Characterisation of a *phoP* P1*vir* transduction defect and the implication for the regulation of the extracytoplasmic stress response in *Escherichia coli*

PhD Thesis for the degree of PhD in

Microbiology

Faculty of Sciences School of Biosciences University of Kent

Amias Alstrom-Moore

2017

Declaration,

No part of this thesis has been submitted in support of an application for any degree or qualification

Of the University of Kent, or any other University or Institution of Learning.

Amias Alstrom-Moore

01/01/2018

Acknowledgements

I would like to take this opportunity to thank everyone who has supported me throughout the course of this PhD. I would like to express my special appreciation and thanks to Dr Ian Blomfield, you have been a tremendous mentor for me. Your guidance, support, encouragement has been invaluable to me.

Secondly I would like to thank Alex Moores for her invaluable help and supervision along the course of my research. Next I would like to thank the Blomfield and Shepherd labs. Louise, Sara, Charlie and Claudia, you made my PhD a glittering experience! To all my friends that helped me along the way in a myriad of ways, read my work, gave me somewhere to stay, made me laugh after a particularly hard day in the lab; Chris, Sarah, Igor, James, Chris, Richard and Jack, thank you.

Finally, a special thanks to my family the time and support you have given me has been without measure. Words cannot express how grateful I am for all of the sacrifices that you've made on my behalf. Tara, Meme, Zubie and Hari this one's for you.

Abstract

The sensor kinase PhoQ and its cognate response regulator PhoP constitute a two-component system, which is primarily responsible for sensing and responding to Mg^{2+} starvation in *Escherichia coli*. Additionally, there is growing evidence of regulatory links between PhoPQ and constituents of the outer membrane. Furthermore, it has been shown that PhoPQ is regulated negatively by MicA, an sRNA controlled by sigmaE. Encoded by *rpoE*, sigmaE is an alternative sigma factor that is activated in response to extracytoplasmic stress, specifically misfolded outer membrane proteins. Surprisingly, it was not possible to generate *ΔphoP* mutants, using P1*vir* transduction under standard conditions and kanamycin as the selective agent. Furthermore, a statistical analysis of these results indicates they cannot be explained by chance alone. The results show that PhoP is required for sigmaE activity in an RseA-independent manner, thereby suggesting that PhoP is a chief regulator of sigmaE activity. It is likely that diminished sigmaE activity in a *phoP* mutant, extracytoplasmic stress and OM deformation, caused by the reagents used in P1*vir* transduction, are responsible for the inability to transduce the *phoP* allele. Finally, evidence has been found relating to a second mechanism through which PhoP directly represses *rpoE* expression, thereby introducing further complexity into the regulator relationship that exists between sigmaE and PhoP.

Abbreviations

Aa	Amino acid
Ala	Alanine
Amp	Ampicillin
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
Asx	Aspartate or asparagine
bp	Base pair
Cam	Chloramphenicol
Cys	Cysteine
ddH_2O	Double deionised water
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
E. coli	*Escherichia coli*
EAEC	Enteroaggregative *Escherichia coli*
ECF	Extra-cytoplasmic function
EHEC	Enterohaemorrhagic *Escherichia coli*
EIEC	Enteroinvasive *Escherichia coli*
EPEC	Enteropathogenic *Escherichia coli*
ETEC	Enterotoxigenic *Escherichia coli*
GlcN-6-P	Glucosamine-6-phosphate
GlcNA	N-acetylglucosamine
Gln	Glutamine
Glu	Glutamate
Gly	Glycine
Hep	L-glycero-d-manno-heptose
His	Histidine
H-NS	Histone-like nucleoid-structuring protein
Ile	Isoleucine
IPTG	Isopropyl β-d-1-thiogalactopyranoside
Kan	Kanamycin
KNTase	Kanamycin nucleotidyltransferase
lac	Lactose
LB	Luria-Bertani
Leu	Leucine
LPS	Lipopolysaccharide
Lrp	Leucine-responsive regulatory protein
Lys	Lysine
MCL	Multiple cloning site
Met	Methionine
mRNA	(Messenger) ribonucleic acid

NTD	N-terminal domain
NTP	Nucleoside triphosphate
nts	Nucleotides
OD	Optical density
OM	Outer membrane
OMP	Outer membrane protein
ONPG	Ortho-nitrophenyl-β-galactosidase
ORF	Open reading frame
P3	Third promoter
PCR	Polymerase chain reaction
Phe	Phenylalanine
Pro	Proline
RBS	Ribosome binding site
RNAP	RNA polymerase
rpm	Revolutions per minute
SD	Shine-Dalgarno
SDS	Sodium dodecyl sulphate
Ser	Serine
sRNA	(Small) ribonucleic acid
TAE	Tris/acetate/edta
TCS	Two-component system
TE	Tris/edta (buffer)
Tet	Tetracycline
Thr	Threonine
tl	Translational
Tris	Tris(hydroxymethyl) aminomethane
tRNA	Transfer RNA
Trp	Tryptophan
Tyr	Tyrosine
UPEC	Uropathogenic *Escherichia coli*
UTI	Urinary tract infection
UTR	Untranslated region
Val	Valine

Contents

Chapter 1: Introduction 1

1.1 *Escherichia coli* 2

1.2 The *E. coli* inner and outer membrane 4

1.2.1 Peptidoglycan 11

1.3 Gene regulation 14

1.3.1 RNA polymerase and sigma factors 15

1.3.2 Regulatory RNAs 20

1.3.3 Transcription factors 21

1.3.4 Two-component systems 25

1.4. PhoPQ 29

1.4.1 The PhoPQ regulon 32

1.4.2 PhoPQ regulation 36

1.5. Specific stress responses 37

1.5.1 Heat and cold shock 38

1.5.2 Envelope stress 40

1.5.3 General stress response 42

1.6 Horizontal gene transfer and bacteriophages 44

1.6.1 The lytic and lysogenic cycles 45

1.6.2 P1*vir* transduction 49

1.7 Project Aims 53

Chapter 2: Materials and Methods 54

2.1 Bacterial strains, plasmids and oligonucleotides 55

2.2 QIAquick kits 63

2.2.1 Media, storage and growth conditions 63

2.2.2 Luria-Bertani media 63

2.2.3 Sucrose media 64

2.2.4 Agar Supplements 64

2.3 Centrifugation 65

2.4 P1*vir* transduction 65

2.4.1 P1*vir* transduction-derived procedures 67

2.5 Isolation of genomic DNA 68

2.5.1 Isolation of plasmid DNA 69

2.5.2 Restriction endonuclease digestion of DNA 69
2.5.4 Production of chemically competent cells 71
2.5.5 Lambda red re-combineering 72
2.5.6 The polymerase chain reaction 74
2.6 DNA purification 76
2.6.1 Curing Keio collection mutants 76
2.7. Allelic exchange 77
2.8 β-galactosidase assay 80
2.9 DNA analysis by agarose gel electrophoresis 81
2.10 Whole-cell lysate preparation 82
2.10.1 Markwell Assay 83
2.10.2 Western Blot Procedure 84
2.11 Automated growth curve analysis by plate reader 86
Chapter 3: The *phoP* transduction defect 88
3.1 Introduction 89
3.2 The *phoP* transduction defect 90
3.2.1 Investigating the *phoP* transduction defect 91
3.3 Complementation of the *phoP* transduction defect 94
3.3.1 Construction of the *lac*UV5-*phoP* strain 95
3.3.2 Testing the complementation mutant *lac*UV5-*phoP* 104
3.3.3 Construction of a *phoP* mutant via lambda red recombination 106
3.4 Discussion 107
Chapter 4: Exploring the mechanism of the *phoP* P1*vir* transduction defect 111
4.1 Introduction 112
4.2 Transduction of *mgtA* 113
4.2.1 Metal ion starvation in *phoP* mutants 113
4.2.2 Testing the effect of citrate on the viability of *phoP* mutants 117
4.2.3 Testing the effect of transduction reagents on the viability of *phoP* mutants 120
4.3 *phoP* mutants can adapt to the conditions of P1*vir* transduction 123
4.3.1 *phoP* mutants can be transduced if the standard incubation times are changed 125
4.3.2 The inability to create suppressor-free *phoP* transductants supports the notion of genetic compensation for loss of the gene 127
4.4 *phoP* mutants can be transduced, if the standard temperature is changed 129

4.4.1 *phoP* transductants can be isolated under standard conditions in suppressor mutant backgrounds 131
4.4.2 The effect of Procaine on *phoP* transduction 136
4.5 Discussion 137
Chapter 5: The regulation of sigmaE by PhoPQ 141
5.1 Introduction 142
5.2 PhoPQ positively regulates sigmaE activity 143
5.2.1 The effect of *mgrB* mutation on sigmaE activity 145
5.2.2 The effect of *phoP* loss on sigmaE, demonstrated via complementation 146
5.3 Directly observing the effect of loss of *phoP* on cellular sigmaE concentration 148
5.4 PhoP influences sigmaE activity independent of RseA 152
5.5. Searching for a PhoP-Box in the *rpoE* promoter region, using *in silico* analysis 154
5.5.1 Optimising the PFM threshold for consensus sequence prediction 158
5.5.2 Searching for PhoP binding sites in the *rpoE* promoter region 163
5.6 SigmaS positively regulates sigmaE activity, independent of RseA 166
5.6.1 *rpoS* mutation acts as an anti-suppressor for the *phoP* Transduction defect 168
5.7 Discussion 169
Chapter 6: Discussion 174
6.1 Background Summary 175
6.2 *phoP* is necessary for successful P1*vir* transduction under standard conditions
6.3 The mechanism of the *phoP* transduction defect 181
6.4 The Absence of a *phoQ* transduction defect 184
6.5 Postulations on the mechanism by which PhoP regulates sigmaE 185
6.6 Final model 189
6.7 Future work 191
6.7.1 Characterising *phoP* mutants from the published literature 191
6.7.2 Exploring the Absence of a *phoQ* transduction defect 192
6.7.3 Further tests required to characterise the *phoP* transduction defect 193
6.7.4 Further characterising the relationship between PhoP and sigmaE 194
Appendix 197
Appendix A 198
Appendix B 202
Appendix C 205
Appendix D 207

Appendix E 210
Appendix F 217
References 224

Chapter 1: Introduction

1.1 *Escherichia coli*

Escherichia coli (*E. coli*) is a gram-negative rod-shaped facultative anaerobe associated with the lower gut of animals, though it can persist in the environment[1]. Furthermore, it is a commensal organism in humans and colonises the mucosal layer of the lower intestines hours after birth[1]. This colonisation benefits both organisms, as *E. coli* is a successful competitor of pathogens attempting to establish themselves within its niche. These commensal *E. coli* can only cause disease in cases where the host has become immunocompromised, or where gastrointestinal tract barriers are breached[1]. However, more virulent strains of *E. coli* are responsible for many common human afflictions. When a specific set of virulence factors is accumulated by a bacterium, that organism is able to adapt to a new function and persist as a pathogen. Numerous strains of *E. coli*, known as pathotypes[3], have adapted to a variety of niches via the use of an assortment of virulence factors[1].

Pathotypes arise when beneficial virulence factors are incorporated into the bacterial genome. Genetic analysis of various pathotypes has shown the importance of horizontal gene transfer in the successful evolution of virulence and adaptation to new environments[3]. Newly acquired DNA can be obtained from a variety of sources such as plasmids, bacteriophage or transposons[4], and the new genetic material acquired by a potential pathogen often becomes integrated into the chromosome at insertion hotspots, which in turn can form pathogenicity islands[3].

Pathotypes are able to cause a wide array of diseases in and outside of the gut and according to their virulence mechanism, they are categorised into two groups: diarrhoeagenic *E. coli* and extraintestinal *E. coli*. Diarrhoeagenic *E. coli*, which is one of the most common etiologic agents of diarrhea[1], is further sub-classified into six groups, namely enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC) and diffusely adherent *E. coli* (DAEC, diarrhoeal disease and enteric

disease)[1], [3]. Pathotypes associated with extraintestinal infections are uropathogenic *E. coli* (UPEC) and meningitis-associated *E. coli (*NMEC)[3]. Pathogenic *E. coli* can also be categorised by serogroup, a type of grouping based on the lipopolysaccharide O antigen and the flagella H antigen[3]. However, pathotypes can contain many different serotypes and serotypes can be part of several different pathotypes. However, serotyping remains an important method for establishing the virulence of an *E. coli* isolate[3], [4].

Diarrhoeagenic *E. coli* pathotypes can be characterised by the cytoskeleton rearrangements it causes the host cell, such is the case with EHEC, EPEC, and the production of enterotoxins and endotoxins[3]. Enterotoxins are diverse in conformation and mechanism of action. These toxins can cause both diarrhoeal and non-diarrhoeal diseases[3]. For example, the Shiga toxin, produced by Shiga-toxin producing EHEC (known as STEC and verotoxigenic *E. coli* (VTEC)), is responsible for causing haemorrhagic colitis, bloody diarrhoea, non-bloody diarrhoea and hemolytic uremic syndrome (HUS)[3]. However, many non-diarrhoeal symptoms are caused by endotoxins, also known as lipopolysaccharide. LPS can cause a number of symptoms but is most often associated with endotoxic shock[5].

Pathotypes utilise many of their surface structures as virulence factors[3] and can damage the host cell but benefit the pathogen[3]. Fimbriae are bacterial surface structures that are important virulence determinants[3]. They are large, rod-like appendages anchored to the bacterial outer membrane. Similar in structure to pili, fimbriae are distinct, as they play no role in conjugation or the transfer of genetic material between bacteria[3].

Some pathotypes utilise the type III secretion system, which is found in a variety of gram-negative bacteria[3]. The type III secretion system is made up of structural proteins and translocators, the latter of which function to translocate effector proteins into the host cell cytoplasm[3]. The type III secretion system gives pathogens the ability to inject toxins and effector proteins directly into the host cell. its basic structure is that of an inner and outer membrane ring[6], each of which helps to provide a continuous path from the cytoplasm to the exterior of the

cell. Finally, a needle-like structure, associated with the outer membrane ring, projects out from the bacterial cell surface[1]

Fast-growing, robust and oxygen-tolerant *E. coli* has become a favoured organism, widely used to research the many fundamental and basic principles upon which modern microbiology is based. Much of what is known about microbial DNA replication, gene expression and protein synthesis is derived from work on this organism[7], with *E. coli* K12 and its derivatives being some of the most widely used strains in this regard[8].

E. coli's relativity simple nutritional needs and rapid growth rate make it an invaluable asset for molecular biology and biochemistry and due to its long history of study, it is one of the most well researched and best understood organisms, though many aspects of its cellular physiology and genetic regulation still require illumination.

1.2 The *E. coli* inner and outer membrane

All cells have a cytoplasmic membrane, a fluid lipid bilayer with associated proteins. Gram-positive, monoderms, have a single cytoplasmic membrane[9] and utilise a (comparatively thicker) peptidoglycan barrier and the cytoplasmic membrane only as their primary barrier between the environment and the cytoplasm[10]. While the cytoplasmic membrane is conserved in all domains of life, many bacterial cells have developed complementary structures and architecture through the cell wall[11], which contributes greatly to protecting (from detrimental environmental conditions) and structuring the cell[12]. Gram-negative, diderm, bacteria have a secondary lipid bilayer, known as the outer membrane (OM). In between these two membranes is the periplasmic space, which contains a thin layer of peptidoglycan. The OM and inner membrane (IM) are structurally distinct, in that although both are composed of phospholipids and proteins, the IM contains an symmetric phospholipids bilayer, while the OM has an asymmetric bilayer[13]. Another important distinguishing feature of the OM is the presence of lipopolysaccharide (LPS)[13].

Lipids play a significant role in bacterial membranes. Most bacterial lipids are derivatives of long-chain fatty acids and can be split into two groups[14], namely neutral lipids, which are hydrophobic, and complex lipids[14], which are amphipathic and contain a hydrophobic moiety, usually in the form of a charged head group[14]. Phospholipids are the most abundant lipid in bacterial cell membranes[15] and consist of two fatty acid chains linked to a charged phosphate group[15]. They are important for membrane structure and flexibility, and their modification, such as variations in fatty acid residues and head group composition, can help the cell adapt to certain conditions[16]. The majority of lipids in the inner membrane are a class of phospholipids known as Phosphatidylethanolamine[17], while the outer membrane has a much greater variety of lipids, containing phosphatidylcholine, phosphatidylethanolamine, cholesterol and others in smaller amounts[14].

Phospholipids provide the basic structure of the bacterial cell membranes but proteins are employed to carry out most of the specific functions. These proteins can be broken down into two groups, i.e. integral membrane proteins and peripheral membrane proteins[18], the latter of which only associate with the membrane temporarily, are not embedded within the lipid bilayer and are often associated with integral membrane proteins[18], which are embedded into the lipid bilayer and permanently associated with the membrane[18]. Some integral membrane proteins are also known as transmembrane proteins, spanning the inner and outer leaflet of the membrane[19]. Portions of these proteins are exposed on both sides of the leaflet. Transmembrane proteins are amphipathic, with hydrophobic-sided chains interacting with the fatty acids in the phospholipids bilayer and the hydrophilic portion exposed to the aqueous environment of the cellular compartments[18].

The inner membrane, which encloses the cytoplasm, and its associated proteins are involved in a wide variety of different processes, such as the movement of molecules, macromolecules and proteins between cellular compartments, environmental sensing and the biosynthesis of a range of molecules such as lipids,

peptidoglycan, LPS, etc.[20]. The inner membrane contains lipoproteins and both integral and peripheral proteins[20]. Peripheral proteins may exist on either side of the membrane, and those residing on the cytoplasmic side interact with the proteome of the cytoplasm and play a role in cell metabolism[21]. Inner membrane proteins (IMPs) are complex and diverse in function.

The OM (Figure 1.1) acts as a barrier for diffusion and grants additional resistance to antibiotics and detergents[13]. This is of particular relevance for pathogens, where the OM bestows greater protection for surviving inside parts of the mammalian body, for example the gastrointestinal tract. Moreover, because of the nature of the OM, diffusion of many important metabolites, such as amino acids, ions, oligosaccharides and waste products[22], is facilitated by integral OM proteins (OMPs)[23]. They are important for transport and assembly of structures on the inner and outer faces of the OM[24]. The majority of OMPs are β-barrel proteins[24]. To date, only two OMPs have been categorised as essential in *E. coli*, namely the LptD-LptE complex, which is responsible for inserting LPS into the OM, and the BamA[24]. Although the structures of OMP, i.e. β-barrels, are similar, they are nevertheless quite diverse. Non-specific OMPs, for instance, function as porins and form channels for hydrophilic molecules up to 600Da in size[11]. Although denoted as non-specific, some of these porins show a preference for molecules based on their charge and size[11]. OmpF and OmpC, for example, allow for the diffusion of positively charged molecules[11], while, conversely, PhoE has a preference for negatively charged solutes[25].

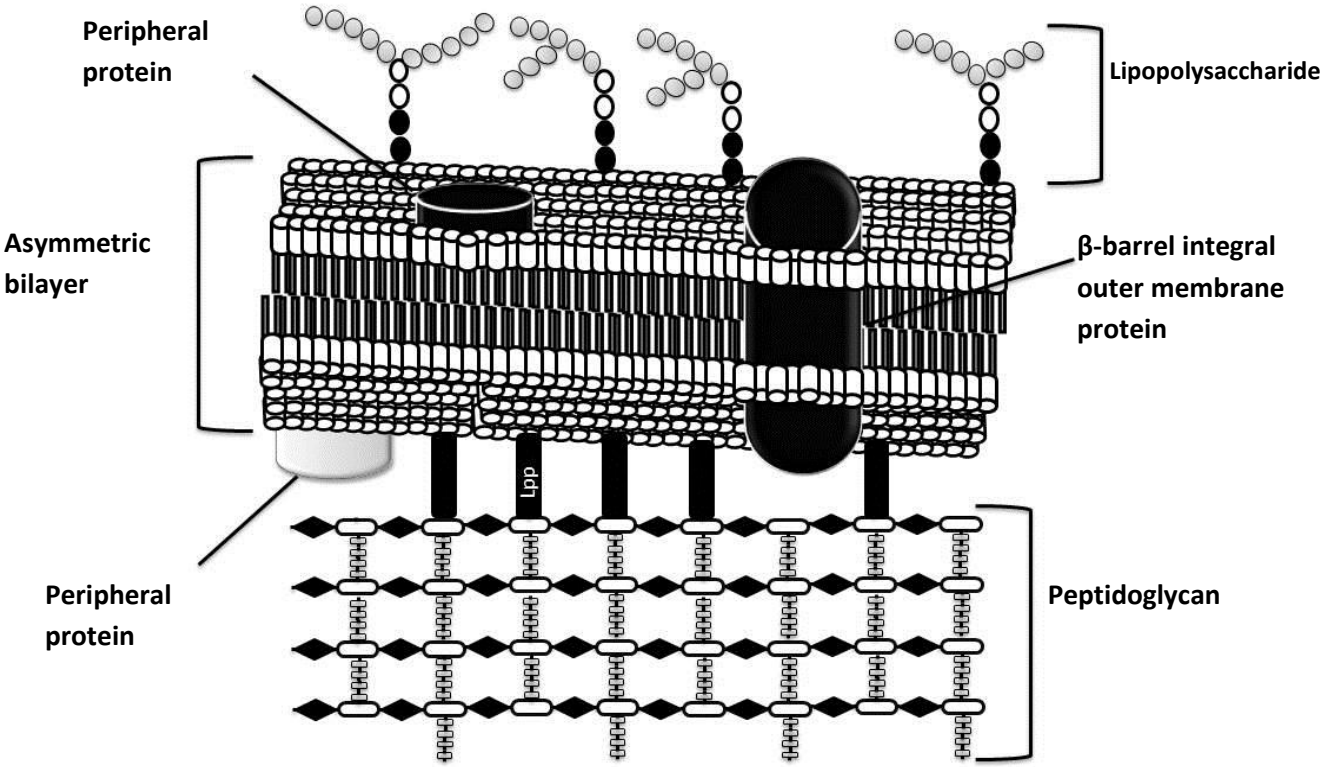


Figure 1.1. Showing a cross-section of the OM and peptidoglycan chains in the periplasmic space.

Other OMPs form channels for specific (usually large) molecules; LamB allows for the diffusion of maltose and other sugars and FhuA for Fe^{3+}[26]. Some of these specific channels allow passive diffusion, while others require facilitation from active transport systems. OMPs can also be used as channels for the exportation of solutes and proteins from the cell, and some have enzymatic capabilities, such as the PhoP-regulated OmpT[27]. Other OMPs play a role in maintaining the structure of the OM, an example of which is the BAM complex, while others help form the LSP, such as LptD[28].

OMPS are synthesised in the cytoplasm. Therefore, OMP (or OMP precursors) must be transported through the inner membrane, the periplasm, and inserted correctly into the OM[24]. This is achieved by periplasmic chaperones and complex, membrane-embedded translocation proteins[29]. Transportation of unfolded OMPS through the IM is facilitated by SecYEG complex[16].

The Sec-translocon is a heterotrimeric complex[29], The main channel of the translocon is formed from SecYEG and peripheral protein, SecA, which is an ATPase, and provides energy for the translocation of the proteins[29]. The cytoplasmic chaperone SecB delivers unfolded OMPS to secYEG for translocation across the membrane[29], though proteins may be delivered co-translationally and targeted to the SecYEG by an N-terminal signal peptide[29]. Once in the periplasm, OMPS are prone to aggregation and misfolding. To prevent this issue, the OMPS are escorted through the periplasm by chaperones that bind to unfolded OMPS. Additionally, the periplasm also contains protease, which degrades misfolded OMPs[30]. Survival protein A (SurA), Seventeen kilodalton protein (SKp) and DegP are three well studied examples of periplasmic proteins that perform these functions[30].

SurA is peptidyl-prolyl isomerase and required for folding OMPs, as well as interconverting the cis and trans isomers of peptide bonds[31]. Strains lacking *surA* see an increase in misfolded OMPs and a decrease in OMP density, compared to *surA* positive strains[31]. SKp binds to unfolded OMPs and acts as an anti-aggregator and chaperone[32], and it also targets OMPs with rich hydrophobic and aromatic residues, binding them and shielding the protein from the surrounding environment[32].DegP is a protease that seeks misfolded and aggregated[33]. Once chaperoned through the periplasm, OMPs are delivered to the BAM (β-barrel assembly machinery) complex.

BAM is responsible for the insertion and folding of OMPs and consists of BamA and four lipoproteins, referred to as “BaMBCDE.” While BamA is an essential and integral transmembrane protein, BaMBCDE are anchored to the inner leaflet of the OM[34]. The BAM complex recognises the C-terminal motif, which acts as a signal for insertion[34].

Other OM proteins that have important functions are lipoproteins much work has been done on Lpp – the most abundant protein in the cell and which anchors the peptidoglycan to the OM[35].

Lipopolysaccharide (LPS) is a glycolipid[36] found exclusively in gram-negative bacteria, embedded in the outer leaflet of the OM. The LPS is essential for almost

all gram-negative bacteria and is a large constituent of the OM[36]. It is made of three domains: lipid A (endotoxin), core oligosaccharide (subdivided into inner and outer variants) and a variable O-antigen (Figure 1.2). Lipid A consists of two glucosamines with attached fatty acids, making LPS a saccharolipid. The inner core has a conserved structure, while the outer core is more variable in this respect[37]. However, the LPS core structure is known to be variable among *E. coli* serotypes. There are five well defined core structures, namely R1, R2, R3, R4 and K-12. These core differ slightly in regard to the sugar residues that make up the inner and outer core[38]. While these core types are found among commensal isolates certain core structures are associated with pathogenic strains. For example the R1 core structure is frequently encountered in extraintestinal isolates, while R3 is predominate STEC and VTEC [38].

A Kdo (3-deoxy-d-manno-oct-2 ulosonic acid) group is the first residue that links the inner core to lipid A[36]. Modification of the Kdo group can affect the cell's resistance to antimicrobial peptides[39]. EptB modifies the Kdo group by adding a phosphoethanolamine (pEtN) moiety[39], which reduces the anion charge of the molecules, decreases electrostatic repulsion between the LPS molecules. EptB is positively regulated by sigmaE[39]. The Kdo groups are followed by heptose, which are phosphorylated, then finally, in *E. coli* K-12, three glucose residues.

The immunogenic O-antigen is at the terminal end of the LPS molecule and also acts as a virulence factor to help the cell evade and escape the host immune system[40]. The O-antigen extends, from the cell surface, into the environment and constitutes the hydrophilic portion of the LPS[37]. The O-antigen consists of repeating oligosaccharides, making units of 3-5 sugars. These units can repeat up to 40 times[37].

The interaction between LPS molecules, facilitated by divalent cations binding to negative charges on the Kdo carboxylate and glucosamine phosphates groups, can cause neighbouring molecules to bridge the stabilisation of the overall OM structure[37].

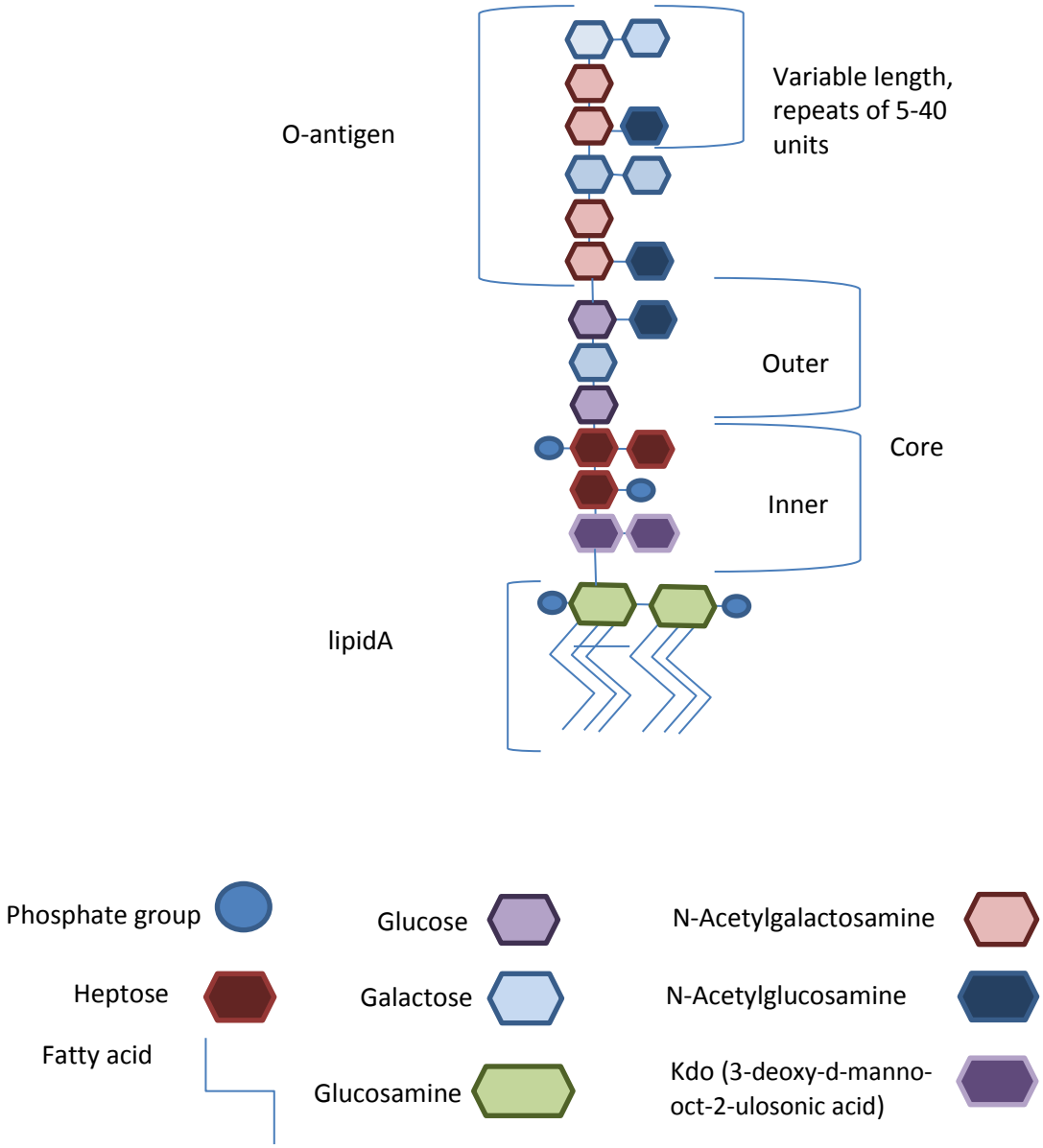


Figure 1.2. *E.* coli K-12 LPS structure

LPS is synthesised in cytoplasm and on the cytoplasmic face of the inner membrane. LPS synthesis starts with lipid A[37]. Lipid A is synthesised from the precursor UDP-N-acetylglucosamine. Subsequent reactions by LpxA, LpxC and LpxB convert UDP-N-acetylglucosamine into UDP-diacyl-GlcN[37]. The uridine-5'-monophosphate motiy is cleaved from UDP-diacyl-GlcN by LpxH and it is converted to lipid X (2,3-diacylglucosamine-1-phosphate). Next LpxB encodes an inverting glycosyl transferase[37]. LpxB enables the formation of 2',3'-diacylglucosamine-(β,1'-6)-2,3-

diacylglucosamine-1-phosphate (Disachride-1-P) from LipidX and UDP-diacyl-GlcN. The kinase, LpxK phosphorylates disaccharide-1-P, this forms lipid IVA[37]. The enzyme KdtA incorporates two 3-deoxy-D-manno-octulosonic acid (Kdo) residues into lipid IVA, subsequent reactions by LpxL and LpxM form Kdo2-lipid A[37].

