S., S. Roels, and R. Losick, *Sporulation operon spoIVF and the characterization of mutations that uncouple mother-cell from forespore gene expression in Bacillus subtilis.* J Mol Biol, 1991. **221**(4): p. 1237-56.

305. Ricca, E., S. Cutting, and R. Losick, *Characterization of bofA, a gene involved in intercompartmental regulation of pro-sigma K processing during sporulation in Bacillus subtilis.* J Bacteriol, 1992. **174**(10): p. 3177-84.

306. Campo, N. and D.Z. Rudner, *SpoIVB and CtpB are both forespore signals in the activation of the sporulation transcription factor sigmaK in Bacillus subtilis.* J Bacteriol, 2007. **189**(16): p. 6021-7.

307. Zhou, R. and L. Kroos, *Serine proteases from two cell types target different components of a complex that governs regulated intramembrane proteolysis of pro-sigmaK during Bacillus subtilis development.* Mol Microbiol, 2005. **58**(3): p. 835-46.

308. Ades, S.E., *AAA+ Molecular Machines: Firing on All Cylinders.* Current Biology, 2006. **16**(2): p. R46-R48.

309. Alba, B., et al., *DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigma'E-dependent extracytoplasmic stress response.* Genes & Development, 2002. **16**: p. 2156-2168.

310. Kanehara, K., K. Ito, and Y. Akiyama, *YaeL (EcfE) activates the sigmaE pathway of stress response through a site-2 cleavage of anti-sigma'E, RseA.* Genes & Development, 2002. **16**: p. 2147-2155.

311. Flynn, J.M., et al., *Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation.* Genes Dev, 2004. **18**(18): p. 2292-301.

312. Missiakas, D., J.M. Betton, and S. Raina, *New components of protein folding in extracytoplasmic compartments of Escherichia coli SurA, FkpA and Skp/OmpH.* Mol Microbiol, 1996. **21**(4): p. 871-84.

313. Rouviere, P.E. and C.A. Gross, *SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins.* Genes Dev, 1996. **10**(24): p. 3170-82.

314. Raivio, T.L. and T.J. Silhavy, *The sigmaE and Cpx regulatory pathways: overlapping but distinct envelope stress responses.* Curr Opin Microbiol, 1999. **2**(2): p. 159-65.

315. Connolly, L., et al., *The response to extracytoplasmic stress in Escherichia coli is controlled by partially overlapping pathways.* Genes Dev, 1997. **11**(15): p. 2012-21.

316. Rigel, N.W. and T.J. Silhavy, *Making a beta-barrel: assembly of outer membrane proteins in Gram-negative bacteria.* Curr Opin Microbiol, 2012.

317. Chaba, R., et al., *Design principles of the proteolytic cascade governing the sigmaE-mediated envelope stress response in Escherichia coli: keys to graded, buffered, and rapid signal transduction.* Genes Dev, 2007. **21**(1): p. 124-36.

318. Maeda, H., N. Fujita, and A. Ishihama, *Competition among seven Escherichia coli sigma subunits: relative binding affinities to the core RNA polymerase.* Nucleic Acids Res, 2000. **28**(18): p. 3497-503.

319. Grigorova, I.L., et al., *Fine-tuning of the Escherichia coli sigmaE envelope stress response relies on multiple mechanisms to inhibit signal-independent proteolysis of the transmembrane anti-sigma factor, RseA.* Genes Dev, 2004. **18**(21): p. 2686-97.

320. Chaba, R., et al., *Signal integration by DegS and RseB governs the σ E-mediated envelope stress response in Escherichia coli.* Proc Natl Acad Sci U S A, 2011. **108**(5): p. 2106-11.

321. Li, X., et al., *Cleavage of RseA by RseP requires a carboxyl-terminal hydrophobic amino acid following DegS cleavage.* Proc Natl Acad Sci U S A, 2009. **106**(35): p. 14837-42.