The core oligosaccharide region of LPS varies between *E. coli* strains but a common feature is a core set of sugars, which variable regions attached. These are a linear set of six sugars and are most often e Kdo, Hep, D-glucose and D-galactose. The core is synthesised by several enzymes, WaaG, WaaB, WaaO, WaaR, WaaF, and WaaC[37] . The enzyme encoded by *waaL* is essential for attachment of core-lipid A complex to the O-antigen[37].

Sugar-nucleotide act as precursors for the synthesis of the O-antigen. The precursors are synthesised by glycosyltransferases and polymerases and finally attached to the core-lipid A complex[37]. O-antigen synthesis happens on the periplasmic face of the inner membrane[37]. The core-lipid A is transported to the periplasm by, integral IMP, MsbA[37]. The O-antigen is transported across the inner membrane by the Wzx protein[37]. Assembly is carried out in the periplasm by enzyme encoded by *waaL* (O-antigen ligase)[38]. Several proteins expressed from the *lpt* operon, such as chaperones, IMPs and OMPs function to transport the nascent LPS to the OM[37]. LptBFG form a ABC transporter that, with assistance from LptA and LptC, translocate the LPS molecule to the inner leaflet of the OM[37]. It is believed that LPS molecules are transported to outer leaflet of the OM by LptD and LptE[37].

1.2.1 Peptidoglycan

The periplasm is interposed between the inner and outer membranes and is an oxidising, viscous, gel-like compartment[12] containing a diverse soup of proteins and molecules necessary for vital functions[12].

Peptidoglycan, situated in the periplasm, forms a sacculus and gives the envelope its rigidity. Cells lacking peptidoglycan are prone to lysis[12]. The peptidoglycan

sacculus has a (hydrated) thickness of approximately 2.5-7nm and has high elasticity[12], which allows the surface area of the sacculus to stretch – necessary to withstand turgor osmotic pressure placed on sacculus by the cytoplasm[12]. Moreover, the sacculus contains pores (approximate radius of 2.1nm) which aid the diffusion of proteins up to 50KDa in size, thereby acting as a molecular sieve[12].

Peptidoglycan is formed from cross-linked chains of glycans, which in turn are connected via short peptides[21] and formed alternating units of β-1,4-linked *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) to create the bulk of the peptidoglycan structure[21].

Peptides link the chains GlucNac & MurNac together with neighbouring chains[21]. The peptides are amide-linked to the lactoyl group of MurNAc[21]. *E. coli* has a peptide structure L-alanine (L-Ala)-gamma-D-isoglutamic acid (D-iGlu)-meso-diaminopimelic acid (m-Dap)-D-alanine (D-Ala)(Figure 1.3)[21].

Synthesis and translocation of peptidoglycan is a complex, multifaceted, process. The first steps take place in the cytoplasm, while the subsequent steps take place on the inner and outer side of the inner membrane[21].

In the initial steps, precursors such as UDP-Nacetylglucosamine, UDP-N-acetylmuramic acid, UDP-N-acetylmuramyl-pentapeptide and D-glutamic acid are synthesised. Biosynthesis of UDP-GlcNAc from fructose-6-phosphate is facilitated by proteins transcribed from the *glm* operon (GlmS, GlmM and GlmU)[41]. UDP-GlcNAc can be catabolised into GlcNAc-6-phosphate and then into fructose-6-phosphate, which is carried out by NagA and NagB, respectively[41]. It has been found that the *nagA* promoter responds to the concentration of extracellular divalent cations via the regulator PhoP[42]. In this way, PhoPQ influences the availability of the peptidoglycan precursor UDP-GlcNAc.

In the next set of reactions, UDP-GlcNAc is used as a precursor for UDP-MurNAc[41]. This step is the first committed move towards peptidoglycan formation. These reactions are carried out by MurA and MurB in a two-step process that yields the product UDP-MurNAc[41]. Like UDP-GlcNac and UDP-MurNAc,

smaller peptides, used to link glycan chains of peptidoglycan, are also synthesised in the cytoplasm. These include D-Ala, D-Ala-D-Ala and D-Glu[41].

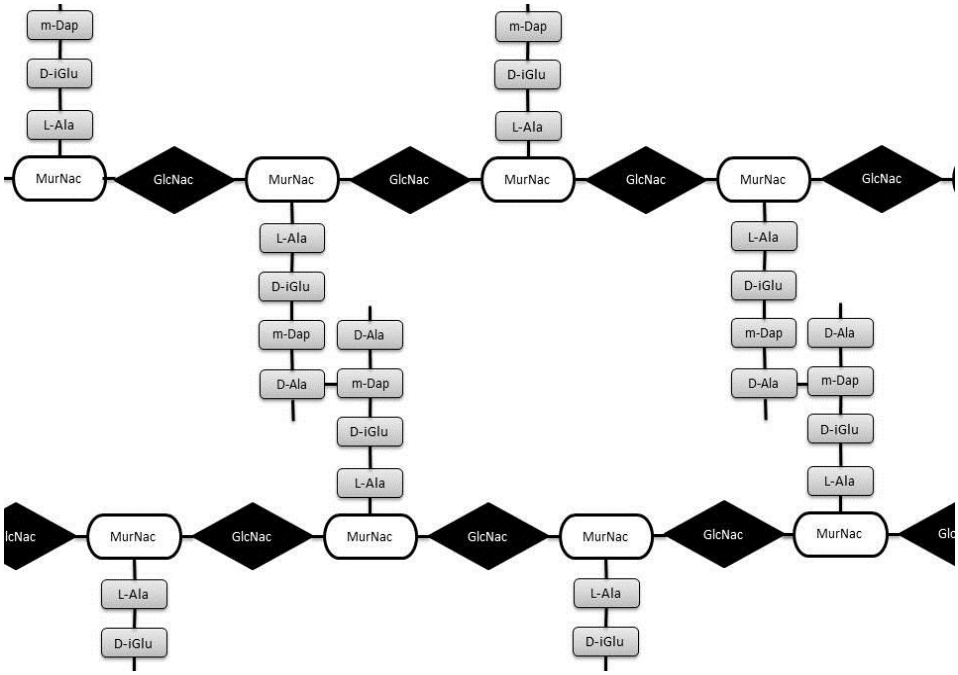


Figure 1.3. The peptidoglycan structure shown above is made of alternating chains of *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc). MurNAc has a stem of peptides, L-alanine (L-Ala)-gamma-D-isoglutamic acid (D-iGlu)-meso-diaminopimelic acid (m-Dap)-D-alanine (D-Ala). This peptide stem cross-links the glycan chains horizontally and vertically (not shown) to create the peptidoglycan's 3D structure.

The later steps in peptidoglycan synthesis take place within and at the interface of the cytoplasmic membrane[41], including the translocation and polymerisation of the peptidoglycan precursors. Lipid II is the final peptidoglycan precursor before the peptidoglycan polymer is formed and consists of one GlcNAc-MurNAc-pentapeptide subunit[21]. Initially anchored in the inner leaflet, lipid II is flipped to the outer leaflet of the cytoplasmic membrane by the proteins RodA and FtsW[33]. PG synthases then polymerise the glycan chains via glycosyltransferase (GTase) reactions and form the peptide cross-links that make up peptidoglycan[41].

1.3 Gene regulation

Gene expression in bacteria is influenced primarily by two extracellular factors, namely nutrient availability and environmental conditions. These two factors must be monitored closely by the cell, as they can change quickly and dramatically, and so overcoming these changes is often necessary for the transmission, survival and pathogenesis of an organism. For example, *E. coli* must move from the nutrient-rich environment of the gut to the harsh outside environment, in order to colonise new hosts successfully. To survive such changes, the bacteria must utilise a number of methods to effect genetic and phenotypic changes that help them adjust rapidly to their new conditions.

Gene expression can be controlled at both the transcriptional and post-transcriptional levels. Promoter specificity, transcription factors or sigma factors, can influence the expression of genes on transcriptional level. While post-transcriptional regulation is often carried out by small non-coding regulatory RNAs

Gene regulation in *E. coli* is arranged hierarchically. Genes, with a similar function, are grouped into operons. Individual operons are often regulated by cis- and trans-acting regulators, or by the abundance of a pathway-specific precursor or end product[43]. Multiple operons that contribute to a similar function in the cell, and share a common regulator, are grouped into regulons[44]. These shared regulators often recognise a DNA target sequence that is shared by all members of the regulon[45], multiple collections of which are grouped into stimulons. Stimulons, i.e. multiple overlapping regulons, allow the cell to react broadly to changes such as osmolarity or nutrient deprivation. Hierarchical regulation allows for broad and specific responses, fine-tuning gene expression and the integration of multiple environmental and nutritional signals into a global transcriptional network[46].

1.3.1 RNA polymerase and sigma factors

RNA polymerase has broad specificity, though not all promoters are bound with the same affinity[47]. The promoter sequence has an effect on the affinity of RNA polymerase binding, in that promoters that bind to RNA polymerase with a high affinity and are transcribed with a high frequency are strong, while those that are transcribed at a low frequency are weak. A highly conserved characteristic of promoter architecture are the –10 and –35 hexamers, located 10 and 35 base pairs upstream from the transcription start site. The consensus sequences of these hexamers are 5'-TATAAT-3' and 5'-TTGACA-3. Deviation from the consensus sequences or the optimum spacer between them, 17bp, is associated with weak promoters[48].

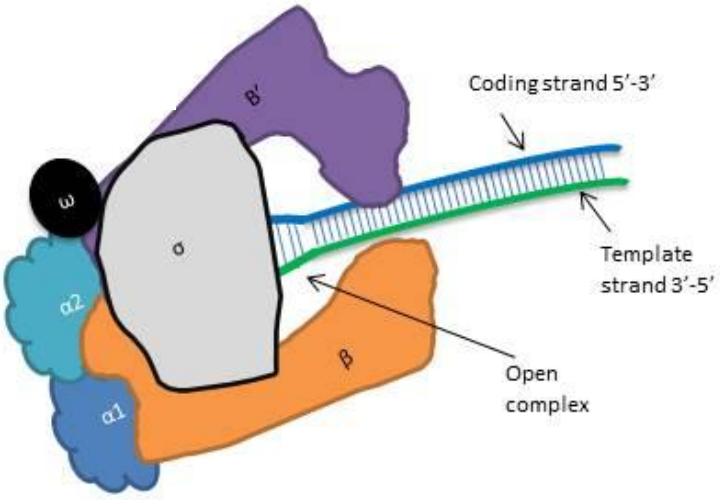
Other than the -10 and -35 sequences, some promoters have a -16, discriminator and UP element. The -16 motif (3-4 bp in length) is found adjacent to the -10 element[49]. The discriminator is the region between -10 region and transcription start site and effects isomerisation of the double stranded DNA molecule during transcription initiation[49]. The UP element is a region upstream of the -35 sequence and binds the α-C terminal domain of RNA polymerase[49].

Transcription requires the recruitment of RNA polymerase, but the available concentration of RNA polymerase is limited in the bacterial cell[50]. A proportion of the cell's RNA polymerase is bound inactively to the cell's DNA, while the majority of active RNA polymerase is occupied by creating transcripts of stable RNA needed for translation[50]. RNA polymerase holoenzyme is composed of a several subunits, two alpha, one beta and one beta'[50]. RNA also requires dissociable subunits called "sigma factors". RNA polymerase can bind to DNA without sigma factors; however, without these subunits, the enzyme will be unable to differentiate between sequences, as it relies on sigma factors for consensus sequence recognition[51]. The beta subunit is encoded by the *rpoB* gene and has two catalytic sites[52], the first of which is utilised to incorporate the "first nucleotide" in the RNA message, while the second site is for the addition of the "second nucleotide",

which is the elongation site[52]. The beta' subunit is a large protein encoded by the gene *rpoC*. This subunit has a DNA binding site, namely a 25 Å groove[47]. The beta' subunit positions the non-transcriptional strand away from the transcriptional strand in such a way that it does not inhibit the process[52]. Assembly requires an additional subunit, omega[47], which is required for assembly of the holoenzyme. The subunits are assembled sequentially in the following order: alpha, alpha, beta, beta' and finally the addition of the sigma factor -α2ββ'σ [53]. Sigma factors bind to the core RNA polymerase in such a way that they are in contact with the RNA polymerase subunits β and β'[54], which in turn forms a holoenzyme. The sigma factor, bound to RNA polymerase, directly interacts with DNA in the -35 and -10 regions[51], following which the complex then binds the promoter region, and RNA synthesis ensues. The association between the sigma factor and RNA polymerase is reduced during elongation and termination, as the sigma factor is unnecessary for these steps[52].

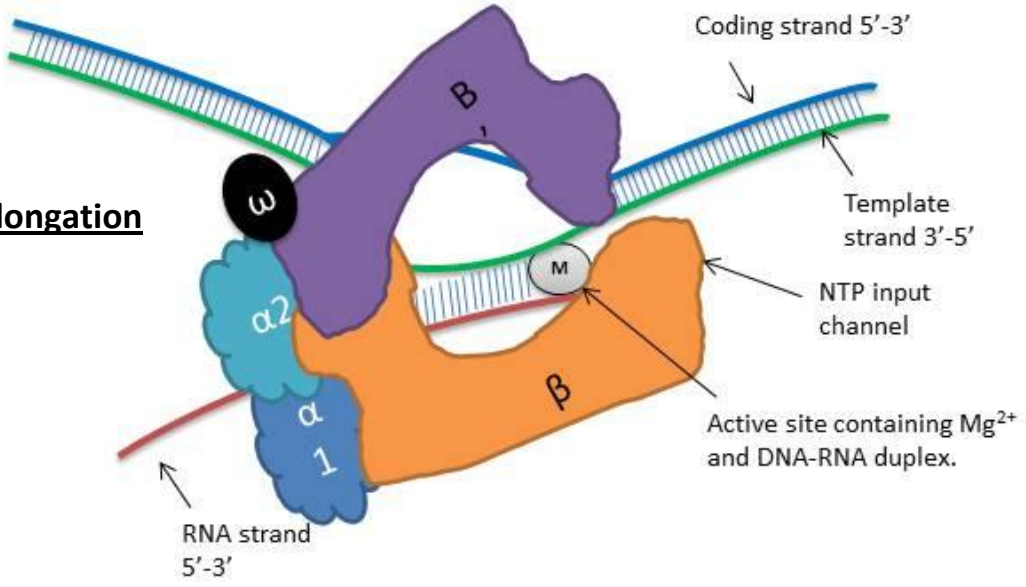
Transcription starts at the promoter of the DNA template. The holoenzyme binds to duplex DNA and rapidly slides along the double helix in search of a promoter, forming temporary hydrogen bonds, in search of the promoter sequences[53]. Once a promoter is found, a segment of the helix must be open. RNA polymerase forms a complex with the section of open DNA[54]. RNA synthesis, unlike DNA synthesis, starts de nova. After forming the first phosphodiester bond, the sigma factor is lost. The newly synthesised RNA forms a hybrid helix with the template DNA strand[52]. This RNA-DNA helix is about 8 bp long, but about 17 bp of DNA are unwound throughout the elongation phase[67]. Termination is facilitated by stop codons, which are often from a palindromic GC-rich region followed by an AT-rich region[67].The palindromic sequence self-complements, forming a hairpin loop that causes stalling. During the termination phase of transcription, the formation of phosphodiester bonds ceases, the RNA-DNA hybrid dissociates, the melted region of the DNA rewinds and RNA polymerase releases the DNA (Figure 1.4)[50].

Initiation



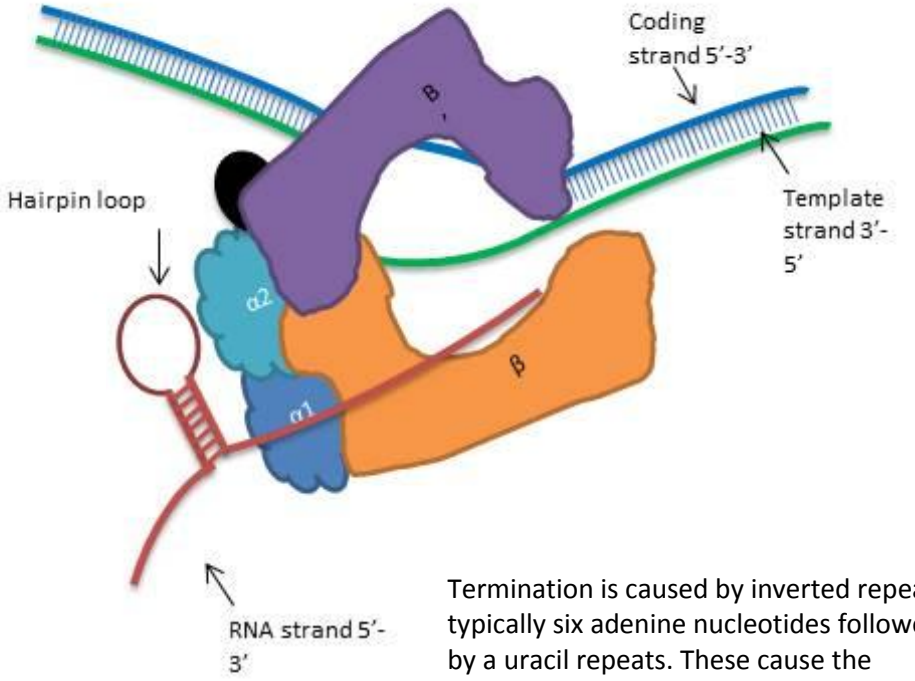
During initiation, -10 and -35 hexameric repeats in promoter regions are recognised by the σ factor. σ^{70} recognises the consensus sequences TATAAT and TTGACA. The β and β' subunits adopt a closed conformation during initiation. The DNA RNA polymerase open complex when the βσ subunit triggers melting of the DNA. NTP bind to the template strand, as subsequent NTPs enter they form phosphodiester bonds. When chain of NTPs reaches a length of 10pb the σ factor will fall away allowing closer association between RNA polymerase and the DNA molecule.

Elongation



During elongation the RNA chain will form a duplex with the template strand. Nucleotides are added to the 3' of the RNA chain at the active site. The active site is composed of sites j and j+l. RNA chain is situated in the j site. NTP are held in j+l site. Mg^{2+} initiates nucleophilic attack on the NTP allowing for the formation of a phosphodiester bond between the NTP and the RNA chain. The elongation complex will move one base pair down position in the newly added nucleotide in the j site. The RNA-DNA duplex will pivot as nucleotides are added. This allows the duplex to break apart allowing the RNA chain to exit and the DNA strands to reanneal.

Termination



Termination is caused by inverted repeats, typically six adenine nucleotides followed by a uracil repeats. These cause the formation of a hairpin loop in the RNA chain. This causes transcription to stall and the RNA chain to separate from the template strand and terminating transcription.

Figure 1.4. RNA polymerase's distinct subunits α2ββ'σ and RNA polymerase holding the open DNA (black) complex, and formation of the RNA(Red)-DNA duplex. Outlining the three steps of RNA synthesis, initiation, elongation and termination

RNA polymerase acts as a gateway for regulation. The steps in RNA synthesis can be regulated, though this regulation can affect the elongation and termination steps, such as through attenuation (conditional termination) and riboswitches[55], albeit most regulation happens in the initiation step. An important part of initiation regulation is the recruitment of sigma factors to the holoenzyme. While specific transcription factors allow *E. coli* to respond to discrete intra- and extracellular signals, sigma factors provide a much broader response by acting as dissociable subunits of RNA polymerase[54]. *E. coli* contains two structurally distinct sigma factor families: the sigma54 family recognises promoters that have conserved elements near -12 and -24 upstream from the transcriptional start site, while the sigma70 family is more numerous and includes the primary and essential sigma70 factor[56]. All bacteria contain a primary sigma factor, necessary for the

transcription of genes associated with survival and growth[54]. These primary sigma factors regulate the transcription of essential genes important for healthy cell growth[69]. The sigma70 family also contains alternative sigma factors that are important regulators of various stress responses[56]. *E. coli* has six alternative sigma factors, namely sigmaS, sigmaN, sigmaH, sigmaF, sigmaE and sigmaFecI (Table 1.1)[56]. An alternative sigma factor redirects RNA polymerase to promoters within their regulons. RNA polymerase cannot bind more than one sigma factor, so there is competition between sigma factors[57], which makes the tight regulation of alternative sigma factors necessary. Regulation is accomplished via a number of methods, including protein synthesis and stability.

Sigma factors	**Gene**	**Primary Function/regulon**
σ^{70}	*rpoD*	Housekeeping
σN	*rpoN*	Nitrogen-regulated genes
σH	*rpoH*	Heat-shock genes
σE	*rpoE*	Envelope stress response
σF	*rpoF*	Flagella synthesis/chemotaxis
σS	*rpoS*	Starvation/general stress response
σFecl	*fecL*	Iron transport

Table 1.1. The different sigma factors in *E. coli*, the genes encoding them and their function.

1.3.2 Regulatory RNAs

RNAs are often utilized as regulators within the cell. RNA regulators are less resource intensive for the cell to produce and do not need to be translated. They can be deployed quickly, to shut off gene expression. So, regulatory RNAs offer several advantages compared to protein regulators.

Small, non-coding RNAs (sRNAs) are an important regulatory tool ranging from a few nucleotides to hundreds[58]. The majority of sRNAs act post-transcriptionally and negatively regulate target genes by base pairing to the mRNA message[58].Great numbers of sRNAs are dependent on Hfq, which is an RNA binding protein. Hfq binding increases sRNA stability and, in many cases, it is the Hfq-sRNA complex which allows sRNAs to become active and bind to their targets[59]. The majority of sRNAs bind to the 5'UTR of their targets and often occlude the ribosome-binding site. This method, along with the increased frequency of RNaseE-mediated degradation, is the most common way of preventing the sRNA-mRNA duplex from being translated[60].

A minority of sRNAs activate expression of their targets. This is achieved by sRNA binding disrupting secondary structures of the target that would otherwise sequester the ribosome-binding site[61].

sRNAs can be either trans or cis-encoded. Trans-encoded sRNAs have limited complimentary with their targets[62], [63]. Hfq-binding is used to facilitate binding in spite of limited complimentary[63]. Cis-encoded sRNAs are encoded on the opposite strand of their target genes, because of this they are, largely, complementary to their targets. Although, many, cis-encoded sRNAs have the ability to function in trans[63].

Certain sequences at the 5'end of mRNAs that can adopt different conformations in response to environmental signals such as temperature changes, ligand binding, ribosome stalling, etc. are called "riboswitches," which have two parts, namely the aptamer (the ligand binding region) and the expression platform, the latter of which can adopt alternative secondary structures which regulate gene expression by affecting translation[64].

Examples of these secondary structures are hairpin structures that disrupt transcriptional terminators or antiterminators, or which occlude or expose ribosome-binding sites[63]. The majority of riboswitches require ligand binding to function.

1.3.3 Transcription factors

Transcription factors (TFs) influence the frequency which RNA polymerase binds to a promoter[65]. There are approximately 300 genes in *E. coli* that encode DNA-binding proteins known as "transcription factors"[66]. The majority of these proteins bind specific promoter or enhancer sequences[45]. Transcription factors are activated in response to ligand binding or covalent modification, which allows TFs to respond to endogenous or extra-cellular signals[65]. The active form of many TFs is a homo-dimer or homo-multimeric[65]. Endogenous signals, to which TFs respond, are often metabolites produced by the cell, while examples of exogenous signals are extremely varied and include metabolic by-products of other cells, ions or changes in temperature or osmolality[49]. Often TF networks will include transcription factors that are sensitive to both types of signals, this is logical as exogenous and endogenous signals are often closely linked. For example, the availably of a metabolite synthesised internally may depend on the availability of an extra-cellular precursor molecule[49]. LacI and TrpR are, well studied, examples of transcription factors that are regulated by exogenous and endogenous metabolites, respectively.

Most transcription factors can be characterised by their domains. The majority of TFs have two domains, a DNA-binding domain and a regulatory domain[65]. However, some TFs contain three domains and a minority contain one or four domains. 90% of transcription factors, found in *E. coli*, have a single DNA-binding domain, the other domains being regulatory or related to auxiliary function[65]. Two-component systems are a broad exception to this rule as the DNA-binding domain and the regulatory domain are on two separate proteins.

The majority of DNA-binding domains, in *E. coli*, can be grouped into 11 different families, which contain variations of the helix-turn-helix motif, alpha helixes joined by peptide chains[65]. Although this motif is common in DNA-binding domains it is not exclusive to them. The size of DNA-binding domain families varies, the largest family 'winged-helix' contains 123 members while TrpR is its family's singular member[65].

Transcription factors are regulated both transcriptionally, via expression and subsequently by their total protein concentration, but also via their regulatory domains. The non-DNA binding domains can be categorised by function. Small molecule-binding domains are the most common, second to enzymatic domains[65]. There are protein-protein interaction domain families, which facilitate binding with other TFs, RNA polymerase or covalent modification, and receiver domains found in the response regulators of Two-component systems[65]. These diverse Non-DNA-binding domains can be found disrupted among the DNA-binding domain families[65].

Transcription factors can affect gene expression via a number of mechanisms. Some TFs influence the activity of RNA polymerase by remodelling the holo-enzyme to recognise promoters[67]. For example SoxS is an 'appropriator' that is expressed in *E. coli* in response to oxidative stress, which binds to the C-terminal domain of both α-subunits of RNA polymerase and promotes transcription of genes with specific sequences called Sox-boxes[67].

The majority of transcription factors affect gene expression by interacting with the promoter region of a gene, although the mechanism they use is diverse. The regulatory control that a TF can exert is dependent on TF concentration and TF-DNA affinity[67]. TF can bind to consensus sequences with a strong or weak affinity, weak consensus sequences require a higher concentration of TF for regulation, while the opposite is true for strong consensus sequences[67].

TFs can positively or negatively regulate their targets[68]. They can block RNA polymerase through steric hindrance or by recruiting co-repressors[27] that decrease the affinity of RNA polymerase for the gene's promoter region[69].

Conversely, TFs can also recruit co-activators or increase the concentration of RNA polymerase for the gene's promoter[26].

To accomplish repression via steric hindrance TFs bind to operators, a short sequence of DNA, this allows the TF to physically block the -10 and -35 regions of the promoter[70]. Promoters often contain several operators next to each other, which allow several of the same TF to bind for an aggregative effect or different TFs to bind so gene expression can be mediated by multiple signals. Some TFs bind to distal operators and can affect the local DNA topology, causing loops that prevent RNA polymerase binding[70]. Some repressors function as 'anti-activators', which disrupt the activity of TFs that promote transcription, which is accomplished by a combination of DNA and/or protein binding[70].

There are several classes of transcription factors that enhance transcription (activators). Class I activators bind to operators and recruit RNA polymerase[70]. Class I operators normally bind positions –61, –71, –81 or –91 upstream of transcription start site[70]. The C-terminal domain of RNA polymerase's α-subunit interacts with a region on the activator (called the activation region), this protein-protein helps to stabilise RNA polymerase DNA-binding, specifically the transcription closed complex[70]. Class II activators bind to operators that overlap the -35 region of the promoter. They can interact with either the N-terminal domain of the α-subunit or the sigma factor of RNA polymerase[70]. This interaction can facilitate the transition of RNA polymerase-DNA from closed to open complex. A third class of activators bind to operators between the -10 and -35, promoter regions that bind activators of this class usually have non-optimal spacing between these sequences. Binding of these types of TFs helps to better orientate the region so the local DNA topology can better facilitate RNA polymerase binding[70].

TF can function as activators of some genes and repressors of others, for example some genes that are regulated by the Factor-for-inversion stimulation protein (FiS) have two adjacent TF operators[70]. A weak negative consensus sequence, that promotes repression, inside the promoter region and strong positive operator outside of the promoter which activates the gene[70]. In low concentrations, the TF

will bind the strong promoter which will activate transcription, however when the transcription factor is abundant the positive site will be saturated and weak consensus sequence binds the TF and repressor transcription[70].

TFs are organised into local networks, which represent interconnected regulations of target genes. In *E. coli*, the most common types of local networks are single input modules, in which a TF regulates a group of genes, wherein affinity for the TF can vary between genes in the group[71]. These systems are highly sensitive to the concentration of TF regulating them[72]. There are also multiple input modules in which genes respond to several TFs. In this way, the expression of a single gene can be influenced by multiple signals[72] and local transcription networks can take on different motifs, a common type of which is the feedforward motif. In this local network, a TF will regulate a gene and an intermediate TF, which also regulates the target gene[72]. The TF and the intermediate TF may have different effects on the target gene, i.e. one may be a repressor and the other an activator, or they may have the same effect, but the most common formation seen in this arrangement has both the TF and the intermediate TF acting as activators[73]. This ensures gene transcription only when a signal is persistent, thus preventing fluctuation. The second most common type of configuration observed is one in which the TF is an activator and the intermediate is a repressor[72]. This type of configuration forces the pulse-like expression of the target gene. These local networks are integrated into the greater whole, forming a global network of TFs.

There is a huge difference in the number of genes that TFs regulate, ranging from a large number of genes to just one only. The global TF network is organised hierarchically, wherein 20% of *E. coli* TFs regulate the transcription of the majority of genes in the global TF network[72], including TFs associated with redox sensing (ArcA), iron transport (Fur), environmental sensing (CspA) and carbon metabolism (Lrp)[72]. The majority of TFs are fine-tuners and regulate a small number of genes[72]. These networks function to maintain homoeostasis and help the cell rapidly respond to changes in the environment.

1.3.4 Two-component systems

Cells have developed complex signal transduction systems that allow them to respond to both exogenous and endogenous stimuli, assuring the tight control of transcription, translation and, ultimately, phenotypic changes to help them adapt to adverse growth conditions[74]. This system enables the cell to adapt its phenotype in response to stress and deprivation by coupling gene expression and intracellular states to changes in gene expression[75]. One of the most important ways this is achieved in *E. coli* is via two-component systems (TCSs), which are activated in response to one more extracellular stimuli and usually have large and complex regulons that overlap with many other regulators and transcription factors.[76] TCSs usually consist of an, integral, membrane-bound sensor kinase and a cytoplasmic DNA-binding response regulator[74]. In most instances, the membrane-bound sensor kinase will autophosphorylate a histidine residue when it is activated by a specific signal[29], which will then cause the cytoplasmic sensor kinase domain to phosphorylate a key aspartic residue of the response regulator, thus causing a conformational change[29]. This conformational change allows the response regulator to bind efficiently to DNA [74], [77]. These systems are very common and exist in all domains of life[78]. In *E. coli*, there are 29 histidine kinases and 32 response regulators[77].

Phosphotransfer – the phosphorylation of the response regulator by its cognate histidine kinase – is an integral part of the TCS activation process. In the majority of these cases, phosphoryl group transfer happens on a one-to-one basis, though one-to-many (one histidine kinase activating many response regulators) or many-to-one (many histidine kinases activating a single response regulator) transmissions are not uncommon[79]. The latter type of system allows different signals or stressors to activate the same gene or the coordination of multiple signals so that subthreshold activation of many pathways triggers the response regulator.
Specificity between histidine kinases and their cognate regulators is important, because unwanted cross-talk between histidine kinases and non-cognate regulators can cause noise during signal transduction, and there is impetus to avoid this to

maintain the integrity of signal transduction[79]. The specificity of phosphotransfer is driven primarily by molecular recognition and the specificity of the histidine kinase[74], which is dictated by key amino acid residues that are necessary for docking the kinase with the response regulator[80].
While response regulators often hydrolyse on their own, they often have dedicated phosphatases, even those with short phosphoryl group half-lives[79], which prevents unnecessary noise in signal transduction. In many cases the response regulator's cognate histidine kinases will act bi-functionally as both kinases and phosphatases. However, many TCS are also regulated by other phosphatases[81].

Prototypically, histidine kinases are homodimeric transmembrane proteins, with each protomer having several domains. The sensor domain differ greatly between sensors kinases, though some common structural folds infer a conserved method of signal sensing[82]. The extracystolic sensor domain is connected to two transmembrane helices[34], the N-terminus of which are located in the cytoplasm and connected to the HAMP domain. The HAMP domain, approximately 50 residues in length, consists of two α-helices (referred to as "Hα1" and "Hα2"). The HAMP domain is then connected to the dimerisation and histidine phosphorylation domain (DHp), which consists of two helices called "Dα1" and "Dα2". Finally, at the C-terminus of the second DHp domain, is the catalytic (and ATP-binding) domain. The DHp and catalytic domain is referred to collectively as the "kinase domain," which has several conserved residues, such as a conserved histidine residue that is phosphorylated and donates a phosphoryl group to activate the response regulator (Figure 1.5)[74]. Signal transduction through the sensor kinase starts by stimulating the sensor domain, which causes a cascade of conformational change that starts with the sensor domain, following which the signal is then transmitted through the transmembrane helixes to the HAMP domain and then finally to the kinase domain[77], [82]. Signal transduction through the transmembrane helixes is thought to occur via a screw-like motion[82].

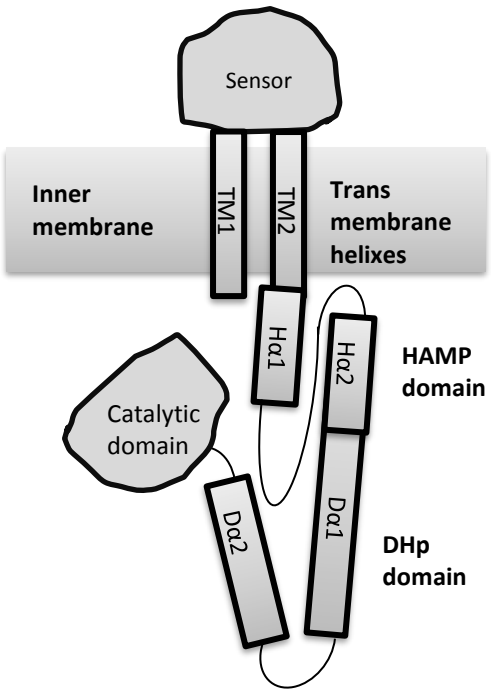


Figure 1.5. Example of a prototypical sensor kinase protomer. Showing the extracytosolic sensor domain, the transmembrane helices (TM1&2), the HAMP domain, consisting of two α-helices (Hα1&2), the dimerisation and histidine phosphorylation (DHp) domain, which consists of "Dα1" and "Dα2", and the catalytic domain. These domains are connected by small linker polypeptides. Adapted from Wang et al. 2012[82].

Response regulators have two domains: a receiver domain (this domain accepts the phosphoryl group from the sensor kinase) and an effector domain. Some response regulators, such as CheY, only have a receiver domain, and conformational changes in the single domain allow for interaction with the target[78]. Receiver domains have well-conserved structures[83], while effector domains have more diverse functions, reflected by their structural diversity. The defining feature of the receiver domain is the conserved aspartate phosphorylation site, as this motif accepts the phosphoryl group from the sensor kinase[83].

All receiver domains adopt a conserved alpha/beta fold, in which a five-stranded beta sheet is surrounded by five amphipathic alpha helices, with helices 1 and 5 on one side and 2,3 and 5 on the other[83]. Alpha helix 1 is essential for binding to the DHp domain of the sensor kinase. Many conserved residues are situated at the C-terminus of β1 and β3. Here, at the end of the β3 strand, the conserved aspartate residue which functions as a phosphoacceptor, essential for phosphorylation of the response regulator, can be found[83]. Moreover, the two Asp (or Glu) residues that

follow β1, along with the Asp residue between β3 and α3, are also important for phosphorylation. These three residues allow for divalent cations, Mg^{2+} or $Mn^{2+,}$ binding which is essential for phosphorylation[83].
β4 ends with a conserved Thr/ Ser residue, along with the Phe/Tyr residue associated with β5, these two residues facilitate signal transduction by interacting with the phosphoryl group delivered to the β3 Asp residue and are known as switch residues[83]. The switch residues facilitate the conformation change, of the response regulation, necessary for activation(Figure 1.6)[83].

These conformational changes, propagated through a series of highly conserved residues, cause the receiver domain to homodimerise or affect protein-protein interaction and allow the response regulator to activate. The majority of bacterial effector domains function as TFs with DNA binding capability [77], although others have a different mechanisms, such as enzymatic activity[83].

TCS with DNA-binding output domains can be sub-classified into three groups based on the mechanism through which they bind DNA, OmpR-like, NarL-like and NtrC-like, although significant diversity can exist between members of the same group [84]. NarL-like response regulators are autoinhibitory[84], while the phosphorylation of the NarL N-terminal receiver domain releases the otherwise sequestered C-terminal output domain that subsequently binds specific DNA promoter sites, in order to repress or activate gene expression[84]. NtrC-like phosphorylation causes a rearrangement from an inactive dimer to an altered active dimer[84]. OmpR-like phosphorylation causes dimerisation, leading to tighter binding of the target promoter, PhoP, from the PhoPQ TCS is an example of a response regulator from this group[85].

The second most common type of response regulator domain is involved in cyclic-di-GMP signalling[84]. Cyclic di-GMP is an important messenger in bacteria, helping to regulate many functions such as biofilm formation, motility and virulence[86]. These, non-DNA binding domains exhibit enzymatic activity, they can synthesise cyclic-di-GMP and are called “GGDEF” domains[84].

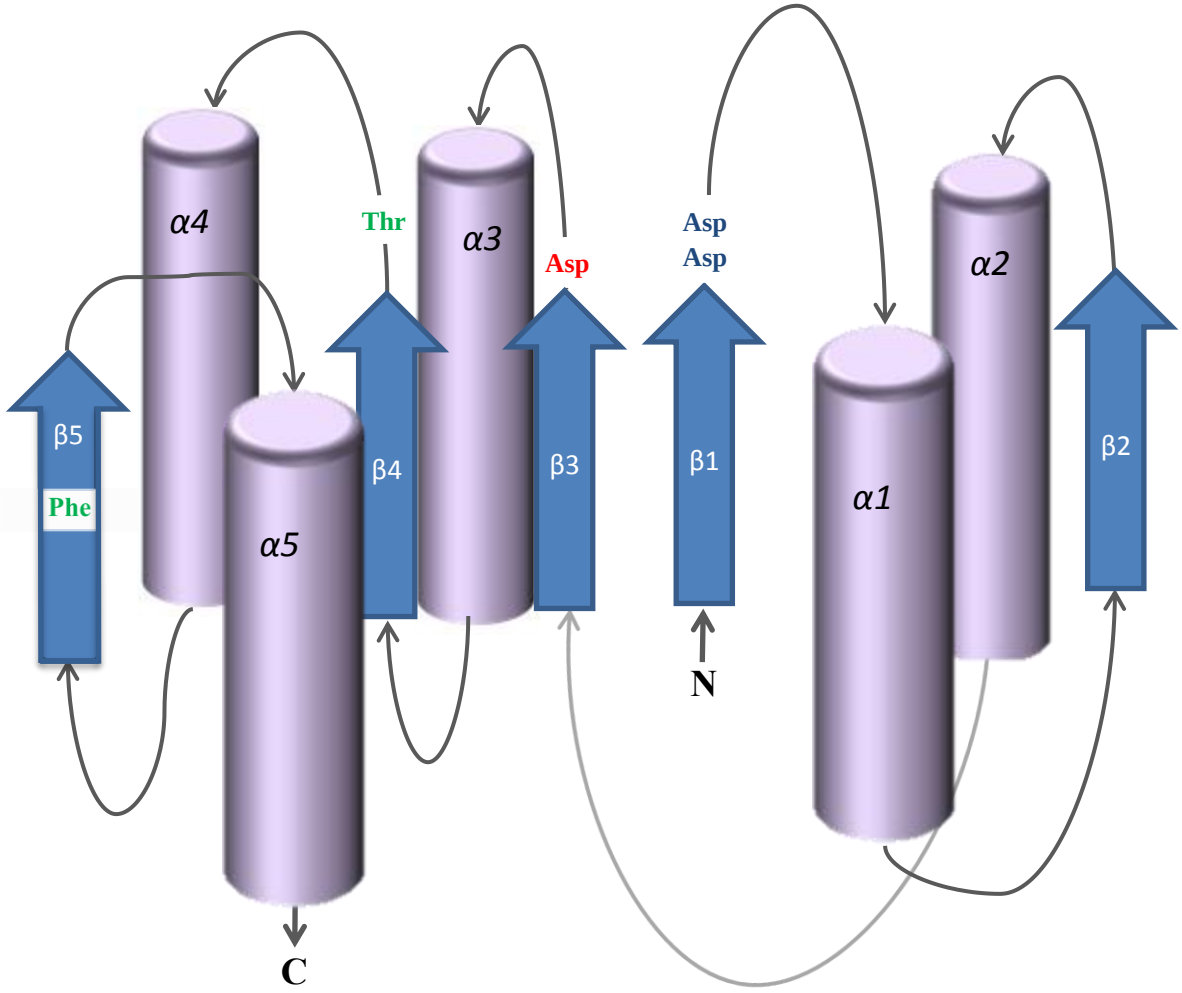


Figure 1.6. Cartoon representation (not to scale) of a prototypical response regulator receiver domain. α-helixes (α1-5) are depicted as purple cylinders, and β-sheets (β1-5) are represented by thick blue arrows. The N-terminus and C-terminus are denoted by N and C, respectively. α-helixes and β-sheets are linked by amino acid sequences portrayed as grey arrows. Where grey arrows cross, the lighter arrow indicates it is further from the foreground. Specific amino acids outlined in the text are denoted by three-letter abbreviations. The phosphoacceptor aspartate (Asp) residue, following β3, is highlighted in red. Two metal ion-binding aspartate (Asp) residues following from β1 are highlighted in light blue. Switch residues, denoted above as threonine (Thr) and phenylalanine (Phe), situated between β4 and α4 and in the middle of β5, respectively, are highlighted in green. Adapted from Bourret et al.[83].

1.4. PhoPQ

PhoPQ consists of the membrane-spanning sensor kinase, PhoQ, and PhoP, the cytoplasmic response regulator[39]. First discovered in *Salmonella enterica*, PhoQ was initially thought to be a regulator of PhoN, an acid phosphatase [87], [88]. PhoPQ showed homology to other two-component systems, and it was soon

discovered that it consisted of a single transcriptional unit and that PhoP and PhoQ were transcriptional regulator and membrane-bound sensor kinases, respectively. Later, PhoQ homologs were found in *E. coli*. The *E. coli* homolog of PhoPQ is 86% identical to protein first discovered in *Salmonella enterica* [76]. PhoPQ responds primarily to low concentrations of Mg^{2+}, though other metal ions, Ca^{2+} and Mn^{2+}, also induce PhoPQ, albeit to a lesser degree [88]. *phoPQ* is expressed as a polycistronic transcript from one of two promoters. P1 is active in a PhoP-dependent manner, while P2 is constitutive, ensuring that there is a low level of *phoPQ* expression when Mg^{2+} concentrations are high[88].

Like many sensor kinases, PhoQ exists in *E. coli* as a transmembrane homodimer and contains four main regions. The cytoplasmic histidine kinase and HAMP domains[89], the periplasmic sensor domain and membrane-spanning antiparallel helices(Figure 1.7)[89]. Periplasmic Mg^{2+} starvation will initiate signal propagation, which in turn causes the autophosphorylation of PhoQ from ATP at the conserved histidine 277 residue, situated in the cytoplasmic kinase domain[89]. This is followed by the transfer of the phosphoryl group to the conserved aspartate 51 residue on the PhoP receiver domain[89]. PhoQ can also act as a phosphatase, causing dephosphorylation of the PhoP aspartate 51 residue when Periplasmic Mg^{2+} is prevalent[89].

1 MKKLLRLFFPLSLRVRFLLATAAVVLVLSLAYGMVALIGYSVSFDKTTFRLLRGESNLFYTLAKWENNKLHVELP 75

76 ENIDKQSPTMTLIYDENGQLLWAQRDVPWLMKMIQPDWLKSNGFHEIEADVNDTSLLLSGDHSIQQQLQEVREDD 150

151 DDAEMTHSVAVNVYPATSRMPKLTIVVVDTIPVELKSSYMVWSWFIYVLSANLLLVIPLLWVAAWWSLRPIEALA 225

226 KEVRELEEHNRELLNPATTRELTSLVRNLNRLLKSERERYDKYRTTLTDLTHSLKTPLAVLQSTLRSLRSEKMSV 300

301 SDAEPVMLEQISRISQQIGYYLHRASMRGGTLLSRELHPVAPLLDNLTSALNKVYQRKGVNISLDISPEISFVGE 375

376 QNDFVEVMGNVLDNACKYCLEFVEISARQTDEHLYIVVEDDGPGIPLSKREVIFDRGQRVDTLRPGQGVGLAVAR 450

451 EITEQYEGKIVAGESMLGGARMEVIFGRQHSAPKDE 486

Figure 1.7. The PhoQ protein sequence taken from *E. coli* MG1655. The N-terminus is highlighted in blue, transmembrane helixes in orange, the sensor domain in green, the HAMP domain in grey and the kinase domain in purple. The conserved histidine 277 residue is highlighted in green[89]**.**

PhoP is an OmpR-like response regulator, part of the OmpR/PhoB subfamily of response regulators, characterised by an N-terminal receiver domain and a C-terminal helix-turn-helix DNA-binding motif[74]. Once phosphorylated, the active PhoP dimerises and modifies gene expression by binding directly to the gene promoters within its regulon. The PhoP protein comprises two domains, namely a DNA-binding domain and a receiver domain. While the crystal structure of the receiver domain has been resolved, the secondary structure of the DNA binding domain still requires elucidation(Figure 1.8).

1 *MRVLVV*EDNALLRHHLKVQIQDAGH*QVDDA*EDAKEADYYLNEHIP*DIAIV*DLGL*PDE*DGLSLIRRWRSNDVSLP*IL* 76

77 *VLT*ARESWQDKVEVLSAGAD*DYVT*KPFHIEEVMARMQALMRRNSGLASQVISLPPFQVDLSRRELSINDEVIKLT 151

152 AFEYTIMETLIRNNGKVVSKDSLMLQLYPDAELRESHTIDVLMGRLRKKIQAQYPQEVITTVRGQGYLFELR 223

Figure 1.8. The PhoP protein sequence taken from *E. coli* MG1655. The receiver domain is highlighted in blue, the α-helix in bold and β-folds underlined in italic. The DNA-binding domain is highlighted in pink and the conserved aspartate 51 residue is highlighted in green.

Among *E. coli*, there are a number of common PhoP homologues that are found in both pathogenic and non-pathogenic isolates (Appendix Figure F.1). The most frequently encountered example is one in which leucine residue 45 is replaced with isoleucine, though a fraction of these may also have arginine residue 65 replaced by cysteine.

PhoP is well conserved among the order enterobacteriales (Figure 1.9), particularly among members of the Enterobacteriaceae family[90]. However, PhoPQ homologues that perform a similar role, namely enhancing survival in low Mg^{2+} conditions, have been well characterised in *Pseudomonas aeruginosa*[91], *Mycobacterium tuberculosis*[92] and *Vibrio* species[93].When comparing the PhoP sequence of *E. coli* MG1655 to PhoP homologues found in other species of the order enterobacteriales, it can be seen that the *E. coli* K-12 MG1655 PhoP sequence is somewhat conserved (Figure 1.9).

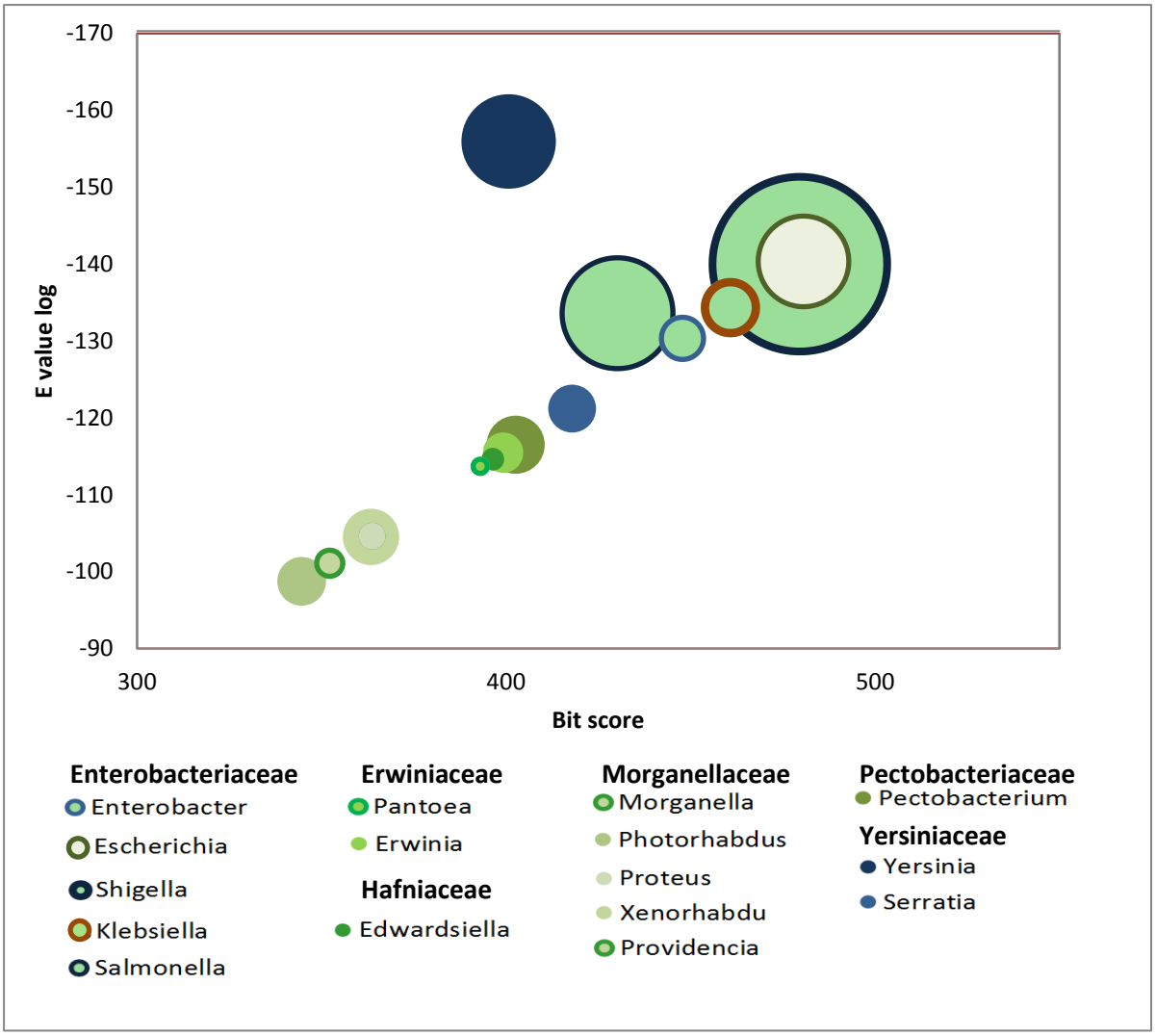


Figure 1.9. Bubble graph comparing similarities between different PhoP homologues found in species of the enterobacteriales order. The data were generated from a HMMER analysis of PhoP sequences taken from MG1655, using tools available at ebi.ac.uk/Tools/hmmer along with an E-value threshold of <1E-30. The diameter of the bubbles represents the range of scores and of sequence found in each species,as absolute value. Scores range from 344.6 to 479, which is equivalent to 86% and 100% similarity to the *E. coli* MG1655 PhoP sequence.

1.4.1 The PhoPQ regulon

PhoPQ's most obvious role in the cell is to assist with metal ion homeostasis during divalent cation starvation. *E. coli* has a number of Mg^{2+} transporters, namely CorA, MgtE and MgtA. PhoPQ regulates the P-type TPase magnesium transporter MgtA[94]. CorA and MgtE are constitutive Mg^{2+} transporters, while MgtA is only expressed when the cell is starved of Mg^{2+}[94].

However, PhoPQ has an expansive regulon, playing a role in, acid and antimicrobial peptide resistance, and LPS modification[42].PhoP regulates a large number of genes indirectly through other TCSs, regulatory sRNAs and transcription factors. It has been found that in a *phoP* mutant, under non-Mg^{2+} starved conditions, the expression of 31 genes significantly altered compared to the wild type[95]. Four genes were upregulated and 27 genes were downregulated[22]. However, microarray experiments carried out in a low Mg^{2+} environment found that the expression of 232 genes was altered in a PhoPQ-dependent manner[42].

Analysis of the PhoPQ TCS showed that it binds to a distinct sequence, (T/G)GTTTA-nnnnn-(T/G)GTTTA, which constitutes the PhoP box, which is often found -35 upstream from the start codon[96]. Of the 232 genes found to have their expression influenced by PhoPQ, 26 of these genes were found to have the PhoP box tandem repeat[42]. These 26 genes have a variety of functions, including pili chaperones, utilisation of alternative carbon sources, helicases, lipoproteins and LPS modification[5]. Moreover, it was found that PhoP binds to these genes with different affinities[42]. PhoP bound *mgtA* and *vboR* with the highest affinity, and *hemL* and *nagA* with the lowest affinity[42]. Out of these 26 genes, nine, *phoPQ*, *mgtA*, *mgrB hemL*, *nagA*, *rstAB*, *slyB*, *vboR* and *yrbL,* were denoted by Minagawa et al. as the Mg^{2+} stimulon in *E*. *coli*[42].

PhoP exerts indirect regulatory control over most of its genes in part due to cross-talking with otherTCSs. It was found that PhoP-regulated genes share similar expression profiles with a number of different TCS. EvgAS-PhoPQ and PhoPQ-yedWV were found to have the highest regulatory overlaps, though the signal to which EvgSA responds is unknown [97]; however, it has been found that cells in which EvgsA is constitutively expressed show an increased multi-drug resistance phenotype and acid resistance[98]. A small inner membrane spanning protein, SafA, connects EvgS and PhoQ and facilities signal transduction between the two systems[99]. Thirteen PhoPQ-regulated genes show enhanced expression when EvgAS is active: *crcA, hemL, mgtA, ompT, proP, rstA, rstB, slyB, ybjG, yrbL, mgrB* and *phoPQ* itself[94].

While PhoPQ is not considered as a virulence determinant in commensal strains of *E. coli*, it has been found to be associated with virulence in some pathotypes[1]. UPEC strains (CFT073) lacking PhoPQ were attenuated and no longer able to cause infection in mice. It was found that, in this strain, 36 flagella and chemotaxis genes (not normally found in commensal strains) and acid fitness genes were all repressed in the absence of PhoPQ[90]. Furthermore, PhoPQ has an established role in resisting cationic antimicrobial peptides (CAMPs) in both pathogenic and non-pathogenic strains.

CrcA, regulated by PhoPQ, transfers a palmitate residue from a phospholipid to the proximal unit of lipid A. Acylation of lipid A prevents insertion of CAMPs, and subsequent disruption, of the bacterial membrane[100]. CAMPs are able to bind to the hydrophilic region of LPS via electrostatic interaction[100]. Once bound CAMPs 'flip' so that the hydrophobic region is buried in the lipid A layer of outer leaflet of the OM. Acylation of lipid A helps to diminishes the rate of insertion of CAMPS into the lipid A layer by increasing hydrophobic interaction between neighbouring lipid A acyl tails[100].

It has been found that *phoP* mutants are unable to survive treatment with mastoparan or cecropin P1, while survival is diminished following exposure to magainin 2. However, survival, when treated with melittin, was not dependent on *phoP* in *E. coli* but it was in *Salmonella enterica* serovar Typhimurium (*S*. Typhimurium)[101]. It has been suggested that the difference in the PhoP-dependent response to antimicrobial peptides might be due to the proposed difference in cross-talk between the TCS PmrAB and PhoPQ in *E. coli* and *S*. Typhimurium[102].