322. Kanehara, K., K. Ito, and Y. Akiyama, *YaeL (EcfE) activates the sigmaE pathway of stress response through a site-2 cleavage of anti-sigma’ E, RseA.* Genes & Development, 2002. **16**: p. 2147-2155.

323. Alba, B., et al., *DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigma’ E-dependent extracytoplasmic stress response.* Genes & Development, 2002. **16**: p. 2156-2168.

324. Levchenko, I., et al., *Versatile modes of peptide recognition by the AAA+ adaptor protein SspB.* Nat Struct Mol Biol, 2005. **12**(6): p. 520-5.

325. Bohn, C., J. Collier, and P. Bouloc, *Dispensable PDZ domain of Escherichia coli YaeL essential protease.* Mol Microbiol, 2004. **52**(2): p. 427-35.

326. Kanehara, K., K. Ito, and Y. Akiyama, *YaeL proteolysis of RseA is controlled by the PDZ domain of YaeL and a Gln-rich region of RseA.* EMBO J, 2003. **22**(23): p. 6389-98.

327. Walsh, N., et al., *OMP Peptide Signals Initiate the Envelope-Stress Response by Activating DegS Protease via Relief of Inhibition Mediated by Its PDZ Domain.* Cell, 2003. **113**: p. 61-71.

328. Clausen, T., et al., *HTRA proteases: regulated proteolysis in protein quality control.* Nat Rev Mol Cell Biol, 2011. **12**(3): p. 152-62.

329. Wilken, C., et al., *Crystal structure of the DegS stress sensor: How a PDZ domain recognizes misfolded protein and activates a protease.* Cell, 2004. **117**(4): p. 483-94.

330. Zeth, K., *Structural analysis of DegS, a stress sensor of the bacterial periplasm.* FEBS Lett, 2004. **569**(1-3): p. 351-8.

331. Cowan, S.W., et al., *The structure of OmpF porin in a tetragonal crystal form.* Structure, 1995. **3**(10): p. 1041-50.

332. Hagan, C.L., T.J. Silhavy, and D. Kahne, *beta-Barrel membrane protein assembly by the Bam complex.* Annu Rev Biochem, 2011. **80**: p. 189-210.

333. Rezuchova, B., et al., *New members of the Escherichia coli sigmaE regulon identified by a two-plasmid system.* FEMS Microbiol Lett, 2003. **225**(1): p. 1-7.

334. Wollmann, P. and K. Zeth, *The structure of RseB: a sensor in periplasmic stress response of E. coli.* J Mol Biol, 2007. **372**(4): p. 927-41.

335. Kim, D.Y., et al., *Structural basis for the negative regulation of bacterial stress response by RseB.* Protein Science, 2010. **19**(6): p. 1258-1263.

336. Kim, D.Y., et al., *Crystal structure of RseB and a model of its binding mode to RseA.* Proc Natl Acad Sci USA, 2007. **104**(21): p. 8779-84.

337. Hasselblatt, H., et al., *Regulation of the E stress response by DegS: how the PDZ domain keeps the protease inactive in the resting state and allows integration of different OMP-derived stress signals upon folding stress.* Genes & Development, 2007. **21**(20): p. 2659-2670.

338. Sohn, J., R.A. Grant, and R.T. Sauer, *Allostery is an intrinsic property of the protease domain of DegS: implications for enzyme function and evolution.* J Biol Chem, 2010. **285**(44): p. 34039-47.

339. Sohn, J., R.A. Grant, and R.T. Sauer, *Allosteric Activation of DegS, a Stress Sensor PDZ Protease.* Cell, 2007. **131**(3): p. 572-583.

340. Sohn, J., R.A. Grant, and R.T. Sauer, *OMP peptides activate the DegS stress-sensor protease by a relief of inhibition mechanism.* Structure, 2009. **17**(10): p. 1411-21.

341. Monod, J., J. Wyman, and J.P. Changeux, *On the Nature of Allosteric Transitions: A Plausible Model.* J Mol Biol, 1965. **12**: p. 88-118.

342. Cui, Q. and M. Karplus, *Allostery and cooperativity revisited.* Protein Sci, 2008. **17**(8): p. 1295-307.

343. Kinch, L.N., K. Ginalski, and N.V. Grishin, *Site-2 protease regulated intramembrane proteolysis: sequence homologs suggest an ancient signaling cascade.* Protein Sci, 2006. **15**(1): p. 84-93.

344. Koide, K., K. Ito, and Y. Akiyama, *Substrate recognition and binding by RseP, an Escherichia coli intramembrane protease.* J Biol Chem, 2008. **283**(15): p. 9562-70.

345. Inaba, K., et al., *A pair of circularly permutated PDZ domains control RseP, the S2P family intramembrane protease of Escherichia coli.* J Biol Chem, 2008. **283**(50): p. 35042-52.

346. Feng, L., et al., *Structure of a site-2 protease family intramembrane metalloprotease.* Science, 2007. **318**(5856): p. 1608-12.

347. Koide, K., et al., *Environment of the active site region of RseP, an Escherichia coli regulated intramembrane proteolysis protease, assessed by site-directed cysteine alkylation.* J Biol Chem, 2007. **282**(7): p. 4553-60.

348. Akiyama, Y., K. Kanehara, and K. Ito, *RseP (YaeL), an Escherichia coli RIP protease, cleaves transmembrane sequences.* EMBO J, 2004. **23**(22): p. 4434-42.

349. Saito, A., et al., *Post-liberation cleavage of signal peptides is catalyzed by the site-2 protease (S2P) in bacteria.* Proc Natl Acad Sci U S A, 2011. **108**(33): p. 13740-5.

350. Shah, S., et al., *Nicastrin functions as a gamma-secretase-substrate receptor.* Cell, 2005. **122**(3): p. 435-47.

351. Qiu, D., et al., *ClpXP proteases positively regulate alginate overexpression and mucoid conversion in Pseudomonas aeruginosa.* Microbiology, 2008. **154**(Pt 7): p. 2119-30.

352. Kim, D.Y., et al., *Crystal structure of the protease domain of a heat-shock protein HtrA from Thermotoga maritima.* J Biol Chem, 2003. **278**(8): p. 6543-51.

353. Cao, M., et al., *Antibiotics that inhibit cell wall biosynthesis induce expression of the Bacillus subtilissigmaW and sigmaM regulons.* Mol Microbiol, 2002. **45**(5): p. 1267-76.

354. Wiegert, T., et al., *Alkaline shock induces the Bacillus subtilis sigma(W) regulon.* Mol Microbiol, 2001. **41**(1): p. 59-71.

355. Ellermeier, C.D. and R. Losick, *Evidence for a novel protease governing regulated intramembrane proteolysis and resistance to antimicrobial peptides in Bacillus subtilis.* Genes Dev, 2006. **20**(14): p. 1911-22.

356. Heinrich, J. and T. Wiegert, *YpdC determines site-1 degradation in regulated intramembrane proteolysis of the RsiW anti-sigma factor of Bacillus subtilis.* Mol Microbiol, 2006. **62**(2): p. 566-79.

357. Zellmeier, S., W. Schumann, and T. Wiegert, *Involvement of Clp protease activity in modulating the Bacillus subtilissigmaw stress response.* Mol Microbiol, 2006. **61**(6): p. 1569-82.

358. Heinrich, J., K. Hein, and T. Wiegert, *Two proteolytic modules are involved in regulated intramembrane proteolysis of Bacillus subtilis RsiW*. Mol Microbiol, 2009. **74**(6): p. 1412-26.

359. Makinoshima, H. and M.S. Glickman, *Regulation of Mycobacterium tuberculosis cell envelope composition and virulence by intramembrane proteolysis.* Nature, 2005. **436**(7049): p. 406-9.