In S*. enterica*, the protein PmrD allows for cross-talking between PhoPQ and PmrAB, the latter of which is a TCS that senses cationic antimicrobial peptides, high Fe^{3+} concentrations and acidic PH[101]. PmrAB also regulates LPS modification and resistance to antimicrobial peptides and regulates genes such as *eptA* and *arnT*. These genes encode phosphoethanolamine (pEtN) transferase and a 4-amino-4-

deoxy-L-arabinose (L-Ara4N) transferase, respectively, and these enzymes are able to modify the lipid A group of the LPS, by adding pEtN or L-Ara4N groups[103]. This helps diminish the charge on the lipid A anionic phosphates, reduces the net charge of the LPS molecule and helps to protect the outer membrane from attack by positively charged cationic antimicrobial peptides[103]. In *E. coli*, it was found that PmrAB-regulated genes were transcribed in a low Mg^{2+} environment in the absence of PhoPQ[103], therefore inferring that PmrAB is not dependent on PhoPQ in *E. coli*. Conversely, it has also been found that PmrD does have a similar role in both *E. coli* and *S. enterica* and that a second, unknown system exists and is responsible for transcribing PmrAB under conditions of low Mg^{2+} in the absence of PhoPQ, thereby explaining the results found initially[101]. This suggests that a similar link between PmrAB and PhoPQ may exists in both *S.* Typhimurium and *E. coli.*

A link between these two systems, PhoPQ and PmrAB, in *E. coli* would not be surprising, as PhoPQ has a well-established role in regulating both LPS modification and resistance to antimicrobial peptides. Aside from CrcA PhoPQ also regulates negatively regulates EptB, a phosphoethanolamine transferase via the sRNA MgrR. When not repressed by the PhoPQ controlled MgrR, EptB will add a phosphoethanolamine group to the outer Kdo residue of the LPS[94], [104]. Interestingly, it has been observed that *E. coli* cells that are unable to perform this LPS modification are less virulent.

Additionally, PhoP also regulates *ompT*. This protein is activated under conditions of stress, heat, ethanol and highly denaturing conditions[27]. OmpT will degrade highly denatured proteins and it also favours antimicrobial peptides as substrates[105]. Unsurprisingly, SigmaH is also implicated in OmpT regulation. OmpT is negatively regulated by OmrA and OmrB, which in turn are activated by the TCS EnvZ-OmpR, and it has been associated with virulence factors in urinary tract infection and the regulation of nucleotide excision repair during the SOS response[105]. OmpT, a prototypical member of the omptin family, is a surface membrane serine protease with a 10-strand antiparallel β-barrel conformation and extracellular loops that extend beyond the membrane[36]. The enzymatic activity

of this protein was found to depend on the presence of LPS, a fully acylated lipid A, is a requirement for OmpT activation[36]. It has been suggested that this mechanism for OmpT activation is used to stop OmpT attacking cellular proteins while it is being translocated to the OM from the cytoplasm[36].

In addition to antimicrobial peptide resistance, PhoPQ plays a critical role in the control of acid resistance. *gadAB,* part of the glutamic acid decarboxylase (GAD) system, encodes glutamate decarboxylase proteins and protects *E. coli* from the toxic effects of low pH (pH 2.5 to 4.5)[106]. GadW acts as a transcriptional regulator of genes involved in the GAD system, while *gadW* is unregulated in response to acidic conditions. Moreover, *gadAB* transcription has also been found to be influenced by factors such as RpoS, CRP, HN-S and EvgA[107]. PhoP also regulates periplasmic proteins HdeA and HdeB, active under acidic conditions, these chaperone-like proteins mitigate aggregations of periplasmic proteins [96].

PhoPQ also plays a role in N-acetyl-D-glucosamine metabolism via its regulation of *nagA*, which encodes the protein Nacetylglucosamine-6-phosphate deacetylase[18]. This protein serves to catalyse the first cytoplasmic reaction of N-acetylglucosamine metabolism, which is a multi-step process that starts with N-acetyl-D-glucosamine-6-phosphate and through a series of steps can yield glucose-6-phosphate, an important metabolite which can be used in glycolysis or the synthesis of nucleotides via the pentose phosphate pathway[108].

1.4.2 PhoPQ regulation

Mg^{2+} has several important biochemical roles within the cell, one of the primary ones being stabilisation of ribosome. The 70s unit of ribosome of *E. coli* contains over 170 Mg^{2+} atoms, and Mg^{2+} is essential for its formation[109] by stabilising the secondary and tertiary structures of rRNA. Mg^{2+} starvation causes the disassociation of ribosomal subunits and the unfolding of rRNA[110]. Mg^{2+} (and Ca^{2+}) also helps neutralise the charge between LPS molecules on the outer membrane and acts as a cofactor for several enzymes, some of which are involved DNA replication, transcription and translation[109].

Mg^{2+} plays an essential role in regulating PhoQ. The periplasmic domain of PhoQ responds to low Mg^{2+} concentrations, the presence of which stabilises the inactive conformation of PhoQ[87] and releases the inhibition of the DNA binding domain on the PhoP response regulator[76]. PhoQ can also respond to Ca^{2+} starvation. The domain responsible for binding Mg^{2+} and Ca^{2+} seems to be distinct, and evidence suggests that the absence of both of these metal ions has a stronger effect on PhoPQ induction compared to each of these alone[42]. Once active, PhoP will act as a transcriptional regulator, enhancing its own expression[4].

Aside from Mg^{2+} concentration, PhoPQ is regulated post-transcriptionally by two sRNAs, i.e. the sigmaE-controlled MicA and GcvB[111]. GcvB is regulated positively by GcvA and negatively by GcvR. These proteins regulate the *gcv*THP operon, responsible for encoding enzymes for the glycerin cleavage system (important for maintaining glycerin concentrations in the cell)[111]. GcvB regulates approximately 20 mRNA targets, including the periplasmic transporter components DppA and OppA[39], [98]. MicA is positively expressed during extracytoplasmic stress, under the control of sigmaE, linking PhoP with the extracytoplasmic stress response[112].

PhoPQ is also regulated by MgrB, a PhoPQ-regulated protein sensitive to the redox state of the cell[113]. PhoPQ also auto-regulates itself from PhoPp1 and requires both PhoP and PhoQ for activity; conversely, a second promoter, PhoPp2, has shown to be constitutively actively independent of the phosphorylation state of PhoP, which provides a constant basal level of expression and provides a pool of unphosphorylated PhoP to be activated by PhoQ during cation starvation[115],[121].

1.5. Specific stress responses

Bacterial stress can encompass any physical or biological stimulus that hinders unrestricted growth. This broad definition exemplifies physical changes in temperature, pH, osmolality and biological changes such as nutrient deprivation[114], [115]. Indeed, nutrient-limited stress is induced as a cell

population moves from the exponential to the stationary phase[116]. During nutrient-limited stress, metabolic processes are shut down and the rate of growth is slowed [117]; in fact, bacteria rarely have access to ideal growth conditions outside of specially created laboratory environments. *E. coli* has many mechanisms that allow the organism to respond to single stresses (examples of which are mentioned above). These single stress-induced responses usually allow the bacteria to eliminate the stressor or diminish its effect while also repairing damage caused by the stressor[114], [117]. This is achieved by using a number of different methods pertaining to gene regulation and expression.

1.5.1 Heat and cold shock

High temperatures have a variety of effects on cellular physiology, such as breakage in DNA and damage to the cytoplasmic membrane and ribosomes[118]. However, the primary signal for the activation of the heat shock response (HSR) is unfolded proteins[118], which has been demonstrated by the overproduction of unfolded proteins activating the HSR, in the absence of a temperature shift [119]. When *E. coli* is exposed to heat above what is optimal for growth, proteins will be less likely to take the proper conformation. The protein structure is dynamic and can be influenced by environmental conditions, and so increasing the temperature can cause a shift in conformational equilibrium, which results in incorrect protein folding and protein aggregates[120]. The cell responds by upregulating a group of genes called “heat shock genes” (HSGs), which causes heat shock proteins (HSP) to be overexpressed, such as proteases and chaperones that mitigate the damage of miss or unfolded proteins[121]. Heat shock especially causes damage to DNA and rRNA, so many HSP function to protect these molecules as well as tune transcription and translation[122]. The HSR has several phases, induction-transcription of HSP, adaptation-, fine-tuning of the response consistent with the present stimuli, and finally steady-state heat shock proteins are maintained at a level appropriate for the growth temperature[118].

It has been observed that cells can pre-adapt to heat stress. Cells that have been exposed to a high, non-lethal temperature exhibit a longer lag phase when they are exposed to a lethally high temperature, compared with cells that have not gone through this process[123]. The primary regulator of the HSR is sigmaH, and like all sigma factors, it has a large regulon that includes genes responsible for transcription, translation, DNA modification, protein chaperones, protease, etc.[118].

As the temperature drops, cells must adapt to the changing biochemistry that low temperatures bring. Biochemical reactions slow down as liquid in cells becomes more viscous and water ionisation decreases (thereby decreasing the availability of H^+ and OH^- ions)[124]. This affects water diffusion and many biochemical reactions that rely on these ions. RNA secondary structures become more stable, ribosome assembly is reduced and DNA becomes more negatively supercoiled, these factors can have a negative effect on gene expression[124]. Additionally, membrane lipid fluidity is reduced[125].

In response to a downshift in temperature, *E. coli* employs a number of proteins, such as cold shock proteins (CSPs), which mitigate damage and loss of function[126]. Unlike the heat shock response, the cold shock response is not mediated by a single, dedicated transcription factor (such as sigmaH) but is the product of several inputs[124]. When the temperature decreases, CSPs are transcribed at an increased rate (making 10-13% of transcribed proteins under low temperature conditions)[124]. One of the most important CSPs is the protein CspA, which is a major cold shock protein and part of a family of homologous proteins, which includes, CspB, CspC, CspD, CspE, CspF, CspG, CspH and CspI[127]. The function of all of these proteins is not known, and it is in fact thought that some of them may be redundant. CspA binds both single- and double-stranded DNA and RNA[127], helping to unwind tightly folded nucleic acid molecules, destabilise secondary structures and melt double-stranded RNA molecules, thereby making them available for RNA polymerase[127]. CspA binding also hinders the RNaseE degradation of mRNAs[128], functions as a transcriptional activator and increases

the expression of a number of other cold shock proteins[127]. CspA also induces *rpoS*, linking the cold shock response to the general stress response[127].

1.5.2 Envelope stress

The cell envelope stress response is activated under conditions that inhibit or perturb the components or function of the cell envelope[129]. This can include stimuli that are associated with general stress, such as heat, ethanol, oxidative stress or even starvation[129]. Under such conditions, the envelope stress response will accompany cytoplasmic stress responses[116]; however, this response and cytoplasmic stress responses are distinct [129].

There are five major envelope stress regulators in *E. coli*: sigmaE, Cpx, Rcs, phage-shock protein (Psp) and BaeSR responses[129]. The BaeSR two-component system is induced by toxic compounds, such as indole, and regulates exporter genes and multi-drug transporters[130]. The Psp response is induced by a number of stressors such as ethanol, osmotic shock and the mislocalisation of secretion proteins (that form export porins) to the inner membrane rather than to the outer membrane[131]. Four genes are necessary for this response: *pspA, pspB, pspC* and *pspF*[131]. PspA inhibits PspF, and upon sensing stress PspF is released from inhibition in a PspB- and PspC-mediated manner[131]. PspF is a transcriptional regulator, which, once freed from inhibition, will upregulate PspA, and once activated, PspA will affect inner and outer membrane stability[131].

The CpxAR response is activated under diverse conditions, such as heightened concentrations of alkaline, pH, copper ions and misfolded periplasmic membrane mutations or aggregates[132].CpxAR regulates genes that directly relieve envelope stress, examples of which are periplasmic foldase, DsbA and PpiD and protease DegP[133]. These proteins refold or degrade damaged proteins[133]. Rsc helps to regulate capsule synthesis, but also plays a role in the envelope stress response. Rcs is activated in response to environmental changes such as increased osmolarity,

desiccation, and growth on inorganic surfaces and antimicrobial attack on the outer membrane, such as Beta-lactams or antimicrobial peptides.

SigmaE is an example of an essential alternative sigma factor, although suppressor mutations have been found[134]. SigmaE will respond to misfolded outer membrane proteins, but it also plays a role in the normal growth of the cell. During non-stressed conditions, sigmaE is constituently active at low concentrations as the cell wall grows and expands[134]. SigmaE protects the cell against effects from ethanol, overproduction of outer membrane porins, inactivation of periplasmic chaperones and damage incurred from heat[114]. Due to its extensive role in preventing extracytoplasmic damage, there is a lot of cross-regulation with other alternative sigma factors, such as SigmaH and sigmaS[135]. While both sigmaS and sigmaE are activated in response to misfolded proteins, sigmaE responds uniquely to misfolded OMP[134].

SigmaE has a large regulon, and its main function is maintaining the outer membrane[57]. RNA polymerase, bound with sigmaE, $\alpha_2\beta\beta'\sigma^E$, will recognise and transcribe at least 43 genes[136], some of which encode periplasmic chaperones such as DsbC, FkpA, HtrA, Skp and SurA[136]. SigmaE also regulates genes that encode proteins in the BAM complex[136] and genes important for LPS biosynthesis such as *lpxD*[136].

SigmaE is transcribed from an operon containing four genes, *rpoE, rseA,rseB* and *rseC*. These genes are co-transcribed into a polycistronic transcript. SigmaE is regulated primarily by the anti-sigma factor RseA – an integral inner membrane protein that inactivates sigmaE by steric hindrance[137]. The N-terminal cytoplasmic domain of RseA binds to sigmaE. This domain is necessary for sigmaE to bind RNAP and is in direct contact with the RseA protein. SigmaE and RseA bind with high affinity, and sigmaE is released from RseA sequestering only by proteolysis[137].

When certain outer membrane proteins are misfolded, due to stress or damage, this will reveal a conserved sequence on the carboxyl terminus of OM porins[137].

This sequence will be recognised by the PDZ domain of the DegS protease[137] DegS will degrade the periplasmic domain of RseA. Interaction between DegS and RseA is inhibited by RseB. RseB will prevent DegS binding. However, the action of RseB can be inhibited by LPS binding to RseA.

After cleavage by DegS, the remaining RseA-sigmaE complex will contain a Gln-rich sequence that will act as a substrate for RseP, thus allowing the complex's final degradation by ClpXP [137].

RseC, an integral membrane protein, is thought to act as an anti-anti-sigma factor. It has been found that increased induction of RseC diminishes the inhibitor effect of RseA on sigmaE. The effect of RseC is thought to be mediated through interactions with either RseA or RseB[138].

1.5.3 General stress response

SigmaS is considered the master regulator of the general stress response in *E. coli*.[116]. In *E. coli*, the stationary phase sigma factor, sigmaS, recognises promoter sequences similar to sigma70 [54]. It has been suggested that this overlap may have come about out of necessity, as many genes regulated by sigma70 are necessary for survival during starvation[139]. The regulatory overlap of these sigma factors demands a mechanism of control that allows for diffraction between them, which is done by selective promoter utilisation by the sigma factors. A/T-rich regions in the promoter or flanking regions of a sigma70/S-regulated gene force a preference for one factor over the other[139].

SigmaS is active when a culture moves from the log phase into stationary phase. The regulons controlled by sigmaS are diverse and are activated in response to a range of changing physiological conditions[116]. SigmaS is regulated transcriptionally, translationally and via the transient stability of the protein. Upon the induction of sigmaS, the cell becomes resistant to a variety of stressors, which offers cross-protection, thus suggesting that sigmaS activates many (perhaps all) of the genes it regulates, in response to every individual stress[140]. Evidence of this

effect is seen when carbon starvation, for example, leads to temperature or osmotic shock resistance[140]. Although *E. coli* has specific stress responses and regulatory genes for these responses, there is a core of stress response genes that only require sigmaS for activation[140]. It has been found that many of these core genes are induced in a sigmaS-dependent manner, but only after a specific type of stress[140].

Transcription of *rpoS* is low under optimal conditions, and it increases upon entering the stationary phase. Expression of *rpoS* is tightly regulated, and if transcription does occur during optimal growth, then translation is shut off and any sigmaS that is translated is rapidly degraded[141]. These regulatory blocks are inhibited by nutrient starvation (carbon, phosphate, magnesium) or specific shocks such as osmolarity, low temperature or DNA damage. Regulation of sigmaS exists at the intersection of many regulatory cascades[141].

Several sRNAs regulate sigmaS. The long 5' UTR of the *rpoS* transcript folds into a stem loop that occludes the ribosome binding site and minimises *rpoS* translation[142]. This inhibitory structure is overcome by trans-encoded sRNAs, DsrA (increases sigmaS activity during low temperatures), RprA (activates *rpoS* translation under conditions of osmotic shock) and ArcZ (stimulates sigmaS under anaerobic conditions)[142]. A fourth sRNA, OxyS, negatively regulates *rpoS* translation[142].

SigmaS transcription is also regulated by many proteins and TCS. CRP (cAMP receptor protein) acts as a transcriptional regulator for *rpoS*[143]; moreover, it is a global regulator and primarily controls the utilisation of glucose as the primary carbon source, such as genes that function to transport and catabolise non-glucose sugars[143].

The BarA/UvrY TCS is induced during the exponential phase and induces *rpoS*[144]. The ArcAB TCS negatively regulates *rpoS* during the exponential phase[145], and ArcAB regulates the transistion from aerobic to anaerobic growth[145]. ArcB inhibits the transcription of *rpoS* by binding to sites near its promoter, thus

inhibiting CRP binding. Interestingly, ArcA can negatively regulate *rpoS* in the absence of ArcB[145], which itself regulates RssB (a regulator of sigmaS proteolysis)[145] and forms a branched one-to-many TCS. The ArcB-RssB system regulates sigmaS via protolysis[146]. SigmaS is degraded via the ATP-dependant protease, ClpXp, which degrades sigmaS in an RssB-dependent manner. When active (phosphorylated) RssB can bind to sigmaS[146], this RssB-sigmaS complex enables sigmaS proteolysis by ClpXP[145]. RssB-dependent ClpXP proteolysis can be disrupted by a number of anti-adapters such as IraM and IraP[147], which interferes with sigmaS-RssB binding, thereby preventing proteolysis. IraP and IraM is regulated by a number of regulators, such as PhoPQ and ppGpp, for example[147]. The stringent response utilises (p)ppGpp as a molecular messenger for many starvation conditions[148]. Aside from regulating IraP, (p)ppGpp has a positive effect on *rpoS* expression. Evidence suggests that (p)ppGpp has a stimulatory effect on *rpoS* transcription and mRNA stability[148].

1.6 Horizontal gene transfer and bacteriophages

In the environment, horizontal gene transfer is carried out by three main mechanisms, namely transformation, conjugation and transduction[149]. Horizontal gene transfer, the act of sharing genetic material between non-prodigy cells, is important for pathogenesis and environmental adaption. During conjugation, donor and recipient cells come into physical contact, using conjugation pilus, to allow transfer. During transformation, naked DNA is taken up from the environment[149]. Transduction is the transfer of DNA, using a bacteriophage vector. However, there are other methods of horizontal gene transfer. Some bacterial species have been observed fusing their OM, at which point bidirectional gene transfer ensues[149]. Another example are gene transfer agents, which can transfer random genetic material from a recipient to a donor by delivering them via a capsid, and it is thought that gene transfer agents are derived from bacteriophage that can no longer package their own DNA[149].

Horizontal gene transfer provides an important mechanism by which bacteria can adapt to their environment, thereby enabling the transfer of virulence determinants, antimicrobial resistance and evolution. For example, when the recipient cell uptakes foreign DNA, it can integrate into a recipient chromosome via homologous recombination or via non-homologous recombination[149]. When genetic material is transferred to a recipient cell that has no genetically similar material, additive transfer can take place[149]. If the recipient already has an homolog of the gene, then it may provide increased fitness or allow divergence of function[149].

Horizontal gene transfer mediated by bacteriophage has been integral to the development and evolution of pathotypes and virulence factors in *E. coli*. The artefacts of this genetic transfer can be observed in *E. coli* pathotypes. Incomplete bacteriophage genomes are observable in many pathotypes genomes[1]. Moreover, the Shiga toxin (Stx), produced by EHEC strains, is encoded on a bacteriophage in the EHEC genome[1]. This bacteriophage is fully functional and capable, under the right circumstances, to undergo lytic growth, thus spreading the toxin throughout a cell population[1]. Bacteriophages have also been implicated in the spread of genes resistant to antibiotics, such as tetracycline ampicillin, bleomycin, quinolone and β-lactams, found in the genomes of bacteriophage in several different environments, including hospitals[150].

1.6.1 The lytic and lysogenic cycles

The lytic and lysogenic cycles represent two outcomes of bacteriophage infection (Figure 1.10). Infection by tailed bacteriophage starts with absorption, which is facilitated by viral surface structures that bind to surface molecules on the host cell. In gram-negative bacteria, many surface proteins such as oligosaccharides or LPS (heptose core) can act as receptors for different bacteriophages[151]. Often, these receptors need to be clustered at a specific concentration so that phage tail can be positioned correctly to penetrate the host cell's surface. Some bacteriophages utilise primary and secondary receptors to bind, while many require cofactors for binding. In most cases, these are divalent cations such as Ca^{2+} or Mg^{2+} that help

overcome the competing charge on the host membrane and bacteriophage. Divalent cations can also cause the structure of bacteriophage to shift, readying it for infection, and extend attachment fibres to attach itself to the host[152].

Some receptors are more prevalent under certain environmental conditions, which can affect the efficiency of infection[151]. Host cell populations can develop resistance to bacteriophages by selecting for cells with altered receptors or altered expressions of receptors used by bacteriophages. However, this can be problematic and have a detrimental effect on the cell, as many bacteriophage receptors are essential. Bacteriophages can compensate for host cell's altering receptors in a number of ways. For example, P1 encodes multiple versions of the tail fibres[151].

Once attached to the host cell receptor, the bacteriophage will bind irreversibly. In most cases, the tail tip of the bacteriophage has enzymatic activity that allows for penetrating the peptidoglycan sacculus and inner membrane, ensuring DNA is released directly into the cytoplasm[152]. This DNA can be transferred into the host cell at a rate as high as 3000–4000 base pairs per second, which contrasts to other genetic transfer mechanisms such as conjugation (100 base pairs per second)[151]. DNA inside bacteriophages is densely packaged, inside the procapsid, by an ATP-powered nanomoter. This places the DNA under high pressure, which facilitates DNA injection into the cell[152].

Once inside the host cell, viral DNA is susceptible to host exonuclease and restriction enzymes. Bacteriophages have evolved numerous ways to protect themselves from this threat, such as by lacking restriction enzyme sites in their DNA, using "odd" nucleotides in their DNA (hmdU, hmdC) or rapidly circularising their DNA by means of terminal redundancies[152]. After the DNA is in the cytoplasm, cellular metabolism is redirected to the replications components of the bacteriophage. The host RNA polymerase will transcribe early bacteriophage genes from strong bacteriophage promoters. These early genes help protect the bacteriophage by inactivating host proteases and restriction enzymes, or the production of new sigma factors, to redirect host RNA synthesis. Middle genes that synthesise new bacteriophage DNA are then transcribed. Finally, late genes are

transcribed which encode bacteriophage structural proteins. Bacteriophage DNA is then packaged into procapsids. Assembly of the mature virion involves complex interactions between scaffolding proteins and major viral structural proteins. The viral head is the starting point for virion assembly, serving as a docking site for nucleic acid-packing enzymes and binding of the bacteriophage tail.
The final stage is host cell lysis. Many bacteriophages have two components to facilitate lysis – a lysin capable of lysing peptidoglycan, and holin, a protein that creates pores in the inner membrane for the delivery of lysin[152].

Conversely, many bacteriophages exhibit a temperate response. A prophage is the dormant bacteriophage DNA inside a host cell, and a host cell that carries a prophage is called a “lysogen”[152]. During this kind of infection, bacterial cell and bacteriophage maintain a stable relationship that can remain viable through subsequent generations. The temperate response requires the modulation of bacterial and phage DNA growth on a molecule-for-molecule basis. During lysogeny, the bacteriophage DNA exists as either an integrated part of the bacterial genome or a low copy plasmid. The close relationship between host and viral DNA means prophages are susceptible to mutation, in the same way as host cell DNA, and can become inactive and inert or a beneficial part of the bacterial genome. It is also possible for bacteriophage DNA to exist as an episome, a process called “pseudolysogeny”[152].

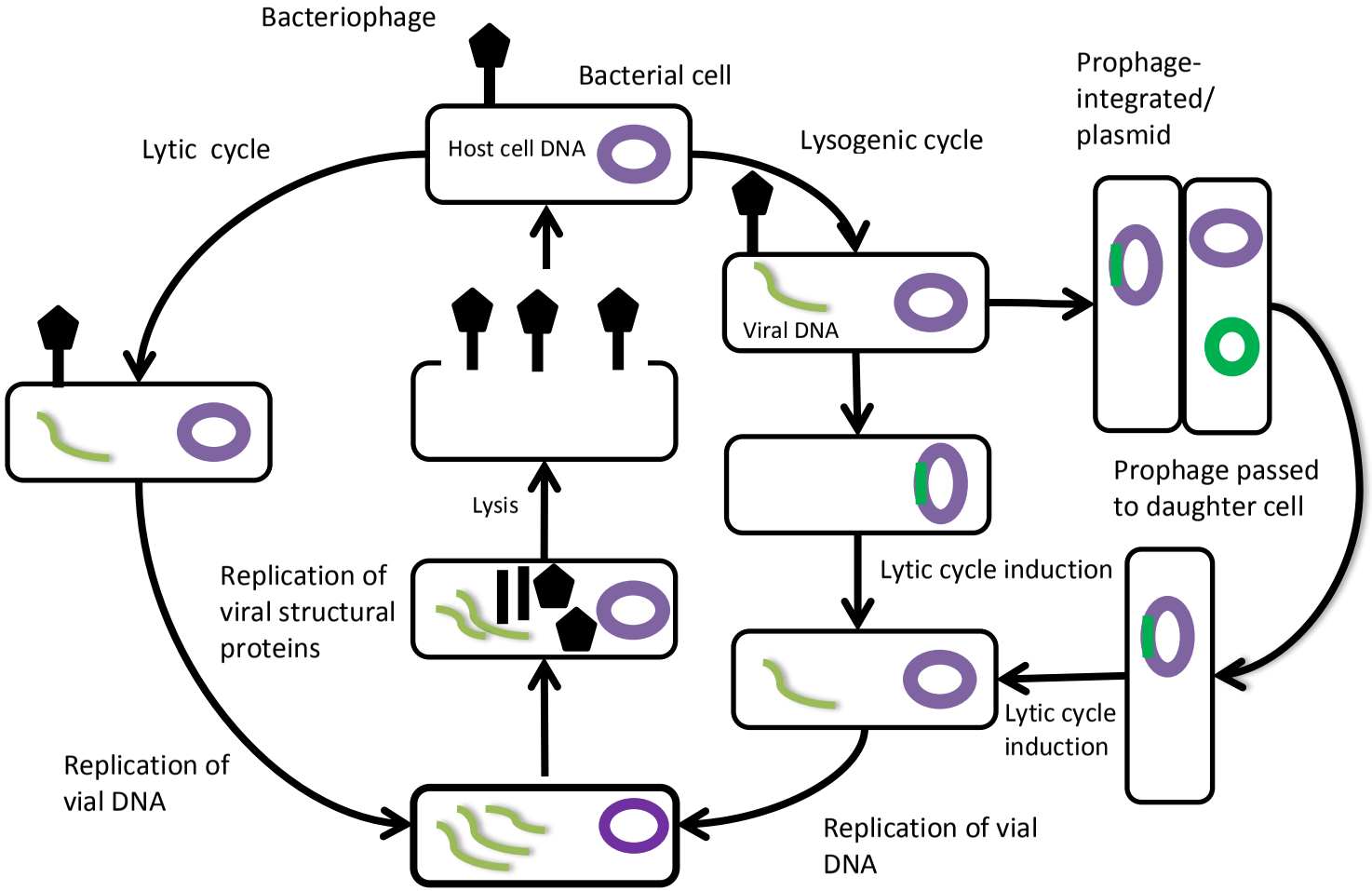


Figure 1.10. The lytic and lysogenic cycle. Adapted from Bacteriphages Aman et al.[269]

The initial stages of the infection of host cells that undergo lysogeny are the same as those found in the first stages of the lytic cycle, starting with bacteriophage absorption into the cell.
Lysogens are normally homoimmune, i.e. immune to infection by the same phage, but they are nevertheless susceptible to injection by heterologous bacteriophages[151]. Host cells can carry multiple (usually heterologous bacteriophages) prophages, a state which is called "multiple lysogeny"[151]. Early genes are transcribed, encoding products such as lytic repressors (C1) and integrase protein that allow for the integration of viral DNA into the bacterial chromosome[151]. As mRNA synthesis declines, the bacteriophage will physically insert itself into the host DNA to form a single integrated molecule[151].

The CI repressor binds a set of operators, repressing the expression of all bacteriophages genes, save its own[151]. Multiple lysogeny is able to occur because most prophages use repressors, which bind prophage DNA and inhibit the transition from lysogenic to lytic infection. Repressors also act on the DNA of new bacteriophages that arrive in the cell[151]. It is possible for lysogenic prophages to switch spontaneously into a lytic state. The reasons for the switch are not completely understood, but it happens at a rate that yields approximately 10^6 bacteriophages particles per millilitre. In any lysogenic culture of moderate density, lytic bacteriophage will always be found[151].

During temperate response, most bacteriophage proteins are not produced, and so the host cell survives the infection[151]. The survival of the host cell is important for temperate bacteriophages, so any detrimental effect on it in the early stages of infection must be mild or reversible. Effects, such as the degradation of the nucleoid, are not observed in temperate bacteriophages, which requires a balance when utilising the host's biochemistry[152].

1.6.2 P1*vir* transduction

Transduction was first reported in 1952, when genetic exchange occurred between different auxotrophic mutants[153]. It was observed that this genetic exchange

differed from other conjugations or transformations, in that it did not require physical contact between a donor and a recipient organism, and it was not disrupted by treatment with DNase[153]. During transduction, foreign DNA is transferred from a host bacterium to a recipient, using a viral vector. There are two types of transduction: generalised, in which the transducing bacteriophage can carry almost any part of the host cell genome[154], and the contrasting specialised transduction, in which the bacteriophage can only move specific parts of the bacterial genome. Transducing (bacteriophage) particles are by-products of normal bacteriophage metabolism[154]. Generalised transducing particles are not normally associated with viral DNA and may contain any segment of the bacterial genome, although not all sections can be packaged and transduced with the same frequency[154]. Specialised transducing particles are associated with lysogeny, and transduced bacterial DNA is always associated with viral DNA[155]. Only bacterial DNA adjacent to the integration prophage will be transduced, while bacteriophage and bacterial DNA are covalently linked[155].

P1 is a temperate bacteriophage and can infect a broad range of hosts. It is the major generalised transducing phage used in *E. coli*[154]. Structurally, P1 has an icosahedral head, a 22nm sheath and six tail fibres. The tail tube has a contractile sheath. A variable part of the tail fibres, encoded by an invertible segment of DNA, determines host specificity[154]. The P1 genome is approximately 93Kbp, double-stranded DNA[156] and contains approximately 117 genes, the majority of which are protein-coding and involved in lytic development[154].

P1 uses terminal glucose on the host cell's LPS, to enter the cell. During infection, calcium ions are necessary for attaching virions to the host cell[156]. Upon entry to the cell, DNA circulates via homologous recombination. The P1 genome has a high percentage of terminal redundancies, i.e. 9-12% (15kbs), compared to other bacteriophages such as P7 (which only has 1% terminal redundancies)[152]. The repeated sequences have been implicated in helping the genome move from a linear to a circular state[156].

Whether or not the bacteriophage will enter into lytic or lysogenic development is dictated by a number intrinsic and extrinsic factors, though this is not completely understood. Environmental factors, the state of the cell and the availability of nutrients interplay with the expression of the C1 repressor [152]. P1 bacteriophages used for P1*vir* transduction are usually mutants lacking the C1 repressor, which prevents them from entering the lysogenic cycle[156].

P1 has several replication origins. Genes expressed during lytic development do so from lytic origin (oriL)[156]. The replication of DNA is initially done via the circular mechanism but later transitions to the rolling circle mechanism. Viral DNA is replicated as concatemers, a protein that recognises pac sites on concatemers and cuts making double-stranded ends[156]. The end of the DNA is packaged inside the head until it is full[156], following which another is packed; five heads full of DNA can be packed sequentially from a single pac site at 100% efficiency; however, efficiency gradually decreases[156]. Often, extra genes are packaged into the head. P1 has a variable assembly mechanism, which means a significant minority of P1 heads are smaller and thus are defective (80% have a head diameter of 85nm, while 20% have a diameter of 65nm)[152]. These defective heads carry a smaller proportion of the genome, but these defective bacteriophage can still form a progeny virus if several complementary variants infect the same host cell. When the head is filled with DNA, a double-stranded cut is made and finally a tail is attached. Once the complete virions are assembled, the host cell is lysed, thereby releasing the viral particles[156].

P1 is able to form transducing particles, which contain bacterial DNA rather than that of the bacteriophage. These transducing particles are released with lytic bacteriophages upon host cell lysis. *E. coli* contains pseudo-pac sites that are recognised by the P1 enzyme pacase[154]. Pseudo-pac sites exist in the bacterial genome at a substantially lower frequency than in the bacteriophage genome[156]. Sections of the bacterial chromosome that are cut by the viral pacase are packaged into phage heads. Transduction can vary up to 25-fold, depending on the transduction marker[152].

There is little specificity, so any part of the host genome can be packaged with almost equal frequency[154]. Some factors, such as distances away from pac sites, affect the frequency a certain segment of DNA will be packaged inside the bacteriophage head. Once inside the recipient cell, DNA is delivered to the recipient cell, as if it were the host viral DNA. This host DNA delivered by the bacteriophage may be degraded by nucleases or integrated with the recipient cell DNA[154]. However, the rate of recombination is very low, approximately 2% of the transduced DNA is recombined into the genome of the recipient cell. The majority of transduced DNA exists in the cytoplasm in a form that does not degrade or recombine (abortive transduction)[154]. Maintenance of abortive transduction DNA is maintained by proteins that are attached to either end of the DNA during packaging and facilitate supercoiling in the DNA[154]. Homologous recombination of DNA from the donor to the recipient cell is facilitated via RecBCD pathways[157].

During homologous recombination, part of the DNA's molecule is exchanged for identical or similar genetic material from an endogenous or exogenous source. Homologous recombination is divided into three steps: presynapsis, synapsis and postsynapsis. During presynapsis, a single stranded region of DNA is generated[157]. Recombination takes place during synapsis, and recombination intermediates are resolved during postsynapsis[157].

The initial stage of homologous recombination, presynapsis, requires the formation of single-stranded DNA (ssDNA)[157]. *E. coli* has multiple pathways through which to initiate homologous recombination. Presynapsis, the production of a single ssDNA, can be initiated by the RecBCD pathway or parallel pathways utilising exonucleases, such as RecE or RecJ, and helicase, such as RecQ[157]. The RecBCD is an enzyme that is part of the main homologous recombination pathway in *E. coli*[157] and acts as a helicase that is able to unwind double-stranded DNA (ddDNA). During unwinding, RecBCD can also function as an exonuclease and degrade double-stranded DNA in an ATP-dependent manner[157]. It also acts as an endonuclease, degrading ssDNA. RecBCD, mediated homologous recombination is stimulated by Chi sites in DNA[157], which are short sequences (5'-GCTGGTGG).

RecBCD will move along the dsDNA molecule. As it moves, it unwinds the DNA's helix. When the enzyme encounters a Chi site, the enzyme cuts the unwound ssDNA and continues along the DNA molecule[157].

The ssDNA, nicked by RecBCD at the Chi site, is held by the protein RecA[158]. Strains deficient in this protein cannot perform homologous recombination. RecA binds to DNA as a polymeric fibre, which is necessary for the synapsis step, i.e. the pairing and exchange of homologous ssDNA, to take place[158]. The second DNA molecule also binds to the RecA filament[157], which is able to bind three or more strands of DNA, thus promoting exchange. When the homologous region to be exchanged is paired with the single-stranded region, together they may form a triplex (ssDNA paired with dsDNA)[158]. When strands are exchanged, it happens in the 5'-3' direction at approximately 3-10 bp/s[157], and as the region of exchange moves towards the ds region of the original DNA molecule, a holiday junction is formed[159].

In the final step, postsynapsis, the holiday junction intermediate is resolved[158]. The RuvC protein, an endonuclease[158], binds to holiday junctions and cuts at the cross-over points to yield two discrete DNA molecules[158].

1.7 Project Aims

PhoPQ's most publicised role involves mitigating the effects of Mg^{2+} starvation. However, its regulon is extensive and its scope comparable to that of a global regulator. PhoP regulates many disparate systems and aspects of cellular metabolism. In the initial stages of this project, a hitherto P1*vir* transduction defect associated with *phoP* was observed. The aim of this project was to confirm and characterise while also increasing confidence in the initial observation.

Transduction exposes the cell to divalent metal ion starvation as well as LPS disruption and subsequent extracytoplasmic stress. This work aimed to investigate the mechanism of the hypothesised transduction defect and consequently provide further insights into the role of PhoPQ in cellular function.

Chapter 2: Materials and Methods

2.1 Bacterial strains, plasmids and oligonucleotides

The strains detailed below, and described in Table 2.1, are derivatives of either *E. coli* K-12 W3110 or MG1655. Many of the strains were purchased from the Keio Collection, The Coli Genetic Stock Centre (CGSC) and are derivatives of the BW25113 strain[160]. Often, these strains were used to produce P1 bacteriophage lysate, which was then used to create mutants of either MG1655 or the laboratory wildtype (MG1655 *fimB-lacZ*). The *rpoH*P3-*lacZ* (CAG45114) reporter fusion strain and the *micA* and *rseA* mutants were obtained from Carol Gross [161]. To cure the Kan^R cassette from the Keio collection mutants, the plasmid pCP20 was used according to the standard procedure[162]. P1 phage lysates used in this study are listed in Table 2.2. Oligonucleotides used in this study are listed in Tables 2.4. Plasmids used in this study are listed in Table 2.3.

Strain number	**Genetic background**	**Origin of strain**
AAEC090	Δ*lacZYA*::*sacB*-Kan^R	(Blomfield et al. 1991)[163]
AAEC189	Λ- F- *endA1 Thi-1 hsdR17 supE44* Δ*lacU169recA mcrA mcrB* Δ*fimB-H*	(Blomfield et al. 1991)[163]
AAEC261A	Δ*lacZYA fimB*-lacZYA	(Blomfield et al. 1993) [164]
BGEC085	W3110	Laboratory collection
BGEC905	Δ*lacZYA* FimB-LacZ	(El-Labany et al. 2003)[165]
CAG25198	*nadB*::Tn10 ΔrseAlacX74 lambda (*rpoH*P3-*lacZ*)	Carol Gross (Guisbert et al. 2007)[161]
CAG45114	Δ*lacX74* lambda (*rpoH*P3-*lacZ*)	Carol Gross (Rhodius et al. 2006)[166]
CAG62192	*micA*::Cam^R	Carol Gross (Gogol et al. 2011)[167]
EG12976	*phoP*::Kan^R	Eduardo Groisman (et al. 2008)[74]
JW0052-1	BW25113 *surA::*Kan^R	Keio Collection (Baba et al. 2006)[160]
JW1115-1	BW25113 *phoQ*::Kan^R	Keio Collection (Baba et al. 2006) [160]
JW1116-1	BW25113 *phoP*::Kan^R	Keio Collection (Baba et al. 2006)[160]

JW1815-1	BW25113 *mgrB::*Kan^R	Keio Collection (Baba et al. 2006)[160]
JW2663	BW25113 *gshA*::Kan^R	Keio Collection (Baba et al. 2006)[160]
JW3818-1	BW25113 *rfaH::*Kan^R	Keio Collection (Baba et al. 2006)[160]
JW5437-1	BW25113 *rpoS::*Kan^R	Keio Collection (Baba et al. 2006)[160]
JW5660-1	BW25113 *eptB::*Kan^R	Keio Collection (Baba et al. 2006)[160]
KCEC1627	FimB-LacZ *gshA*	Joanne Roobal, KCEC928
KCEC3700	*rfaH*::Kan^R	This study; Bacteriophage transduction; JW3818-1 into AAEC261A
KCEC4032	*phoP*:Cam^R	This study; constructed via lambda red using strain AAEC090
KCEC4102	*rfaH,cured*	This study; Kan^R cassette curing; pCP20 KCEC3700
KCEC4138	FimB-LacZ *micA*::Cam^R	This study; Bacteriophage transduction; CAG62192 into BGEC905
KCEC4188	*surA*::Kan^R	This study; Bacteriophage transduction; JW0052-1 into AAEC261A
KCEC4192	FimB-lacZ*, gshA phoP:*Kan^R	This study; Bacteriophage transduction; JW1116-1 into BGEC905
KCEC4302	*phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into AAEC261A
KCEC4418	*rpoH*P3-*lacZ,rseA::Tet*^{R}	Alex moores Bacteriophage transduction; CAG25198into CAG45114
KCEC4418	*rpoH*P3-*lacZ rseA*::Tet^R	This study; Bacteriophage transduction; CAG25198 into CAG45114
KCEC4420	FimB-LacZ*,rseA*::Tet^R	Alex Moores Bacteriophage transduction (CAG25198 into BGEC905)
KCEC4425	FimB-lacZ,*phoP*,cured	This study; Kan^R cassette curing; pCP20 KCEC4642
KCEC4468	*surA,*cured	This study; Kan^R cassette curing; pCP20 KCEC4188

KCEC4534	*micA-lacZYA*	Alex moores (Moores et al. 2014)[168]
KCEC4603	*cydD,phoP*::Kan^R	This study; Bacteriophage transduction; JW1116-1 into MS109 *cydD* cured
KCEC4642	FimB-lacZ,*phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into BGEC905.
KCEC4729	*rfaH,phoP*::Kan^R	This study; Bacteriophage transduction; JW1116-1 into KCEC4102
KCEC4925	FimB-lacZ,*phoP*	This study; Kan^R cassette curing; pCP20 into KCEC4642
KCEC4990	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114.
KCEC4992	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC4994	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC4996	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5021	*micA-lacZYA, ΔlacZYA* Ω, *lac*UV5-*phoP*	This study; allelic exchange ; plasmid from KCEC4534 into KCEC5277
KCEC5022	*micA-lacZYA, ΔlacZYA* Ω, *lac*UV5-*phoP*	This study; allelic exchange ; plasmid from KCEC4534 into KCEC5277
KCEC5023	*micA-lacZYA, ΔlacZYA* Ω, *lac*UV5-*phoP*	This study; allelic exchange ; plasmid from KCEC4534 into KCEC5277
KCEC5123	*rpoH*P3-*lacZ phoQ*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1115-1 into CAG45114
KCEC5124	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28°C; JW1116-1 into CAG45114
KCEC5127	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5129	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C;

		JW1116-1 into CAG45114
KCEC5131	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5133	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5135	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5137	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5139	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5141	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5143	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5145	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5147	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5149	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction at 28 °C; JW1116-1 into CAG45114
KCEC5166	*rpoH*P3-*lacZ phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into CAG45114
KCEC5168	*rpoH*P3- *lacZ phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into CAG45114
KCEC5170	*rpoH*P3- *lacZ phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into CAG45114
KCEC5172	*rpoH*P3- *lacZ phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into CAG45114
KCEC5174	*rpoH*P3- *lacZ phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into CAG45114
KCEC5176	*rpoH*P3- *lacZ phoP*::Kan^R	This study; Bacteriophage

		transduction; JW116-1 into CAG45114
KCEC5178	*rpoH*P3- *lacZ phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into CAG45114
KCEC5180	*rpoH*P3- *lacZ phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into CAG45114
KCEC5239	*rpoH*P3-*lacZ,phoP, rseA::*Kan^R,Tet^R	This study; Bacteriophage transduction; JW1116-1 into KCEC4418
KCEC5259	FimB-LacZ*, phoP,rseA*::Kan^R,Tet^R	This study; Bacteriophage transduction; JW1116-1 into KCEC4420
KCEC5275	*gshA,phoP*::Kan^R	This study; Bacteriophage transduction; JW1116-1 into KCEC928
KCEC5277	*lacUV5-phoP* Δ*lacZYA*	This study; Bacteriophage transduction; KCEC5281 H into AAEC090
KCEC5278	*gshA,phoP*Δ*lac*Ω*lac*UV5-*phoP*	This study; product of allelic exchange between pAM001 and KCEC5281 H
KCEC5281 H	*gshA,phoP*,*lac*UV5Ω*sacB*-Kan^R	This study; Bacteriophage transduction; AAEC090 into KCEC5275
KCEC5340	*rpoH*P3- *lacZ mgrB*::Kan^R	This study; Bacteriophage transduction; JW1815-1 into CAG45114
KCEC5345	*rpoS,*cured	This study; Kan^R cassette curing; pCP20 into KECE5432
KCEC5349	*rpoH*P3-*lacZ rpoS*::Kan^R	This study; Bacteriophage transduction; JW5437-1into CAG45114
KCEC5354	*rpoH*P3-*lacZ rpoS, rseA*::Kan^R Tet^R	This study; Bacteriophage transduction; CAG25198 into KCEC5349
KCEC5409	*ΔphoP lac*UV5-*phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into KCEC5277 with IPTG.
KCEC54329	*rpoS,phoQ*::Kan^R	This study; Bacteriophage transduction; JW115-1 into KCEC5345with IPTG
KCEC928	FimB-LacZ*, gshA*::Kan^R	Laboratory stock (Joanne Roobal, BGEC905)

KECE4323	*phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into AAEC261A, after 48 hours of incubation
KECE4324	*phoP*::Kan^R	This study; Bacteriophage transduction; JW116-1 into AAEC261A, after 48 hours of incubation
KECE4484	*surA*,*phoP*::Kan^R	This study; Bacteriophage transduction; JW1116-1 into KCEC4468
KECE5432	*rpoS*::Kan^R	This study; Bacteriophage transduction; JW5437-1 into AAEC261A
MS109	MG1655 *cydD*::Kan^R	(Mark Shepherd) Laboratory stock, Shepherd Laboratory

Table 2.1. *E. coli* strains used during this study. All strains are derivatives of MG1655, exceptions, such BW25113 are specified.

lysate	**Donor strain**
*eptB::*Kan^R	JW5660-1
gshA::Kan^R	JW2663
Δ*lac*Ω*lac*UV5-*phoP*	KCEC5277
*lac*UV5*ΩsacB*-Kan^R	KCEC5281 H
*mgrB::*Kan^R	JW1815-1
micA::Cam^R	CAG62192
phoP::Kan^R	Eg12976
phoP::Kan^R (1)	JW1116-1
phoP::Kan^R (2)	JW1116-1
phoP::Kan^R (3)	JW1116-1
phoP::Kan^R (4)	JW1116-1
phoQ::Kan^R (1)	JW1115-1
phoQ::Kan^R (2)	JW1115-1
phoQ::Kan^R (3)	JW1115-1
phoQ::Kan^R (4)	JW1115-1
rfaH:: Kan^R	JW3818-1
*rpoS::*Kan^R	JW5437-1
*rpoS:*Kan^R	JW5437-1
rseA::Tet^R	CAG25198
sacB-Kan^R Ω *lacZYA*	AAEC090
surA:Kan^R	JW0052-1
Δ*gshA*, Δ*phoP* Δ*lac*Ω*lac*UV5-*phoP*	KCEC5278

Table 2.2. P1 bacteriophage lysate used in this study and the strains from which they were derived.

Plasmid	Relevant characteristics	Source	Purpose
pCP20	FLP+ Λcl857+ ΛP R TS::Amp^R, Cam^R	Laboratory stock, Cherepanov et al. 1995[162]	Curing procedure
PkD3	oriR6Kgamma, *bla*(ApR),*rgnB*(Ter):: Cam^R	Shepherd Laboratory	lambda red recombination
PKD4	oriR6Kgamma, *bla*(ApR), *rgnB*(Ter)::Kan^R	Shepherd Laboratory	lambda red recombination
pKD46	*repA*101(ts),*araB*p-*gam*-*bet*-*exo*, oriR101, *bla*(ApR)	Shepherd Laboratory	lambda red recombination
pIB462	pIB305:: Cam^R (Blomfield et al. 1991); homologies to *lac* sequences at the multiple cloning site	Laboratory collection, (Blomfield., et al. 1991)[163]	Construction of *lac*UV5-*phoP*
pA001	*lac*UV5-*phoP*:: Cam^R	This study, pIB462 derivative	Construction of *lac*UV5-*phoP*
pAM012	pIB502 *micA*-*lacZYA*	Laboratory stock, Dr Alex Moores[13]	Construction of *micA*-*lacZ*, *lac*UV5-*phoP*

Table 2.3. Plasmids used during this study, their purpose and relevant characteristics

Name	Purpose	Oligonucleotides (5'-3')
*phoP*_λ-forward	Used in the construction of lambda red	CACATAATCGCGTTACACTATTTTAATAATTAAGACAGGGAGAAATAAAAGTGTAGGCTGGAGCTGCTTC
*phoP*_ λ-reverse	Used in the construction of lambda red	GCGAGAGCGGGAAAAAAAGACGCAGTAATTTTTTCAATGGGAATTAGCCATGGTCC
Vector_*phoP*-forward	construction of *lac*UV5-*phoP* –pAM001	AGCTTCAGGGAGAAATAAAAATGCGCGTACTGGTTGT
Insert_*phoP*2-reverse	Used in construction of *lac*UV5-*phoP*–pAM001	GATCCTCATCAGCGCAATTCGAACAG
phoP-forward	Screening for the *phoP* gene	CTTTACCTCCCCTCCCCGCTGG
phoP-reverse	Screening for the *phoP* gene	CCAACAGAAAACGTACCCGCAGC
Cam^R-forward	Screening for the *caT* gene	CACTGGATATACCACCGTTG
Cam^R-reverse	Screening for the *caT* gene	CTGTTGTAATTCATTAAGCA
gshA forward	Screening for the *gshA* gene	TGGCACACTGGCAACAACAGGTC
gshA reverse	Screening for the *gshA* gene	CACGCAGCAGATTACGGTAGGCTTC

Table 2.4. Oligonucleotides used during this study; Screening primers

2.2 QIAquick kits

During the course of this project, several kits were used: QIAquick PCR Purification, QIAquick Gel Extraction, QIAprep Spin Miniprep and MinElute Reaction. When usage is specified, all procedures were done in accordance with the manufacturer's instructions[169]–[172].

2.2.1 Media, storage and growth conditions

Unless otherwise stated, *E. coli* was grown in Luria-Bertani broth, (2.2.2) pH 7, at 37°C, using a shaking water bath at 180rpm (New Brunswick Scientific, Innova 3100). A static incubator (DSI series 300DF) was also used to grow bacteria on solid and liquid media at the temperatures specified in the individual procedures. Strains to be stocked were grown in Luria-Bertani broth overnight. A total of 150µl of 100%, sterilized, glycerol was added to a cryotube, and 850µl of the overnight culture was added and stored at -80°C until needed.

Solutions used in this study were sterilised either by in-house autoclaving (Prestige Medical Autoclave 2100 Series at 121°C for 15 minutes) or by filter sterilisation, using a 0.2µm pore filter (Nalgene). Purification of the double-deionised water (ddH2O) used in solutions was made by multi-deionising and filtrating through Millipore MilliQ® cartridges and then autoclaving for further sterilisation. A static water bath (Grant Instruments Sub Aqua 26) was used for experiments not requiring agitation. Solutions were stored at the temperatures outlined in specific procedures. Media reagents and chemicals were standard laboratory-grade or higher and purchased from Sigma-Aldrich UK unless stated otherwise.

2.2.2 Luria-Bertani media

Luria-Bertani (LB) broth and agar were used as the default liquid and solid media, respectively. LB broth was prepared using the following reagents: tryptone 10g/L (Oxoid), 5g/L yeast extract (Oxoid) and 5g/L NaCl (thermo fisher) dissolved in ddH_20. LB agar was prepared in the same way with the addition of 15g/L nutrient agar (Oxoid). Both LB and LB agar were made in distilled ddH_20 and sterilised by autoclaving, using in-house equipment adapted from standard produces, as

previously described Sambrook et al. 1989[173]. Any additional supplements, such as antibiotics, were added in working concentrations outlined below (Table 2.5).

2.2.3 Sucrose media

Sucrose broth/agar was made in the same way as LB broth or agar, as detailed above. However, 150ml/L of ddH_20 was replaced with a 40% sucrose solution (prepared by dissolving sucrose 400g/L in ddH_20, filter-sterilised), added after cooling, to create a 6% sucrose medium.

2.2.4 Agar Supplements

During this work a number of supplements were added to agar. This was often done by creating a concentrated stock solution, which was then sterilised, and added to the agar after autoclaving.

When Isopropyl β-D-1-thiogalactopyranoside was used, a concentrated 1M stock solution was made using ddH_20. The solution was then filter-sterilised and then added to the molten agar after autoclaving to yield a concentration of 0.1mM.

A stock solution of 1M procaine was made using ddH_20. The solution was then filter-sterilised, and 10ml of agar was removed from 1L after autoclaving and replaced with 10ml of 1M procaine stock solution, which yielded LB-agar containing 100mM of procaine.

Antibiotics (sigma Aldrich) were dissolved into stock concentrations (Table 2.5) and filter-sterilised. Antibiotics were added to media after autoclaving, when necessary. The stock and working concentrations of antibiotics used during this study are outlined in the table below (Table 2.5).

Addition	Dissolve in	Stock solution	Working concentration
Ampicillin (Amp)	ddH_20	125mg/ml	125ug/ml
Chloramphenicol (Cam)	100% ethanol	34mg/ml	25ug/ml
Kanamycin (Kan)	ddH_20	25mg/ml	25ug/ml
Tetracycline (Tet)	70% ethanol	10mg/ml	25ug/ml

Table 2.5. Outlining the different concentrations and solvent used in the preparation of the antibiotic solutions used herein.

2.3 Centrifugation

For larger volumes (50ml+), an ultra-centrifuge (Beckman-Coulter preparative centrifuge Avanti) was used at 4°C, unless specified. Centrifugation of medium volumes (15ml-50ml) was done using a Sigma 2K15 centrifuge (rotor Sigma 256/97 Nr. 12149) set at 7°C, unless indicated. A bench-top micro-centrifuge (Eppendorf MiniSpin, rotor F-45-12-110), used at room temperature, was employed for smaller volumes (1.5ml-2ml).

2.4 P1*vir* transduction

This method and all other derivatives are based on the standard procedure (Miller et al)[174], which uses the P1 bacteriophage as a vector for generalised transduction, transferring DNA, a selectable genetic marker, from a donor cell to a recipient. The majority of strains in this project were constructed using this method.

It was necessary to first generate a lysate from the donor strain. The aim of this step was to obtain a lysate in which the target genetic marker (an antibiotic resistance cassette) had been packaged into transducing particles. An overnight culture of the donor strain (containing the desired selectable chromosomal marker) was grown in LB broth for between 12 and 16 hours at 37°C in a shaking incubator. In total, 200µl of the overnight donor culture was inoculated into 15ml of LB broth after the addition of 50µl of a 40% glucose solution and 50µl 1M $CaCl_2$. The inoculum was then incubated for 30 minutes using a shaking incubator at 37°C. In

total, 200µl of previously prepared phage lysate was added to the inoculum, which was then left to incubate in the same conditions for 2-3 hours. The time waited was to allow the bacteriophage in the previously generated lysate sufficient time to lyse the culture and generate mature bacteriophage that would carry the desired genetic marker of the donor cell. The previously prepared lysate had a different selectable marker (usually tetracycline) than the marker in the donor strain. This was to prevent selecting mutants generated by the transduction of DNA from the previously generated lysate. A no-phage control was also run in parallel. When the inoculum turned clear, compared to the control, 200µl of chloroform was added to the culture and the inoculum was vortexed. The inoculum was then transferred to a 50ml falcon tube and centrifuged at 4500rpm for 10 minutes at 4°C. Next, 5ml of the lysate was added to a 15ml falcon tube containing 100µl of chloroform. The lysate was then stored at 4°C overnight before use. The lysate presumably contains transducing bacteriophage, which carry various parts of the donor strain chromosome, including the selectable marker of interest.

The aim of the subsequent step was to transfer the selectable marker to the recipient strain. An overnight culture of the recipient strains was grown in LB broth at 37°C in a shaking incubator. Then, 200µl of the overnight culture was added to an equal volume of a 10mM $CaCl_2$ and 20mM $MgSO_4$ solution, which was used to enhance the infectivity of the bacteriophage. Next, 100µl of the mixture was aliquoted into three Eppendorf tubes, following which 10µl and 50µl of bacteriophage lysate were added to two of the Eppendorf tubes. The third Eppendorf, containing only the culture and $CaCl_2$, $MgSO_4$ mixture, was used as a control. In addition, 50µl of the bacteriophage lysate alone was added to a fifth tube as a final control. All Eppendorf tubes were then incubated in a static incubator for 30 minutes at 28°C. A solution of 0.1M sodium citrate was made, using LB broth (LB-citrate), and filter-sterilised, after which 1ml of LB-citrate was added to each Eppendorf. The citrate acted as a metal chelator forming Ca-citrate, thus lowering the infectivity of the bacteriophage and inhibiting exponential lysis of culture.

The Eppendorf tubes were then incubated statically for 2.5 hours at 37°C. Then, 100µl of the inoculum was spread onto two LB agar plates (containing the relative antibiotic) for each Eppendorf. The plates were then incubated overnight at 37°C. If any colonies were present the following day, they were re-streaked onto a fresh LB agar plate (containing the relative antibiotic). The re-streaks were then grown overnight at 37°C in a static incubator and stocked accordingly.

2.4.1 P1*vir* transduction-derived procedures

Throughout this work, several derivatives of the standard P1*vir* transduction procedure were used. All of these procedures were based on the original (2.4), albeit with some alterations.

P1*vir* transduction was conducted at 28°C. The lysate from the donor culture for this procedure was generated in the same way as the original (2.4).The initial steps of this P1*vir* transduction were carried out in same way as the original procedure (2.4). After 1ml of LB-citrate had been added to the Eppendorfs, incubation was carried out at 28°C for 2.5 hours rather than at 37°C. Next, 100µl of the inoculum was spread onto LB agar plates (containing the relative antibiotic) for each Eppendorf twice, in the same way as the original procedure, though the plates in this case were incubated overnight at 28°C. If any colonies were present the following day, they were re-streaked onto a fresh LB agar plate (containing the relative antibiotic). The re-streaks were then grown overnight at 28°C, In a static incubator, and stocked.