360. Sklar, J.G., et al., *M. tuberculosis intramembrane protease Rip1 controls transcription through three anti-sigma factor substrates.* Mol Microbiol, 2010. **77**(3): p. 605-17.

Vita
Sarah E. Barchinger

Education

The Pennsylvania State University May 2012
Ph.D., Biochemistry, Microbiology, and Molecular Biology, GPA: 4.0
University of Nebraska-Lincoln May 2005
B.S., High Distinction and Honors, Biochemistry, minors in Chemistry and English, GPA: 3.93

Honors and Awards

National Science Foundation Graduate Research Fellow, 2007-2010
University Graduate Fellow, 2005-2006
Homer F. Braddock Fellow, 2005-2007; Nellie H. and Oscar L. Roberts Fellow, 2005-2007
NSF GRFP Honorable Mention and Eberly College of Science NSF GRFP Incentive Award, 2006
Althouse Outstanding Teaching Assistant Award, Honorable Mention, 2006
Travel Awards: ABASM 2009, Molecular Genetics of Bacteria and Phages Meeting 2009

Publications

Barchinger SE†, Zhang X†, Hester SE, Rodriguez ME, Harvill ET, Ades SE (2012) *sigE* facilitates the adaptation of *Bordetella bronchiseptica* to stress conditions and lethal infection in mice. (Submitted)

Barchinger SE and Ades SE (2012) Regulated Proteolysis: Control of the *Escherichia coli* σ^E-dependent Cell Envelope Stress (Submitted)

Barchinger SE†, Zhang X†, Hester SE, Harvill ET, Ades SE. RseA and RseB negatively regulate *B. bronchiseptica* SigE and contribute to colonization and virulence in mice. (In prep.)

Costanzo A, Nicoloff H, Barchinger SE, Banta AB, Gourse RL, and Ades SE (2008) ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor σ^E in *Escherichia coli* by both direct and indirect mechanisms. *Molecular Microbiology.* 67(3):619-632

Selected Presentations

SE Barchinger, X Zhang, TL Nicholson, SE Hester, ET Harvill and SE Ades. "The SigE system is important for cell envelope stress response and virulence in *Bordetella bronchiseptica,*" Ninth International Bordetella Symposium 2010 (poster, invited oral presentation)

SE Barchinger, X Zhang, SE Hester, ET Harvill and SE. Ades. "Role and Regulation of the *Bordetella bronchiseptica* ECF sigma factor, SigE," *Bordetella* Group Reception at the ASM General Meeting 2009 (poster)

SE Barchinger, X Zhang, SE Hester, ET Harvill and SE Ades. "The SigE system is important for cell envelope stress response and virulence in *Bordetella bronchiseptica,*" Allegheny Branch of the American Society for Microbiology Meeting 2009 (oral presentation)

SE Barchinger, X Zhang, SE Hester, ET Harvill and SE Ades. "The SigE system is important for cell envelope stress response and virulence in *Bordetella bronchiseptica,*" Molecular Genetics of Bacteria and Phages Meeting 2009 (oral presentation)

SE Barchinger, X Zhang, SE Hester, ET Harvill and SE Ades. "Role and Regulation of the *Bordetella bronchiseptica* ECF sigma factor, SigE," Molecular Genetics of Bacteria and Phages Meeting 2007 and 2008 (poster)

Service/Professional Memberships

BMMB Graduate Rep., 2007-2008; Recruitment Tour Guide, 2006-2011; Mentor, 2007-2010
Coordinator, BMMB Graduate Student Seminars (Grad Talks), 2007-2009
BMB Climate and Diversity Committee Member, 2006-2007
PSU BioDays, Volunteer/Coordinator, 2006-2009; Higher Achievement, 2008
PSU Sinfonietta, 2007-2011
Graduate Women in Science Grad School 101 Panel, 2010
Member: Allegheny Branch of ASM, Phi Beta Kappa, National Society of Collegiate Scholars