A P1*vir* "mock" transduction was conducted during this study, in which no lysate was used. An overnight culture was prepared in the same way as in the original procedure (2.4), and then 200µl of the overnight culture was added to an equal volume of a 10mM $CaCl_2$ and 20mM $MgSO_4$ solution and incubated at 28°C for 30 minutes in a static incubator. Next, 100µl of the mixture was aliquoted into three Eppendorfs and incubated at 37°C in a static incubator for 2.5 hours along with a control containing only overnight culture. A dilution series was made from 10^{-1} to 10^{-6}. Finally, 100µl of the 10^{-5} and 10^{-6} dilutions was spread onto plain LB agar plates, two for each dilution, and incubated overnight at 37°C.

Another derivative was carried out, in which P1*vir* transudation was done using protracted incubation. This procedure was the same as the original (2.4), but the final incubation was prolonged from overnight to 36 hours.

Finally, the P1*vir* transduction procedure was used as a basis for a citrate sensitivity test. Overnight cultures were grown in the same way as the original procedure (2.4), and then 200µl of the overnight culture was added to an Eppendorf containing 1ml of 1.0M LB-citrate. Another 200µl of overnight culture was added to a second Eppendorf containing 1ml LB broth. The Eppendorfs were then incubated for 2.5 hours at 37°C. A dilution series was made from 10^{-1} to 10^{-6}, and 100µl of the 10^{-5} and 10^{-6} dilutions was spread onto plain LB agar plates, twice for each dilution, and incubated overnight at 37°C.

2.5 Isolation of genomic DNA

The isolation and extraction of genomic DNA was done using GenElute Bacterial Genomic DNA Kits (Sigma-Aldrich)[171]. In total, 567µl of cells were added directly to 30µl 10% SDS (100ml/L) plus 3µl Proteinase K (4mg/ml) and incubated at 65°C for 2 hours. Then, 100µl of 5M $NaCl_2$ was added and mixed gently, following which an 80µl solution of hexadecyltrimethylammonium bromide (CTAB) was added. The mixture was incubated again at 65°C for 10 minutes, and then 80µl of chloroform was added. The sample was then centrifuged at 10Krpm for 5 minutes. The upper aqueous phase was extracted into a fresh Eppendorf tube. The extract was then mixed with an equal volume of chloroform and further centrifuged for 5 minutes at 10krpm.The upper aqueous phase of the sample was then removed and placed into a fresh falcon tube, after which 600µl of isopropanol was added. The sample was then pelleted for 2 minutes at 10krpm. The supernatant was removed and the pellet was washed with a 70% (700ml/L) ethanol solution. The pellet was then centrifuged for 1 minute at 10Krpm. The supernatant was then removed and the pellet left to briefly air dry, before being suspended in TE and stored at 4°C.

2.5.1 Isolation of plasmid DNA

A QIAprep Spin Miniprep (Qiagen) was used to isolate plasmid DNA. The kit was used as per the manufacturer's instructions. Overnight cultures were grown in 5ml LB broth and in the presence of the appropriate antibiotic, for selection purposes, and at the appropriate temperature for the plasmids (37°C or 28°C) in a shaking incubator. Overnight cultures were centrifuged for 3 minutes at 8000rpm at 28°C, and the resulting pellets were then processed through the Qiagen mini-elute protocol[172], which was followed until the final elution step, in which TE buffer was used to elute the DNA. Then, 50µl of the isolated plasmid DNA was stored at 4°C until required.

2.5.2 Restriction endonuclease digestion of DNA

Reactions and digest conditions, temperature and buffers were used in accordance with suppliers' instructions for each reaction[175]. All restriction endonucleases were purchased from either Promega or New England Biolabs. The reaction products were incubated for 3 hours at 37°C in a temperature-controlled heat block, unless otherwise stated. Purification was done using MiniElute reaction purification or a gel-purification kit (Qiagen)[176]. The constituents of the reaction are shown below (Table 2.6).

Reagent	Concentration of stock solutions	Volume	Final Concentration	Variation
Buffer	10X	2µl	1X	Either 2µl of a single buffer or 1µl of two buffers if two enzymes were used that required separate buffers.
Restriction enzyme	10 units/µl	2µl	1 unit/µl	Either 2µl of a single enzyme or 1µl of two enzymes was used. The concentration (Units/µl) can vary between manufacturers. The concentration and volume specified here refers to enzymes purchased from New England Biolabs, which were the most frequently used during this study.
DNA fragment	Variable	10µl-16µl	Variable	Depending on the size of the fragment different amounts of DNA was used. For plasmid DNA, 10µl was used for smaller PCR-generated fragments up to 16µl was used.
Nuclease-free water	-	0µl-6µl	-	Nuclease-free water was used to make the reaction volume up to 20µl, if needed.

Table 2.6. Reagents used for restriction enzyme digest.

2.5.3 Ligation

The ability of T4 ligase to create phosphodiester bonds between detween the 5'-phosphoryl and 3'-hydroxyl groups of separate double stranded DNA molecules, is a routinely utilised as a tool in molecular biology[177].

In this study PCR products and digested plasmid DNA were ligated using T4 DNA ligase (New England Biolabs or Roche), after digestion with restriction enzymes. A

total sample volume of 10µl was used for each reaction, including the DNA, 1µl of T4 DNA ligase and 1µl T4 DNA ligase buffer, the final volume was made up to 10 µl using Nuclease-free water. Ligations were performed as either double or triple ligations. For double ligation vector to insert ratio used was either 1:1 or 1:3. For triple ligation vector to insert ratio used was either 1:1:1 or 1:2:2. All ligation reactions were incubated at 16°C for 16 hours. The digested insert was ligated into the digested vector, and the product was then transformed into competent cells on plates with the appropriate antibiotic and incubated at either 28°C or 37°C overnight, depending on the sensitivity of the plasmid. If colonies were present, they were grown for 6 hours in LB broth containing the appropriate antibiotic. The plasmid was extracted from these cells using a Miniprep elute kit and then digested with the same restriction enzymes used to cut the vector and fragment initially. The digest was then imaged using gel electrophoresis. If the plasmids released an insert of equal size to the one ligated, the ligation was deemed successful. Correctly constructed plasmids were then stocked and stored at 4°C in nuclease-free water.

2.5.4 Production of chemically competent cells

Artificial transformation of competent bacterial cells allows for the *in vivo* uptake of foreign DNA from the immediate environment. In this work, competent cells were made using the calcium chloride method originally published by Hanahan et al. 1991[178].

Cells to be made competent were Inoculation into 10ml of LB and were grown overnight in a shaking water bath (180rpm) at 37°C. A dilution of the overnight (28µl into 28ml) was made in LB broth and was grown to an OD_{600nm} of 0.5. Five minutes before reaching the desired optical density OD, 3.75ml 100% pre-warmed sterile glycerol was added to the shaking inoculum. Once the desired optical density had been reached, the cells were transferred to a 50ml falcon tube and chilled on ice for 10 minutes, after which the cells were centrifuged (10 minutes, 4000rpm, 4°C). After centrifugation, the supernatant was poured off and the pellet was re-suspended in 25ml of ice-cold solution (0.1M $MgCl_2$ and 15% glycerol) and centrifuged (for 8 minutes at 3800rpm, 4°C). The supernatant was poured off again

and the pellet was re-suspended in 6.25ml of ice-cold T-salts solution (75mM CaCl2, 6mM $MgCl_2$, 15% glycerol) and incubated on ice for 20 minutes. After incubation on ice, the cells were centrifuged again (6 minutes, 3600rpm, 4°C). The supernatant was poured off and the pellet was re-suspended in a 1.25ml T-salts solution, and then aliquots of 100µl, poured into pre-chilled 1.5ml Eppendorfs, were stored at -80°C.

To transform chemically competent cells, 100µl of competent cells was added to 10µl of DNA on ice, by light pipetting. The DNA-cell mixture was incubated on ice for 30 minutes before being transferred to a 42°C water bath for 60 seconds. The DNA-cell mixture was then immediately returned to room temperature and 1ml of LB broth was added. The cells were mixed well by inverting the Eppendorf tube 4-5 times and then incubated for 1 hour at 37°C (28°C for temperature-sensitive plasmids). This gave the cells time to express any relevant antibiotic resistance genes. Next, 200µl of cells were plated out onto LB agar containing the relative antibiotic and incubated at the appropriate temperature overnight. Any colonies were then picked and grown in LB broth at the appropriate temperature, with the relative antibiotic, for between 6-12 hours and then stocked.

2.5.5 Lambda red re-combineering

This procedure was based on the method originally published by Yu et al. 2000 [179]. The aim of this procedure is to replace the wild-type copy of the gene with an antibiotic cassette (Figure 2.1), which it achieves by making use of a strain that contains pKD46, a plasmid that carries a recombinase system used to enhance recombination between a gene fragment and chromosomal DNA. pKD46 carries the λ red genes behind the araBAD promoter and is temperature-sensitive. The araBAD promoter makes the λ red genes L-arabinose-inducible. The λ red genes, *gam*, *exo* and *beta*, inhibit the *E. coli* RecBCD system from degrading double stranded DNA, convert the transformed gene fragment into single-stranded DNA and bind to single stranded DNA to facilitate its ability to integrate into the chromosome, respectively[180].

Primers were designed with a 50-70bp homology to the flanking regions (up- and downstream) of the gene that was to be replaced, adjacent to sequences of homology to an antibiotic resistance cassette (chloramphenicol). These primers were used to PCR a chloramphenicol resistance cassette from the plasmid PKD3, the aim of which was to create a PCR product containing the chloramphenicol resistance cassette with adjacent regions of homology to those flanking the target gene.

A frozen stock of *E. coli*, carrying temperature-sensitive plasmid pKD46, was grown overnight at 28°C in a shaking incubator. In total, 100µl of this culture was then added to 50ml of LB, containing 100mg/ml ampicillin and 10µl of 2mM L-arabinose (oxoid). The culture was then grown at 28°C to OD_{600nm} of 0.7 and then placed on ice for 20 minutes. The inoculum was then centrifuged for 10minutes at 3000rmp at 4°C. The pellet was then re-suspended in 5ml of ice-cold 10% glycerol this step was repeated three times. During the final centrifugation, the pellet was re-suspended in 200µl of a 10% glycerol solution. The mixture was then stored at -80°C until used.

The remaining part of this procedure was carried out on ice, unless otherwise stated. In all, 9µl of purified DNA PCR product was added to the side of the electro-cuvette, and then 40µl of electro competent cells (Thermo Fisher) was added to the cuvette, over the area with the purified DNA. The cuvette was then removed from ice and placed into the electroporator (MicroPulser™) and electroporated using the following conditions: (V) 2450, (µF) 25 (Ω) 200. Once the electroporation was complete, 1ml of LB was added to the cuvette. The mixture was then transferred to an Eppendorf tube and incubated, shaking, for 1 hour at 37°C. Next, 100µl of the mixture was plated onto LB agar containing the appropriate antibiotic. The plates were then incubated overnight at 37°C in a static incubator.

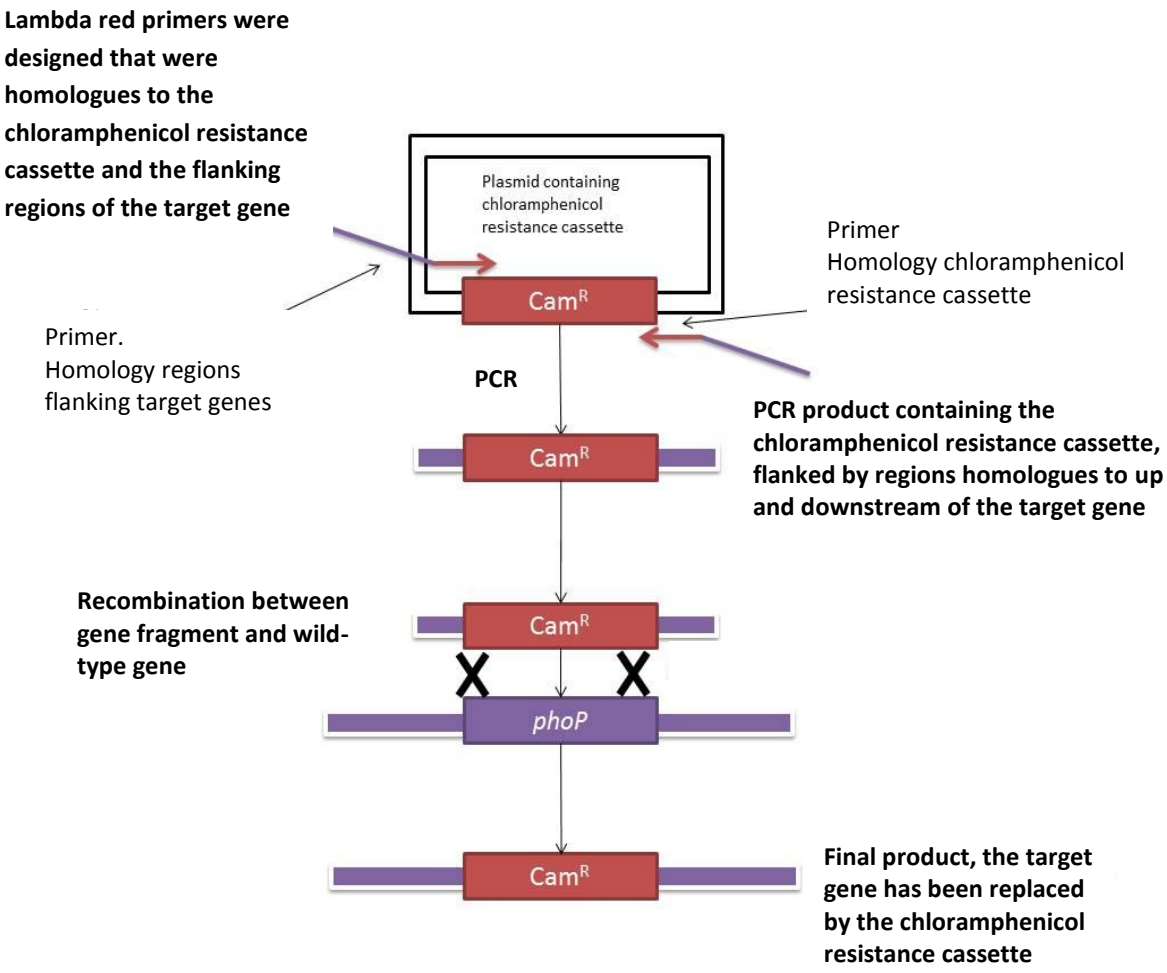


Figure 2.1. Showing the primer design necessary for lambda red recombineering. Generation of the fragment and recombination onto the chromosome. Cam^R refers to the chloramphenicol resistance cassette.

2.5.6 The polymerase chain reaction

The polymerase chain reaction (PCR) has come to be an invaluable workhorse of modern molecular biology[181]. In this study PCR was used to amplify required sections of DNA, where oligonucleotides (primers) were designed to flank the target DNA to initiate amplification. Oligonucleotides were synthesised by Eurofins MWG and stocked at 500µM in ddH2O at 80°C. The stock was then stored at -20°C. A working concentration of stock was made of 50pmoles per microlitre by diluting the original stock using ddH_20.

Reagent	Volumes added to the reaction mix
Forward primer	2.5µl
Reverse primer	2.5µl
ddH_20	17.5µl
Master mix Taq/Q5	25µl
Template DNA	Variable between (0.5µl -2µl)

Table 2.7. Describing the reagents used in the PCR procedure.

The master mix used was ReadyMix™ Taq (MFCD01635810) or Q5 High-Fidelity 2X Master Mix (M0492S, New England Bio labs). Each reaction was set up separately in Eppendorf PCR tubes, using filter pipette tips (Promega). ddH_20 was added first, then the template. The amount of template used was not always consistent but ranged between 0.5µl and 2µl, and was dependent one size of the template DNA. Forward and reverse primers, corresponding to sense and anti-sense strands of the desired amplicon, were added before the master mix (Table 2.7). All reactions were vortexed for 2 seconds before the PCR reaction started. The thermocycler was programmed (Table 2.8) using standard settings, unless otherwise stated.

Temperature (°C)	Time (seconds)	
94	240	
94	30	These steps was repeated 25 times unless otherwise stated
94	30	
60	45	
72	210	
4	∞	

Table 2.8. Describing thermocycler settings used for PCR

2.6 DNA purification

DNA was purified using the Qiagen PCR purification kit, where necessary, after PCR reactions[172]. Following endonuclease digestion, a Qiagen MiniElute enzyme cleanup kit was used, in accordance with the manufacturer's instructions[172], to remove endonucleases and purify DNA products.

2.6.1 Curing Keio collection mutants

The Keio collection is a library of strains containing mutations in non-essential protein-encoding genes, replaced by a Kanamycin-resistant (Kan^R) cassette. Multiple Keio collection gene deletions were combined by initially curing the Kan^R cassette and leaving a clean deletion, before transducing the second mutation. This was done using curing (Cherepanov and Wackernagel, 1995)[162]. with the Plasmid pCP20 (temperature-sensitive to 28˚C and Cam^R/Amp^R). This plasmid encodes the recombinase flippase (Flp), which recognises a pair of 34bp FRT sequences adjacent to the resistance cassette. Flp will bind these sequences and excise the adjacent sequence.

The strain, that was to be cured, was transformed using heat shock with the plasmid pCP20 at a ratio of 10µl/2µl. The mixture was then incubated at 28°C for 2 hours, following which 100µl of the transformants were plated onto amp plates and incubated overnight at 28°C. Colonies were then picked using a sterile metal loop and placed into 5ml of LB. The inoculum was then grown at 42°C for a minimum of 3 hours. The inoculum was then diluted down to 10^{-3} to 10^{-6} and each dilution was plated out in 100µl quantities onto plain LB agar and incubated overnight at 42˚C. Colonies were re-streaked (using a grid system with 6x6 squares) onto plain, Cam and Kan plates (36 colonies of each sample). The plates were then incubated overnight at 37°C. The colonies that grew on LB agar only were successfully cured. Colonies that were resistant to none of the antibiotics used herein were picked and re-streaked onto LB agar plates and grown overnight at 37°C. The colonies were then picked and grown in LB broth for between 6 and 12 hours and then stocked.

2.7. Allelic exchange

Allelic exchange, utilising homologous recombination, allows for the integration of plasmid mutation into the host chromosome carried on a plasmid vector, based on a procedure published by Blomfield, McClain et al. 1991[182] (Figure 2.2).

First, it was necessary to create an intermediate strain and plasmid vector, carrying the mutation of interest. The plasmid used (a derivative of pIB462) was temperature-sensitive, contained a Cam^R resistance cassette and had homology to *lac* flanking the multiple cloning site – these characteristics were necessary for counter-selection, plasmid curing and integration of the gene fragment into the intermediate strain chromosome. This plasmid contained the gene fragment to be exchanged into the recipient cell chromosome. The gene fragment was generated using PCR and cloned into the multiple cloning site of the plasmid. The intermediary strain was created by transducing the *sacB*-Kan^R cassette at the *lac* locus into the strain in which the final mutation was to be inserted.

During the first step, competent cells, made from the intermediary strain, were transformed with a plasmid derivative of pIB462 (containing the mutation of interest). In all, 100µl of competent cells were transformed with 10µl of plasmid, using the standard transformation procedure. After the transformation, the aim was to select for cells that had integrated the plasmid into their chromosome.

This was done by incubating the cells at 42°C overnight. Five pre-warmed LB agar plates were incubated at 42°C, and one plate was incubated at 28°C.

Tranformants were then plated out in 200µl volumes onto five pre-warmed LB agar plates. They were incubated at 42°C, and one plate was incubated at 28°C. The plate incubated at 28°C functioned as a control. The plates were then incubated overnight at their respective temperatures. After 16 hours, they were removed from the incubator.

The aim of the next step was to remove all cells containing the *sacB*-Kan^R cassette by growing on sucrose, thus selecting for successful excision. *E. coli* expressing *sacB*, exposed to sucrose, will lyse due to the build-up of levans in the periplasm[182].

After incubation, a dilution series was made for each of the inoculums, 10^{-1} to 10^{-4}. In total, 100µl of 10^{-1}, 10^{-2} and 10^{-3} dilutions were plated out onto sucrose plates, and then 100µl of 10^{-4} dilution was plated out onto an LB agar plate to be used as a control. The plates were then incubated overnight at 28°C.

The aim of proceeding step was to screen colonies for chloramphenicol and kanamycin resistance. Sensitivity to chloramphenicol and kanamycin infers the loss of the plasmid and s*acB*-Kan^R cassette, respectively (successful exchange results in cells having neither). Three 6cm x 6cm grids were made on three new plates: one LB, one containing chloramphenicol and one containing kanamycin per colony. From the sucrose plates, a colony was transferred to the same grid square on each of the three types of fresh plates. This was done until 36 colonies had been sampled from each clone. All plates were then incubated overnight at 37°C. On the fourth day, colonies that grew only on the plain plate were re-streaked to a single colony to be stocked.

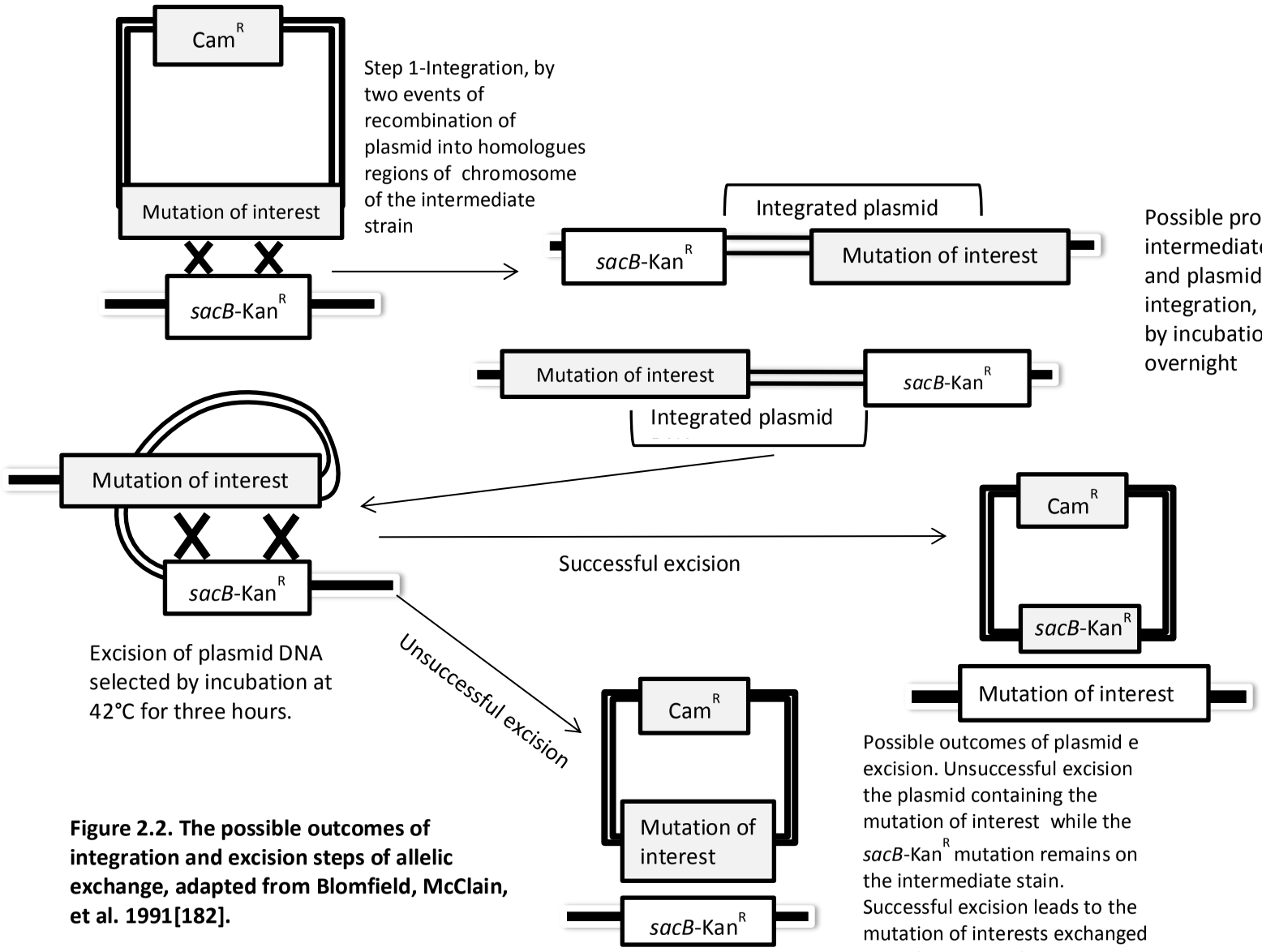


Figure 2.2. The possible outcomes of integration and excision steps of allelic exchange, adapted from Blomfield, McClain, et al. 1991[182].

2.8 β-galactosidase assay

The β-galactosidase assay allows for measuring the expression of a specific gene that is fused to the *lacZ* reporter. The product of the *lacZ* gene, β-galactosidase, will cleave ortho-Nitrophenyl-β-galactoside into o-nitrophenol and galactose. o-nitrophenol formation per unit time is proportionate to the concentration of β-galactosidase and subsequently the expression of *lacZ*. The product of β-galactosidase assays was measured in Miller units according to the published procedure (Miller, 1972)[183]. All β-galactosidase assays were performed in LB broth.

A strain containing the gene of interest, coupled to the *lacZ* gene, was grown overnight in 5ml of LB broth at 37°C in a shaking incubator. Next, 15µl of overnight culture was added to 15ml of warm LB broth. The inoculum was grown to an OD_{600nm} of 1.8 and placed on ice and left for 20 minutes. Next, the OD of the strains was recorded. At this point the OD is expected to be approximately OD_{600nm} of 0.2. Next, 250µl of cell culture was added to 250µl of Z-buffer (Table 2.9), 5µl 0.1% SDS and 10µl chloroform in a glass test tube, this was repeated for four test tubes. In a fifth test tube, a control was prepared in the same way, but by replacing cell culture with LB broth. The tubes were then vortexed for 10 seconds and then placed in a static water bath at 28°C to pre-warm for 5 minutes. The assay was started by adding 100µl 4mg/ml of ortho-Nitrophenyl-β-galactoside (ONPG), filter sterilised, to each tube at 5-second intervals.

After 90 minutes (or until the mixture took on a yellow appearance), 250µl of 1M Na_2CO_3 was added to stop the reaction. The Na_2CO_3 was added in the same order in which the assay was started. The OD of each tube was then measured at OD_{420nm} and OD_{550nm}, and negative controls were used as blanks. The following equation was used to calculate the Miller units.

Miller Units = $1000 \times (OD_{420nm} - 1.75 \times OD_{550nm}) / (OD_{600nm} \times$ Time of Assay (minutes) $\times$ 0.25(volume in ml))

The above equation measures the formation of, yellow, o-nitrophenol and thus reflects activity of β-galactosidase per unit of time. This is achieved by measuring the absorbance of the sample at 420nm, the wavelength at which yellow light is absorbed. The light scattering effect of cell debris is mitigated by measuring the absorbance at 550nm and multiplying it by 1.75, which is equivalent to the absorbance at 420nm. The OD_{600} is a measure of cell density before the start of the assay, ostensibly, before cells had been lysed by SDS. This is value is multiplied by the time of the assay in minutes and the volume in milliliters.

The Miller units for the four tubes was calculated and an average of the four replicates was then taken to calculate a mean value. The Student's T-test was used to compare the mean values obtained, resulting in a confidence interval of 95%. A significant value was one in which the mean value of the first strain was not covered by the second +/- its 95% confidence interval. A 95% confidence interval for four replicates = (3.182σ) / √4). σ was the sample standard deviation, calculated by Microsoft Excel. The number 3.182 was the α/2 critical value for three degrees of freedom.

Reagent	**Concentration**
$NaH_2PO_4.H_2O$	40mM
$Na_2HPO_4.7H_2O$	60mM
$MgSO_4.7H_2O$	1mM
KCl	10mM
β-mercaptoethanol	50mM
Final volume was made up to 500ml using ddH_20	

Table 2.9. The concentrations and reagents of Z–buffer.

2.9 DNA analysis by agarose gel electrophoresis

1% Agarose gels were made, using 0.3gs of Agarose (Bio-Rad) and 30ml of 1x TAE (Table 2.10). The mixture was then microwaved for 2 minutes or until all precipitate

from the mixture had dissolved. The liquid gel was then poured into a gel dock and left to dry. Once solidified, the samples were added to the gel, usually at a volume of between 5µl -15µl per well. A 1KB ladder (promega) was run on the gel so that the DNA fragments could be compared and their size estimated. The gel was then run at 80mA 120volts for 26 minutes, in a solution of 0.5µg/ml ethidium bromide. The gel was then imagined using a UV imaginer. To visualise the DNA fragments and ladder, Kodak 1D software and a UV transilluminator (Hoefer Scientific Mighty Bright Transilluminator UV Light Model UVTM-25) were used. If the DNA was to be used for cloning the band of interest was excised from the gel. After the DNA band had been excised, it was gel-purified, using a Qiagen DNA purification kit as per the manufacturer's instructions[176].

Reagent	**Concentrations**
EDTA 0.5M pH 8.0	1mM
2-Amino-2-(hydroxymethyl)propane-1,3-diol (Tris) (Fisher Scientific)	40mM
Acetic acid	40mM
Final volume was made up to 1L using ddH20	

Table 2.10. Reagents used to make the TAE buffer.

2.10 Whole-cell lysate preparation

Cells were grown overnight in 5ml of LB broth. Then, 100µl of this overnight culture was added to 20ml of LB broth and left to grow to OD_{600nm} of 0.3. The inoculum was then sonicated at 40Hz for a total of 6 minutes, at intervals of 30 seconds. The lysed cells were then subjected to a Markwell assay and diluted down, using ddH_20, to ensure all the samples (if multiple strains were used) had the same total protein count. The lysate was then stored at -20°C until it they were used.

2.10.1 Markwell Assay

This procedure, an assay to determine protein concentration, is based on published work by Markwell,1978[184]. Reagents A (4.9ml) and B (100µl) were mixed in a 49:1 ratio to form Reagent C (Table 2.11), 1ml of which was placed into Eppendorf tubes with 0.25ml of a whole-cell lysate – three Eppendorf tubes were used for each sample and incubated at room temperature for 60 minutes. Next, 75µl of diluted Folin's Solution (mixed 1:1 with dH2O) was added to each tube and vortexed. The tubes were then left to incubate for 45 minutes at room temperature and then vortexed. The OD of each tube, using a spectrophotometer, was taken at 660nm. The average of the three tubes for each sample was taken.

A duplicate of this experiment was run in parallel, using bovine serum albumin (BSA) at different concentrations to form a standard. The data were analysed by plotting a standard curve using the OD of the BSA samples (µg/ml against absorbance OD_{660}). The average OD of the experimental sample was compared to the BSA standard curve. This yielded the sample protein concentration.

Reagents	**Constituents**
Reagent A	20g/l Na_2CO_3, 4g/l NaOH, 1.6g/l Sodium tartrate, 10g/l SDS
Reagent B	4% $CuSO_4.5H_2O$
Reagent C	Mix reagents A & B in a 99:1 ratio

Table 2.11. Describing agents used for a Markwell assay.

2.10.2 Western Blot Procedure

An Invitrogen Life Technologies Western Breeze kit was used to carry out all the Western blots (WesternBreeze Chromogenic Western Blot Immunodetection Kit catalogue no. WB7103)[185]. Blockers/diluents were used throughout the procedure (Table 2.12 and Table 2.13), which was carried out according to the manufacturer's instructions[185].

Initially a gel electrophoresis was performed using an electrophoresis cell (Bio-Rad) on the whole cell lysate (2.10), using a (NuPAGE™ 4-12% Bis-Tris Protein Gels ThermoFisher).

To achieve this, 20µl of whole cell lysate was loaded into the precast gel. The gel was then run at 150V for 50 minutes using NuPAGE MES SDS Running Buffer (ThermoFisher).

After electrophoresis, proteins were transferred onto a nitrocellulose membrane using the iBlot™ system (ThermoFisher) in accordance with manufacture's instructions[186]**.** The iBlot™ system contains an anode stack, cathode stack, blotting filter paper, disposable sponge and Gel Transfer Device. This method was chosen because it was comparatively simpler and easier than other methods, such as semi-dry blotting.

The anode stack was placed on the gel transfer device. Subsequently, the gel was removed from the cassette and placed on top of the anode stack. Next, blotting filter paper was soaked in ddH_20 for 2 minutes and placed on top of the gel, ensuring there were no bubbles present. Then, the cathode stack was placed on top of the filter paper, with the copper electrode aligned to the right of the bottom of the stack. Finally, the the disposable sponge was placed on the inner side of the lid. The gel transfer device was run for 6 minutes at 23V. The nitrocellulose membrane was then removed from the stack.

The membrane was then placed into 10ml of blocking solution in a covered plastic dish and incubated for 30 minutes on a rotary disk set at 1 revolution/sec. The

blocking solution was then decanted. The membrane was rinsed with 20ml of ddH_20 for 5 minutes and then decanted. This washing step was then repeated. The membrane was then incubated with 10ml of primary antibody solution on a rotary disk set at 1 revolution/sec overnight at 4°C. The antibody solution was then decanted and the membrane was washed with 20ml of antibody wash for 5 minutes. This washing step was repeated three times. The membrane was then incubated in 10ml of secondary antibody solution for 60 minutes, following which the secondary antibody was decanted. The membrane was then washed for 5 minutes with 20ml of antibody wash, which was then decanted. This washing step was repeated twice. The membrane was then rinsed with 20ml of ddH_20 for 2 minutes. This step was repeated twice. The membrane was then incubated with 5ml of chromogenic substrate on a rotary disk set at 1 revolution/sec. When purple bands started to develop on the membrane, the chromogenic solution was decanted. The membrane was then rinsed with 20ml of water for 2 minutes, the rinse step was then repeated twice more. The membrane was then left to air dry.

	Ultra-filtered Water	Blocker/Diluent (Part A)	Blocker/Diluent (Part B)	Antibody wash
Blocking Solution	5ml	2ml	3ml	0ml
Primary Antibody	7ml	2ml	1ml	0ml
Antibody Wash	150ml	0ml	0ml	10ml

Table 2.12. Describing the concentrations of blocking solutions used for Western blotting.

Reagent	Constituent
Blocker/Diluent (Part A)	Concentrated buffered saline solution containing detergent
B locker/Diluent (Part B)	Concentrated Hammersten casein solution
Antibody Wash (16X)	Concentrated buffered saline solution containing detergent
Chromogenic Substrate	Ready-to-use solution of BCIP/NBT substrate for alkaline phosphatase
Primary Antibody	Anti-sigmaE (from the Gross lab UCSF)
Secondary Antibody (anti-rabbit)	Ready-to-use solution of alkaline phosphataseconjugated, affinity purified, anti-species IgG

Table 2.13. Describing the reagents and blocking solutions used for Western blotting.

2.11 Automated growth curve analysis by plate reader

Strains were grown overnight in 5ml of LB broth, and then 1µl of the overnight culture was added to 1ml of LB broth pre-warmed to 37°C. During this step any additions to media was made. In this study the growth rate of strains with and without citrate was assayed by adding various concentrations of citrate to the media. Next, 200µl of the inoculum was placed in the well of a BioTek Microplate and then repeated five times, generating replicates for each sample. A blank, of LB broth, was inoculated into at least five wells, to be used as a control and zero state for the plating reader (FluoDia T70). The microplate was then placed in a BioTek

microplate reader and left for up 24 hours, following which the results were recorded.

Chapter 3: The *phoP* transduction defect

3.1 Introduction

PhoPQ plays an important and varied role in the cell, as it regulates a large number of genes. It has been found that over 232 genes, directly or indirectly, respond to PhoPQ[42], but despite the expansive PhoPQ regulon, it has not been considered essential, and *phoP* mutants are made routinely by a variety of methods[112],[96].

Generalised transduction is a useful tool for genetic manipulation, since it allows the movement of genetic elements or chromosomal markers from one bacterium to another. Using a bacteriophage as an intermediary, genetic material is moved from a donor cell to a recipient cell and then exchanged into the chromosome of the recipient via homologous recombination. Although transduction is frequently carried out in nature, in laboratories transduction is usually carried out using a specific, modified, bacteriophage vector, such as P1 in *E. coli*.

Transduction requires control over the infectious nature of the bacteriophage, in order to function efficiently. Without some way of restricting the infectivity of the bacteriophage, virulent bacteriophage will infect and lyse most recipient cells. The method employed to control it involves manipulating available Ca^{2+} concentration[154]. Bacteriophage P1 requires calcium ions to infect a bacterial cell, in order to counteract the negative charge on both the bacterial membrane and the viral capsid.

Without calcium ions, the infectivity rate is lowered due to increased electrostatic repulsion between the bacteriophage and the recipient cell[154]. Once the bacteriophage has had sufficient time to transfer the donor DNA into the recipient bacterium, a metal ion chelator is added (often citrate or Ethylenediaminetetraacetic acid), which lowers the concentration of free calcium sufficiently to effectively prevent further attachment of P1, though not so low as to deprive the cells and induce divalent metal ion starvation.

Surprisingly, previous work carried out in the Blomfield Laboratory found it was not possible to isolate *phoP* transductants under standard conditions (Dr Alex Moore, personal communication).

In agreement with previous work in the Blomfield Laboratory, during initial experimentation it was not possible to make a *phoP* mutant by standard P1*vir* transduction under standard conditions. This transduction incapacity was found to be independent of the wildtype recipient strain or *phoP* mutant donor strain used.

The aim of the work outlined in this chapter was to confirm the existence of a transduction defect associated with the loss of *phoP* during P1*vir* transduction.

3.2 The *phoP* transduction defect

Several attempts were made to transduce *phoP* (JW1116-1) and *phoQ* (JW1115-1) mutations into the W3110 wildtype; it was found that *phoP* could not be transduced, but the opposite was true for *phoQ* (Table 3.1).

Lysate	Repeat 1	Repeat 2	Repeat 3	Average
phoP 10µl	0,0	0,0	0,0	0
phoP 50µl	0,0	0,0	0,0	0
phoQ 10µl	2,4	8,6	3,7	5
phoQ 50µl	1,3	4,2	2,3	2.5

Table 3.1. The number of *phoP* (JW1116-1) and *phoQ* (JW1115-1) transductants generated using P1*vir* transduction in the wildtype W3310. Replicates for each repeat and lysate concentration are displayed in pairs. The average number of transductants is shown for *phoP* and *phoQ* lysates concentrations includes all repeats and replicates.

MG1655 and W3110 are two prototypical wildtypes of *E. coli* K12, and both are exploited widely in molecular biology. They have a high sequence similarity; however, there are genetic differences between the two strains that manifest in the form of insertions, deletions and inversions[187]. Additionally, K12 isolates are known to vary genetically between laboratories [188]. It seemed possible that these genetic differences may affect the strains ability to act as a recipient for *phoP* during P1*vir* transduction. To test the hypothesis that *phoP* transduction defect was

not limited to W3110, the P1*vir* transduction experiment was repeated using the alternative wild type strain MG1655 and the lysate generated from strain JW1116. In agreement with the observation made in 3.2 *phoP* could not be transduced (data not shown).

It was hypothesised that the transduction defect maybe due to the donor strain used. The *phoP* (JW1116-1) lysate utilised in previous experimentation had been generated from a donor strain from the Kieo collection. The Kieo collection distributes single-gene deletion mutants, which have deletions from the second to the seventh-to-last codon. Open reading frames are replaced with a kanamycin resistance cassette flanked by FLP recognition target sites. To confirm that the results demonstrated above were not anomalies caused by the donor strain, or their method of construction, an alternative *phoP* mutant (*phoP*::Tn, provided by the Groisman Laboratory[74] was used (Table 3.2). The same defect was noted using the alternative *phoP* donor strain. The results suggest that the defect initially observed with the *phoP* Keio mutation in W3110 is not specific to the type of *phoP* mutation used.

Lysate	Repeat 1	Repeat 2	Repeat 3	Average
phoP 10µl	0,0	0,0	0,0	0
phoP 50µl	0,0	0,0	0,0	0
phoQ 10µl	9,7	6,4	8,11	7.5
phoQ 50µl	28,31	18,19	8,10	19

Table 3.2. Number of *phoP* (Eg12976) and *phoQ* (JW1115-1) transductants generated using P1*vir* transduction in the wildtype MG1655. Replicates for each repeat and lysate concentration are displayed in pairs. The average shown for *phoP* and *phoQ* lysates concentrations includes all repeats and replicates.

3.2.1 Investigating the *phoP* transduction defect

To further test the validity of this observation the P1*vir* transduction assay, using MG1655 and *phoP* (JW1116-1) and *phoQ* ((JW1115-1), was repeated 25 times. Different cultures of donor strain were used for each repeat and three different preparations of *phoP* and *phoQ* lysates were used throughout the 25 experiments.

25 transductions experiments were done using both 10µl and 50µl lysate for both *phoP* and *phoQ*. Two replicates were done for each concentration of lysate (full data in Appendix Figures B.1A&B.1B).

Lysate	Mean number of transductants	Standard deviation
phoP 10µl	0	-
phoP 50µl	0	-
phoQ 10µl	2	11.43
phoQ 50µl	22.50	2.44

Table 3.3. The average number of *phoP* and *phoQ* transductants isolated in the wildtype (MG1655) over 25 assays, using both 10µl and 50µl of P1 lysate. Full data in Appendix Figures B.1A&B.1B

The data show that it was not possible to create any *phoP* transductants using this method (Table 3.3). However, it was possible to transduce a *phoQ* mutation (donor strain JW1115-1) into the wildtype background, which may suggest a transduction defect associated with *phoP*, albeit not with PhoP's cognate sensor kinase, PhoQ. This was a surprising outcome, as it would be expected that disrupting either component of a TCS, either the response regulator or sensor kinase, would have a similar phenotypic effect.

To test the significance of the results generated above (Table 3.3), a student's t-test was used to compare the means of average number of *phoP* and *phoQ* transductants generated using 50µl of lysate. The null hypothesis tested was that there was no significant difference between the means of the two datasets. The test was performed using the Microsoft Excel data analysis software package. The alpha level chosen was 0.05 (Table 3.4). The analysis shows a small P-value, <0.001, smaller than the alpha level and a large divergence between the t-value and the t

critical value. This suggests that the difference between the means is highly significant, and so the null hypothesis can be strongly rejected.

	Average *phoP* 50µl	Average *phoQ* 50µl
Mean	0	22.54
Variance	0	46.75
Observations	25	25
Hypothesised mean difference	0	
Degrees of freedom	24	
t value	-16.48	
P-value	<0.001	
t critical value	2.06	

Table 3.4. t-Test: Two-Sample assuming unequal variances performed on transduction results from Appendix Figures B.1A&B.1B.

The difference in transducability between these two genes is unexpected but not entirely surprising. It has been documented that some response regulators maintain a certain amount of residual activity in the absence of their cognate sensor kinase[189]. For example, in *S. enterica* it has been shown that PhoP can bind to the *mgtA* promoter in its unphoshorylated state. Although there is a 14% dissimilarity between the *phoP* homologous in *S. enterica* and *E. coli*, it is probable that PhoP functions in a similar way in both organisms. Alternatively, although not documented *in vivo*, it is possible that PhoP may be phosphorylated by a different sensor kinase. Cross-talk between two-component systems is generally considered to be low. However, there are several examples of phosphorylation of response regulators by non-cognate histidine kinases. The response regulators CheY, CusR and YfhA have been shown to be phosphorylated by two, three and four different histidine kinases respectively[190].

Alternatively, it is possible that the *phoP* lysate, used for these assays, had a low concentration of transducing particles, albeit this is unlikely, as three different preparations of *phoP* lysate were used. Nonetheless, because the viral titre was not measured, it remains a possibility. This is further discussed in the final section of this work.

3.3 Complementation of the *phoP* transduction defect

While the literature suggests that it is possible – and common practice – to make *phoP* mutations via P1*vir* transduction[102][112][96], the primary observation of this work disagrees with this notion, thereby suggesting two hypotheses:

I) *phoP* is indispensable during transduction and *phoP* mutants created via transduction have additional suppressor mutations that make *phoP* dispensable.

II) *phoP* is indispensable during transduction.

The transduction defect noted above may in fact be due to the essentiality of the gene. In this case, it must be presumed that existing *phoP* mutants made by others contain additional suppressor mutations that compensate for the loss of the gene. Alternatively, *phoP* may be essential under the conditions of transduction. To help distinguish between these possibilities, a *phoP* mutant was constructed containing an ectopic copy of *phoP* controlled by the *lac*UV5 promoter.

The *lac*UV5 promoter is a modified version of the *lac* promoter. The *lac* operon is vital for the metabolism of lactose in *E. coli*[191]. and is regulated by a dedicated repressor, lacI, and also CRP[174]. The repressor binds to an operator, but this binding is inhibited in the presence of lactose, which allows for the expression of the operon[191]. The presence of CRP-cAMP positively regulates the *lac* operon [191], which in turn is negatively affected by glucose.

The *lac*UV5 has several modifications, such as the inactivation of CRP-cAMP regulation, meaning that strains containing the *lac*UV5 promoter can be grown in rich media[191]. Furthermore, the *lac*UV5 promoter recruits RNA polymerase[191]

more efficiently and is induced by either lactose or lactose-analogue Isopropyl β-D-1-thiogalactopyranoside (IPTG) [191], which is used because, unlike lactose, it cannot be degraded by the cell[191]

It was anticipated that this strain would prove to be a viable recipient for the transduction of *phoP* mutation under inducing (IPTG-present) but not under non-inducing (IPTG-absent) conditions.

Construction involved constructing a plasmid, which contained *phoP* under the control of the *lac*UV5-promoter, and an intermediate strain *gshA,phoP* (KCEC5275). Previous work in the Blomfield laboratory found it was possible to isolate *phoP* transductants in *gshA* mutants (Dr Alex Moores, personal communication) (this is discussed in further detail in chapter 4). Details of the construction of the intermediate strain and plasmid are given below in 3.3.1.

3.3.1 Construction of the *lac*UV5-*phoP* strain

Step 1: Construction of pA001, a temperature sensitive vector containing *phoP* under control of the *lac*UV5 promoter

The plasmid pIB462 was used as a starting vector[163]. This temperature-sensitive plasmid (Figure 3.2) contains a multiple cloning site downstream from the *lac*UV5-promoter as well as regions of homology to the *lac* operon that flank multiple cloning sites, which is necessary for recombination later in the procedure. The pIB462 vector was digested with the restriction endonucleases BamHI and HindIII.

The *phoP* gene was amplified from the wildtype (MG1655), but it was important not to amplify *phoP*'s native promoters that exist -60bp upstream from the gene start[102]. Primers were chosen that amplified as close as possible to the ATG start of the open reading frame while including the ShineDalgarno Sequence (Figure 3.1). Furthermore, primers were designed so that the amplified gene fragment would contain HindIII and BamHi recognition sites. The *phoP* gene fragment was digested with restriction enzymes, HindIII and BamHI to create complementary overhangs with the digested vector. *phoP* was cloned into the multiple cloning site, using the ligation procedure, downstream of the *lac*UV5-*phoP* promoter. The resulting

plasmid was designated pA001 (Figure 3.3 step 1). pA001 was digested with BamHI and HindIII with the aim of isolating and a gene fragment of a size that

corresponded to the original *phoP* gene fragment cloned into the pIB462. This was successful (data not shown) and provided confidence that plasmid, pA001, contained the *phoP* gene fragment.

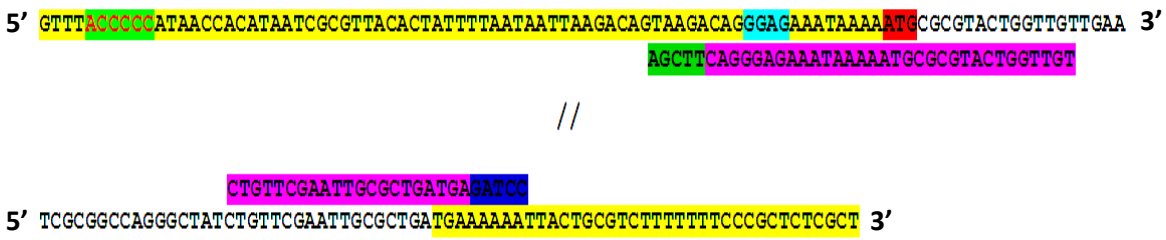


Figure 3.1. Showing the position of the primers (Vector_*phoP*- forward and Insert_*phoP*2-reverse) used to amplify the *phoP* gene on the 5'-3' coding strand. Primers were chosen that were downstream of the *phoP* promoter region. Yellow indicates the up and down stream of the gene (-61- +721). Green indicates the transcription start site. Light blue indicates the Shine-Dalgarn Sequence. Red indicates the start codon. The purple regions indicate the primers used. Blue and green indicate where BamHI and HindII recognition sequences were placed on the primer, in order to engineer these sites into the amplicon.

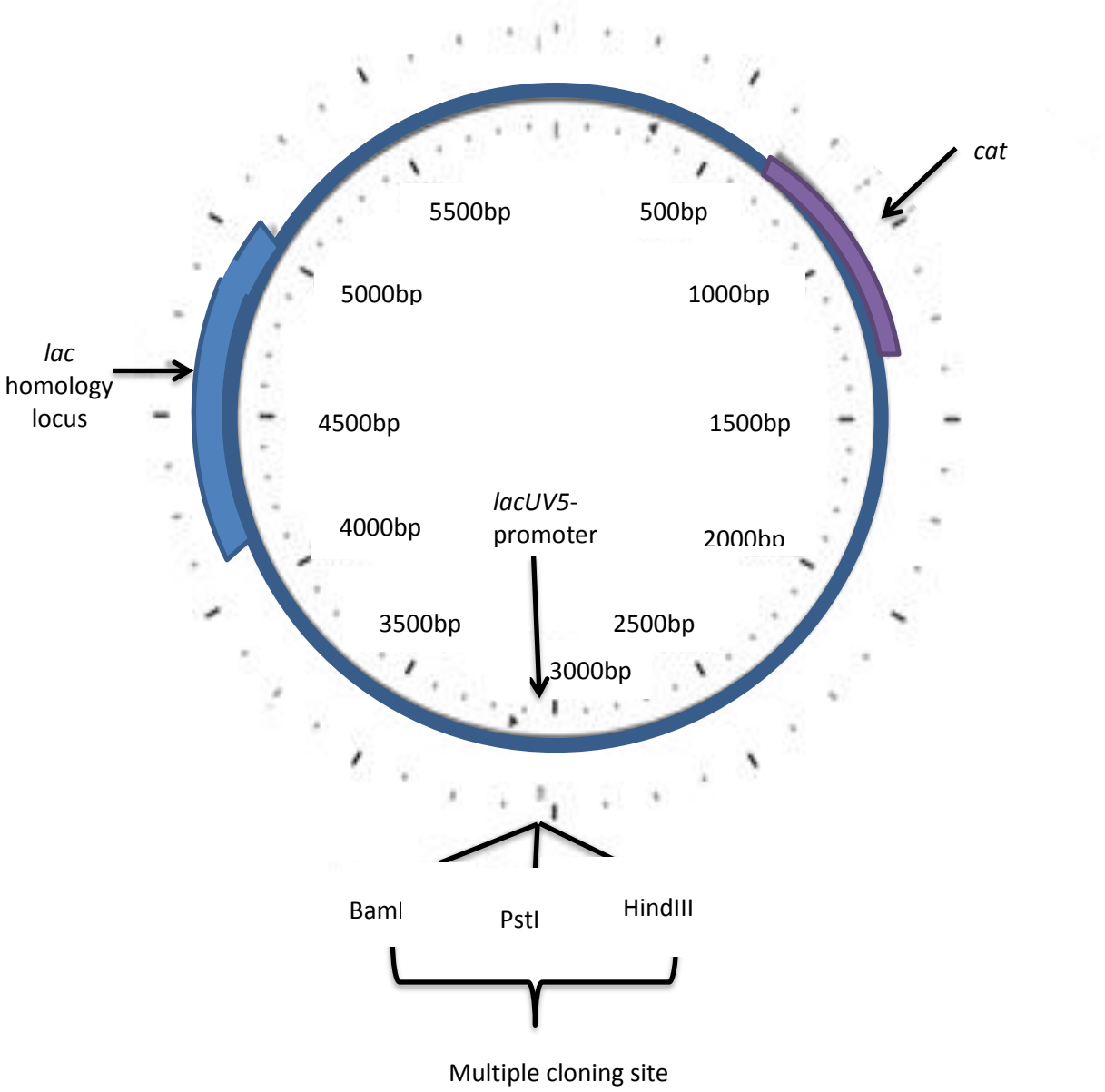


Figure 3.2. Showing the plasmid pIB462 5989 Base pairs in length (Blomfield et al. 1991[182]). The plasmid contains, Chloramphenicol acetyltransferase (Cat) (purple),multiple cloning site (BamH1, PstII, HindIII) and *lac* homology regions downstream of the multiple cloning site (blue).

Step 2: Construction of the intermediate strain

The strain *gshA,phoP* (KCEC5275) was created by transduction (see Chapter 4). A P1 bacteriophage lysate containing the *sacB*-Kan^R cassette was prepared from the strain AAEC090. The *sacB*-Kan^R cassette is composed of the *sacB* gene from *Bacillus subtilis* and a Kanamycin resistance cassette. SacB produces levansucrase, an enzyme that is necessary for the synthesis of a polymer of sucrose called levans[182]. *sacB* expression is lethal in the presence of sucrose in *E. coli* and is commonly used as a counter-selectable marker [182]. The lysate from strain AAEC090 (Δ*lac*ZYA Ω*sacB*-Kan^R) was used to transduce the *gshA,phoP* strain, placing the *sacB*-Kan^R cassette at *lac* and thereby yielding the strain *gshA,phoP,*ΔlacUV5Ω*sacB*-Kan^R (KCEC5281 H) (Figure 3.3 step 2).

Step 3: Allelic exchange between *gshA,phoP*Δ*lac*Ω*sacB*-Kan^R and pA001

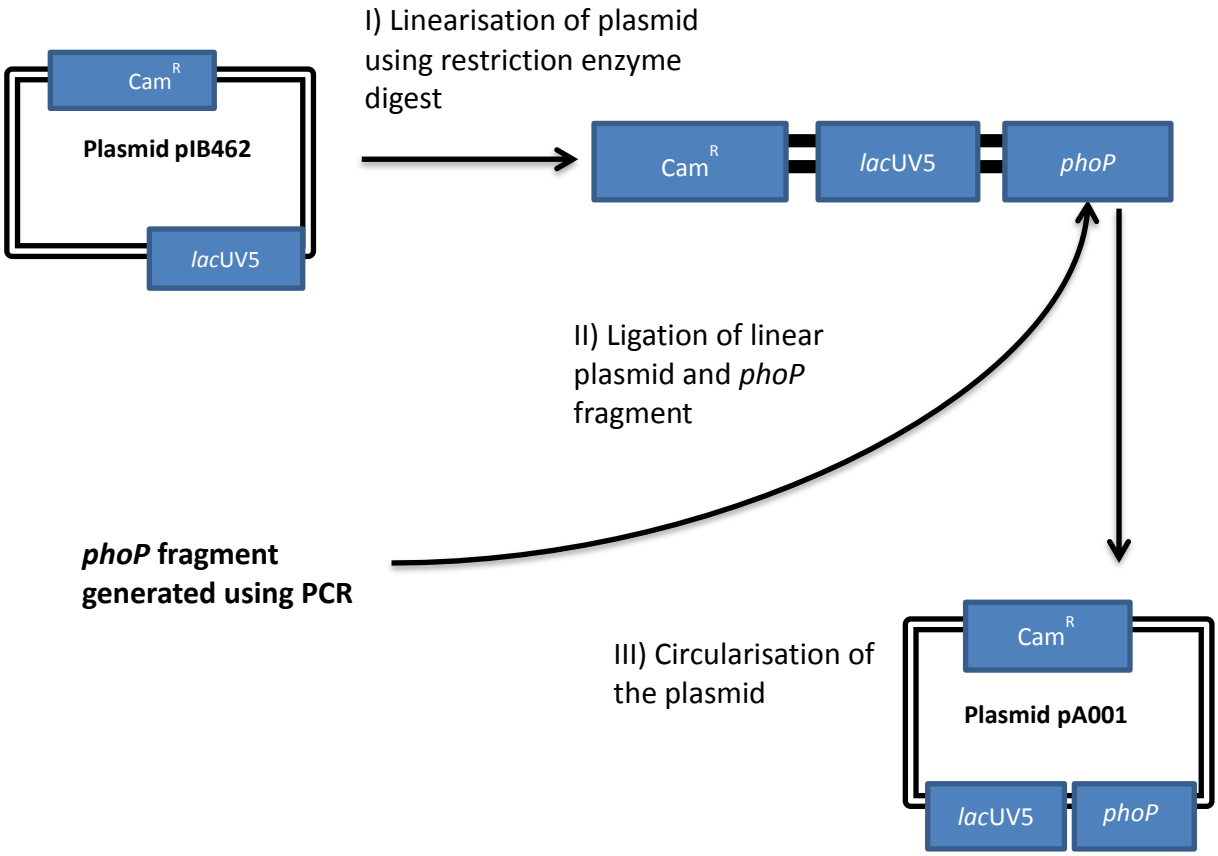
The strain *gshA,phoP*Δl*ac*Ω*sacB*-Kan^R (KCEC5281 H) and the plasmid pA001 were then subjected to allelic exchange. This allowed for the transfer of *lac*UV5-*phoP* from the plasmid into the strain *gshA,phoP*Δl*ac*Ω*sacB*-Kan^R at the *lac* locus, thus replacing the *sacB*-Kan^R cassette. After the allelic exchange, sucrose was used as a counter-selective agent.

Strains that were able to tolerate growing on LB-sucrose agar and unable to survive exposure to kanamycin were assumed to lack the *sacB*-Kan^R cassette. These strains had undergone allelic exchange, and purportedly they contained the *lac*UV5-*phoP* at *lac*. Allelic exchange yielded the strain *gshA,phoP*Δ*lac*Ω*lac*UV5-*phoP* (KCEC5278)(Figure 3.3 step 3). Each strain was assayed for the absence of both wildtype *phoP* and *gshA* by PCR (data not shown). The construct was analysed with three sets of primers, which were used to confirm the absence of *gshA*. External primers were used to confirm the absence of the native *phoP* gene, from 30bp upstream of the ATG start codon (*phoP*-forward and *phoP*-reverse)(Data not shown), while internal primers were used to amplify the ectopic copy of *phoP* and confirm its presence in the strain (Vector_*phoP*- forward and Insert_*phoP*2- reverse)(Appendix Figure A.1A).

Step 4: Transduction of *lac*UV5-phoP into a strain lacking additional mutations

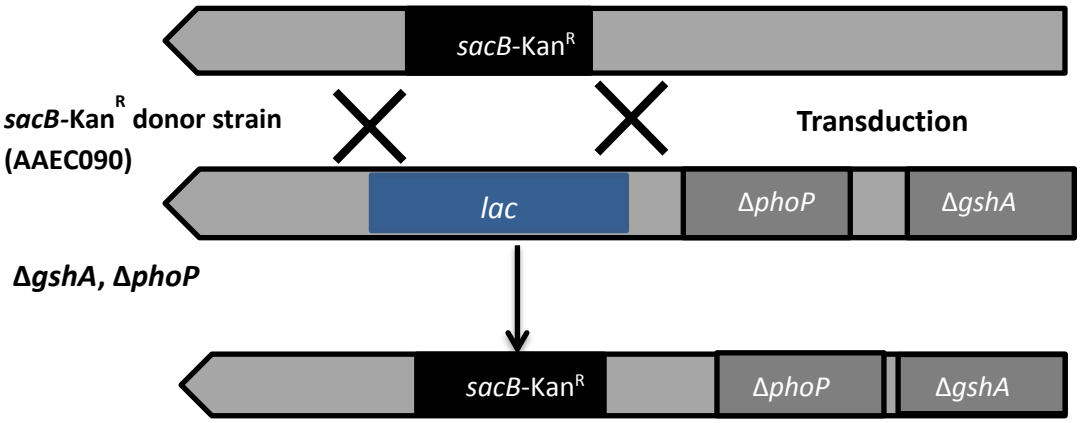
It was necessary to move *lac*UV5-*phoP* into a strain without the *gshA,phoP* mutations. The aim of this step was to backcross *lac*UV5-*phoP* into AAEC090, which has a *sacB*-Kan^R cassette at the *lac* locus and contains the wildtype *phoP* allele(Figure 3.3 step 4).

The strain that was the product of the allelic exchange(Δ*gshA*, Δ*phoP* Δ*lac*Ω*lac*UV5-*phoP*) was used as the donor strain for a P1 lysate, which was then used to transduce the strain AAEC090, which has a *sacB*-Kan^R cassette at the *lac* locus (Figure 3.3 step 4). All transductants that could grow in the presence of sucrose and but were unable to tolerate growth in presences of kanamycin were analysed further by PCR, to confirm their correct chromosomal structure. Strains that contained the wildtype *gshA* and *phoP* genes (data not shown) were successful constructs. External primers were used to differentiate between the wildtype *phoP* gene and the *phoP* associated with the *lac*UV5 promoter .



Step 1:Creation of pA001 plasmid

Plasmid pIB462 was modified via cloning, placing *phoP* under the control of the lacUV5-promoter and yielding the plasmid pA001.
Plasmid pIB462 was cut with BamHI and HindIII endonucleases. The linear plasmid was then ligated with *phoP*, which had also been cut with BamHI and HindIII to create complementary regions.

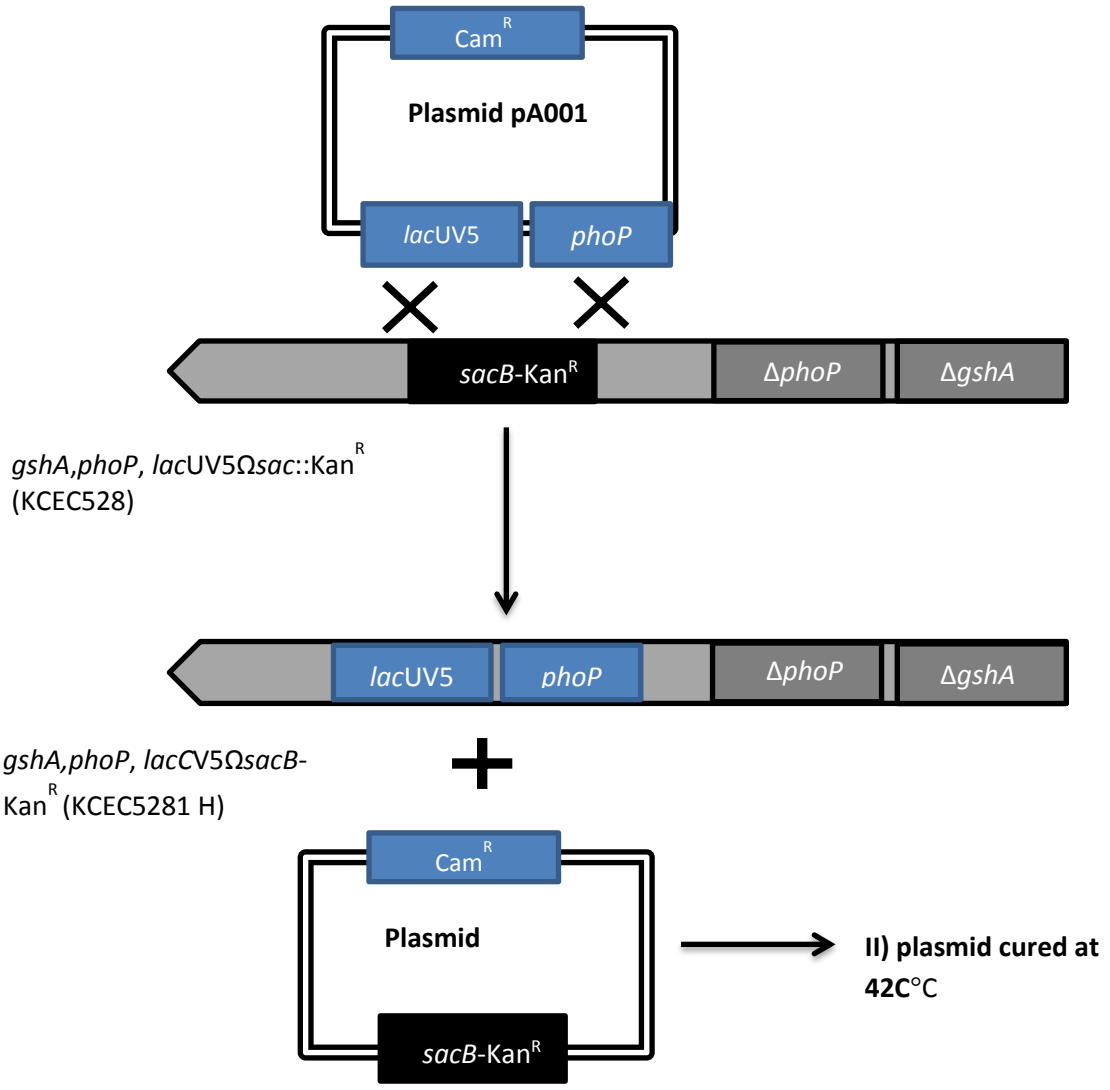


gshA,phoP,
lacUV5ΩsacB -Kan^R
(KCEC528)

Step 2 (Transduction I):Creation of the intermediary strain Δ*gshA*, Δ*phoP* Δ*lacuV5ΩsacB*-Kan^R

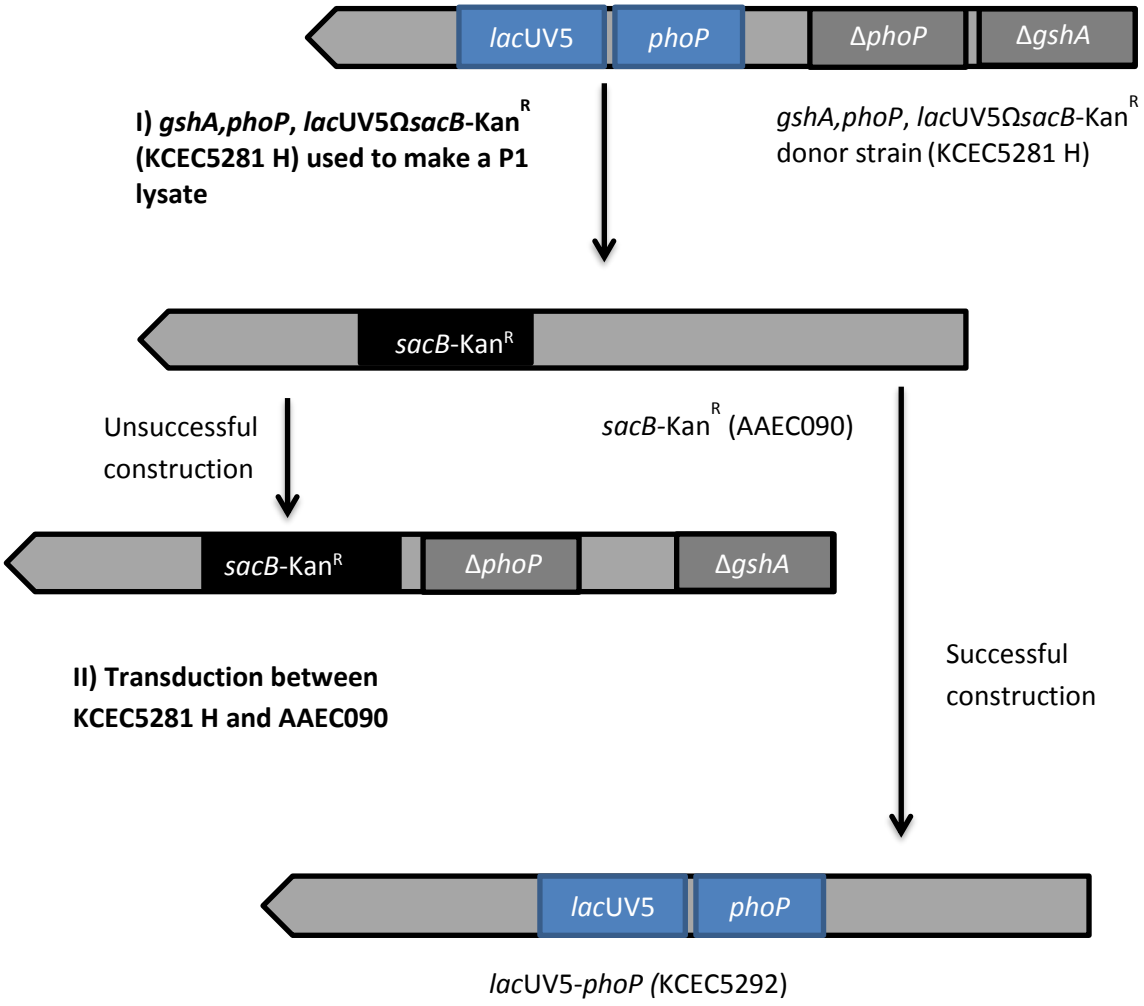
A lysate containing *sacB*-Kan^R was transduced into the strain KCEC5275 Δ*gshA*, Δ*phoP*. This yielded strain Δ*gshA*, Δ*phoP* Δ*lacuV5ΩsacB*-Kan^R (KCEC5281 H). It was necessary to use a Δ*gshA*, Δ*phoP* strain, because *gshA* suppresses the *phoP* transduction defect (see Chapter 4).

The plasmid containing *phoP* under the control of the *lac*UV5 promoter (pA001) was integrated into the chromosome of the strain *gshA*,*phoP*, *lac*UV5Ω*sac*::Kan^{R} (KCEC528). The plasmid was then excised, leaving the *lac*UV5-*phoP* in the chromosome.



<u>Step 3:Allelic exchange between Δ*gshA*, Δ*phoP* Δ*lacuV5*Ω*sacB*-Kan^{R} and pA001</u>

Allelic exchange was then used with the strain Δ*gshA*, Δ*phoP* Δ*lac*U*V5*Ω*sacB*-Kan^{R} and plasmid PA001 lacUV5-p*hoP* produced the strain *ΔgshA, ΔphoP lac*UV5-*phoP.*



Step 4:Transduction 2 creating the strain *lac*UV5-*phoP*

The goal of this step was to place *lac*UV5-*phoP* at *lac*. Using the strain *ΔgshA, ΔphoP lac*UV5-*phoP,* a lysate was made, which was then used to transduce AAEC090, a strain containing the *sacB*-Kan^R selectable marker at the *lac* locus. After the transduction was complete, all transductants were tested using sucrose as a counter-selectable marker. Strains that could not tolerate growth on sucrose were analysed by PCR. Strains that had both *gshA* and the wildtype *phoP* were deemed correct.

Figure 3.3. Step 1-4: schematic outlining the four primary steps required to create the strain *lac*UV5-*phoP* (KCEC5277).

3.3.2 Testing the complementation mutant *lac*UV5-*phoP*

For many genes, essentiality is transient; they are dispensable under some conditions but necessary for survival in some environments. Well studied examples are genes of the *srl, xyl* and *ara* operons. While not considered essential, these genes become indispensable when glucose, *E. coli's* preferred carbon source, is absent and the cell must utilise alternative carbon sources, such as sorbitol, xylose and arabinose, respectively[192].

Many of these transiently essential genes are linked to stress responses and are regulated by alternative sigma factors or transcription factors, placed in the centre of regulatory networks; they govern or work in conjunction with many other regulatory proteins. There is mounting evidence that PhoPQ is centred in the middle of a large and interconnected regulatory network, such as global regulators Crp, Fnr, and ArcAB[96], [190].

PhoP is involved in the regulation of a number of aspects of cell metabolism and physiology, aside from responding to Mg^{2+} starvation. PhoP enhances sigmaS stability by acting as a transcriptional activator of *iraP,* which interferes with RssB mediated degradation by ClpXP[193]. Additionally, PhoP plays a role in LPS modification by acting as a positive regulator of *crcA* and post-transcriptionally inhibiting expression of *eptB*, via MgrB[76]. Both EptB and CrcA influence the charge disruption of the LPS by acetylation of lipid A or by catalysing the addition of palmitate and adding a Kdo group to the inner core, respectively[98], [194].

It is reasonable to assume that, under the conditions experienced during transduction, the viability of *phoP* mutants would be adversely affected. The cell would experience a relative decrease in sigmaS stability. Moreover, there would be a decrease in OM stability due to increased repulsion between LPS molecules. Negative charges that are present on the LPS, phosphate groups associated with the glucosamine residues of lipid A and the carboxyl groups associated with the Kdo groups of the inner core, are essential for LPS crosslinking and, in the presence of divalent cations, function to increase the stability of the OM[59]. However, these

charges are less likely to be neutralised due to chelation and subsequent decrease in availability of divalent cations. This affect will be exacerbated by EptB which will function to increase the negative charge of the LPS in the absence of *phoP*.

The aim of the experiment discussed below was to test the hypothesis that *phoP* is indispensable during transduction. This was achieved by testing the ability of the strain containing the ectopic *phoP* allele, under the control of the *lac*UV5 promoter, to act as a recipient for the *phoP* mutation, under inducing and non-inducing conditions (IPTG+/-).

The *lac*UV5-*phoP* strain was transduced using the *phoP*:Kan^R lysate (JW1116-1), +/- IPTG (0.1mM). 50µl of lysate was used and a total of two replicates were done for each of the three repeats. +IPTG transductants were plated out on LB-agar containing IPTG. It was found that transductants could only be isolated in the presence of IPTG when it would be expected that *phoP* was expressed from the ectopic copy of the gene under control of the *lac*UV5 promoter (Table 3.5).

These key data shows that PhoP is necessary for successful P1*vir* transduction but is not essential for the cell's viability in the strain AAEC090, a MG166 derivative, and presumably other wildtype backgrounds. Moreover, the data suggests that the transduction defect is unlikely to be associated with the transfer or recombination of donor DNA in the recipient because the cells were not exposed to IPTG until after this stage had been completed. This is unsurprising as there is little reason why *phoP* mutants would be unreceptive to the uptake of donor DNA. It is likely *phoP* mutants are hyper-sensitive to the stresses associated with transduction and have a diminished ability to recover. This is likely caused by a reduced ability to respond to Mg^{2+} starvation and decreased activity of sigmaS.

	Number of *phoP* transductants +IPTG	Average	Number of *phoP* transductants –IPTG	Average
Repeat 1	9,13	11	0,0	0
Repeat 2	8,6	7	0,0	0
Repeat 3	10,6	8	0,0	0

Table 3.5. The number of successful *phoP*(JW1116-1 50µl) transductants using the strain *lac*UV5, *phoP* as the recipient strain(KCEC5292) with and without IPTG, and the average of two replicates (shown in pairs) over three assays.

3.3.3 Construction of a *phoP* mutant via lambda red recombination

Two alternative hypotheses can be proposed to explain the *phoP* mutation transduction defect. First, *phoP* may be an essential gene, and its inactivation could make the cell inviable under all conditions. Alternatively, transduction may expose the cell to conditions it cannot tolerate in the absence of *phoP*. The latter hypothesis is supported by the findings from the complementation experiment in the section above (3.2.2), which show that *phoP* mutants remained viable after transduction and *phoP* was, presumably, not being expressed from the *lac*UV5-promoter. Creating *phoP* via another method corroborates the hypothesis that *phoP* is only indispensable during transduction under standard conditions. In an attempt to support this hypothesis, an alternative strain construction method, lambda red recombination, was used.

The chloramphenicol resistance cassette was cloned from plasmid pKD3 (provided by the Shepherd Laboratory) using PCR. Primers were used with flanking regions homologous to, 70bp and 50pb up- and downstream of, the wildtype (MG1655) *phoP*, respectively. This was done to ensure that sufficient homology existed to promote recombination between the chloramphenicol resistance cassette and the locus into which it was to be inserted. The recipient strain (AAEC090) was made competent and transformed with the plasmid pKD46 (provided by the Shepherd Laboratory). pKD46 carries the λ red genes *gam*, *exo* and *beta*, behind the *araBAD* promoter. The transformants were then transformed with the chloramphenicol

DNA hybrid fragment, using electroporation. The resulting transformants were then grown at 37°C on LB-agar plates containing chloramphenicol. Successful deletion of the wildtype *phoP* gene (KCEC4032) was confirmed using PCR analysis (data not shown), which was also used to confirm the presence of the chloramphenicol resistance cassette(Appendix Figure A.2). This experiment confirmed that while it was not possible to make a *phoP* mutant using P1*vir* transduction, this defect could be overcome using lambda red recombineering.

3.4 Discussion

The transduction defect reported herein is well established by the data, demonstrating that *phoP* mutants could not be successfully isolated using P1*vir* transduction in wildtype *E. coli* K-12 (MG1655 or W3110), but *phoQ* mutants could be.

The difference in the numbers of *phoQ* transductants created using 10µl and 50µl (Table 3.1, Table 3.2 and Table 3.3) may suggest that *phoQ*, and presumably *phoP*, mutants are hypersensitive to P1 killing, as it might be expected that the number of transductants created using to 10µl will correlate directly with the numbers created using 50µl of lysate. However, it is more likely that this effect is a result of the optimum volume of bacteriophage lysate for the transduction of *phoQ* being closer to 10µl than to 50µl. The number of transductants generated per microlitre of bacteriophage lysate will have a bell curve-like distribution, around the optimum concentration of bacteriophage lysate[195]. The number of transductants generated by P1*vir* transduction is a function of the concentration of transducing particles in the lysate and the efficiency of recombining DNA in the host cell, which is affected by the locus and differs for different genes[196]. However, without measurements of the *phoQ* lysate's viral titers, it is difficult to hypothesise as to why *phoQ* optimum lysate concentration lies closer to 10µl than to 50µl. It is possible that the lysates used herein were extremely concentrated, containing a high titer of virulent and transducing bacteriophage, in which case a larger volume would have caused increased cell lysis. Alternatively, it is possible that the viral titer was low but recombination efficiency was high.

The number of *phoP* transductants generated was surprising, as no transduction defects associated with *phoP* have been reported. However, this may be because P1*vir* transduction has fallen out of use in favour of new techniques for genetic manipulation, such as recombineering. Moreover, negative results are infrequently published. However there are mentions of *phoP* transductants, made without reported difficulty, going back several decades[102]. There are a number of reasons why this might be the case. One such reason might be associated with changes in methodology over time and variations in laboratory practices. Although there have been no large changes in the published protocol for P1*vir* transduction over many decades, it is undeniable that small variations to the published methodology are common between laboratories. Discussed further in later chapters, small variations in the methodology can have a great effect on the successful transduction of *phoP* using P1*vir* transduction.

Perhaps less likely, it is also possible that genetic variations in laboratory strains could also account for the observable difference between this work and the literature. It is conceivable that SNPs, sufficient to alter the expression or activity of proteins, could change the genetic background of the recipient strain enough to alter transducibilityin unexpected ways. Using generalised transduction, it is possible to transduce most of the *E. coli* genome, but there are some genes that can only transduced at extremely low frequency or not at all, such as with *carA*, which expresses the protein carbamoyl phosphate synthetase[197] and *recD*, respectively[154]. However, the transduction defect associated with *carA* can be overcome in the presence of other secondary mutations.

Furthermore, *E. coli* strains exhibit large amounts of variation between subspecies. The percentage of shared transcribed proteins between EHEC and UPEC is approximately 39%[198]. This is due to different subset of virulence factors accumulated by the different pathotypes. It is thought that genes encoding virulence factors that are present in pathogenic *E. coli* originally evolved in commensal *E. coli* to help increase fitness[6],[22]. For example EHEC (and EPEC) carries the LEE pathogenicity island[3] which encodes several important virulence factors, such as Tir (translocated intimin receptor), intimin and the type III

secretion system, while UPEC does not. Further, a significant strain diversity can be observed among EHEC strains. EHEC is broadly classified into four groups: EHEC1, EHEC2, STEC1 and STEC2[3]. However, EHEC is distinct from STEC, as it has a large virulence plasmid and additional virulence factors, such as the type III secretion system[3]. Conversely a defining characteristic of STEC is the ability to produce the Shiga toxin, which is distinct from toxins that utilise the type III secretion system[3]. Another interesting example of inter and intra species diversity is EIEC. EIEC is a heterogeneous group of pathogens, similar to *Shigella* with regard to its pathogenesis but more closely related to *E. coli*[3]. Metabolically, EIEC shares traits attributed to both *Shigella* and *E. coli.* Almost all strains of EIEC lack lysine decarboxylase, which is common in most *E. coli* strains[3]. However, EIEC has the ability to ferment xylose to produce glucose, which *Shigella* species are unable to do, though it is a common trait of *E. coli*[3]. Nonetheless, like *Shigella*, EIEC has the ability to invade epithelial cells[3], has a low infectious dose and is able to cause disease in otherwise healthy individuals.

Furthermore, even metabolic functions that are conserved between subspecies have the potential to be polymorphic. Different isolates have been found to use a variety of different alleles to express proteins that carry out a similar metabolic function[199]. Moreover, long-term studies of *E. coli* evolution have found a radical potential for *E. coli* to change and adapt in novel and unexpected ways[200]. Therefore, it is possible that some isolates of the same *E. coli* strains could respond quite differently to the loss of *phoP* compared to the strain on which this work was carried out.

The findings presented in this chapter, which suggest that the primary observation – inability to transduce *phoP* under standard conditions using P1*vir* transduction – is a generalised phenomenon among *E. coli* K12 strains. Nonetheless, it is possible that this defect only affects a small percentage of isolates. In the absence of data from a wide variety of isolates and subspecies, it is hard to make this determination with complete confidence. However, the data do strongly suggest that both MG1655 and W3110 (two widely used K-12 wildtypes) are affected by the *phoP* transduction defect.

Nevertheless, it was possible to create a *lac*UV5-*phoP*, *phoP* mutant. This demonstrates that the *phoP* transduction defect can be overcome by complementation with an ectopic, inducible copy of *phoP*. IPTG is necessary in this background for transduction, demonstrating the requirement for PhoP. To confirm that *phoP* is conditionally essential, a *phoP:*Cam^R mutant was generated, using the alternative method lambda red recombination. The ability to construct the *phoP*:Cam^R mutant suggests that it is the processes of transduction that cannot be tolerated by *phoP* mutants. These data support the hypothesis that PhoP is conditionally essential, but it is inessential for the viability of the cell under standard conditions.

It is likely, when considering the role of PhoP in cellular metabolism, that a *phoP* mutant would be less likely to survive environmental stressors. Divalent cations, specifically Mg^{2+}, play dominant roles in the structure and function of biological processes, ranging from preventing ribosome degradation to stabilising the structure of the outer membrane. Indeed, low levels of Mg^{2+} are associated with a low growth rate in many microbial species[110]. It is not unexpected, then, that *phoP* mutants would be more susceptible to the environmental stress associated with P1*vir* transduction.

Chapter 4: Exploring the mechanism of the *phoP* P1*vir* transduction defect

4.1 Introduction

Approximately 303 *E. coli* genes are defined as essential, for growth, under optimal conditions[160]. However, organisms seldom have access to optimal conditions, outside of specifically created laboratory environments, and so many genes in *E. coli* are conditionally essential[201]. Furthermore, genes that are singly dispensable may become essential when others that perform analogous functions are also deleted[202]. For example, cells can survive deletion of either outer membrane chaperones *surA* or *skp*, but deletion of both will lead to an unviable phenotype[203]. Moreover, it is sometimes difficult to define a gene as essential, as the deletion of some genes may allow growth, albeit at an exceptionally arrested rate.

Cells can acquire suppressor mutations that dispense with the need for a particular essential gene, one such example of which is the essential *secA*. SecM is required for the expression of *secA*, and so it is classed as an essential gene; however, mutants have been isolated that constitutively express *secA* and so do away with the need for *secM*[160]. Alternatively, mutations may also exist in the mutated gene; a gene may acquire a mutation effecting a residue of a protein, causing loss or diminished function. Nonetheless, a second mutation, causing a change in a residue located at a different site in the polypeptide chain, may restore wildtype functions[204], known as second site suppressors.

In the previous chapter it was postulated that PhoP performs an essential function during P1*vir* transduction, but it is nonetheless dispensable under optimal growth conditions. In this chapter Mg^{2+} starvation is explored as a possible cause for the transduction defect. Moreover, it was found that changing the conditions, time, temperature and recipient strain transduction of *phoP* was possible. The results below support a model in which changing environmental conditions or the induction of secondary mutations, specifically those which induce sigmaE, suppress the observed *phoP* transduction defect.

4.2 Transduction of *mgtA*

Expression of *phoP* helps to mitigate divalent cations starvation, during transduction the concentration of available divalent cations is limited. It stands to reason that the *phoP* transduction defect could be attributed to divalent cation stress. PhoP positively regulates MgtA[205], which is a P-type ATPase is involved in the uptake of magnesium and nickel into the cytoplasm across the cytoplasmic membrane[42]. It is known that metal ion starvation causes many thousand-fold increases in the transcription of *mgtA*[42], so *mgtA* expression should be enhanced during the chelation step in P1*vir* transduction. The most obvious hypothesis is that the transduction defect can be attributed to cytoplasmic Mg^{2+} starvation.

To investigate this hypothesis, the gene-encoding metal ion transporter MgtA (*mgtA*) was tested. Using the standard protocol for P1*vir* transduction, the PhoP-controlled gene *mgtA* was removed from the wildtype strain (AAEC261A) and replaced with a kanamycin-resistant cassette.

It was possible to knock out the PhoP controlled gene (data not shown), *mgtA*, but it is surprising that the loss of *mgtA* did not induce a lethal phenotype in a similar way to the loss of *phoP*. These data support a model in which cytoplasmic Mg^{2+} stress is not the cause of the *phoP* transduction defect. However, divalent cation stress still remains a valid hypothesis. The transduction defect may be caused by extracytoplasmic Mg^{2+} starvation, leading to LPS disruption, which relies on the positive charge provided by divalent cations to maintain stability. If this were the case, an *mgtA* mutant may alleviate extracytoplasmic divalent cation stress, as fewer cations would be transported into the cytoplasm.

4.2.1 Metal ion starvation in *phoP* mutants

In the previous experiment, it was shown that the transduction defect is not mediated by the expression of *mgtA*, and thus it is unlikely to be attributed to cytoplasmic Mg^{2+} starvation. However, citrate intolerance remains a cogent hypothesis for the transduction defect, as divalent metal ion starvation can cause extracytoplasmic stress and OM stability[206].

Divalent cations associate with the LPS's inner core, specifically the phosphate and carboxylate groups[206], which helps overcome the repulsive force of the negative charges on LPS, allow the molecules to bridge, thereby binding them electrostatically. These links can be disrupted by chelation agents or cationic antibiotics[207], which causes electrostatic repulsion between the LPS molecules[87] and has a deleterious effect on the OM. This can lead to large OM deformation of the inner and outer leaflets of the OM, which can merge and alter the function and viability of the OM[87], [208], causing rupturing and increased permeability[206].

It is reasonable to assume that *phoP* mutants would be especially susceptible to divalent metal chelation by citrate and subsequent OM disruption. Increased expression of *eptB*, during P1*vir* transduction, would presumably cause increased repulsion between LPS molecules. Moreover, PhoP positively regulates YneM, a putative membrane protein encoded on opposite strands of the PhoP-regulated MgrR[209]. The expression of this protein is increased in response to exposure with EDTA, a chelation agent, and SDS[209]. It has been hypothesised that this protein responds to both membrane permeabilisation and extracytoplasmic divalent cation starvation[209]. If true then a decrease in fitness would be expected in *phoP* mutants under the conditions in which it is active.

To explore further the idea that citrate intolerance is responsible for the *phoP* transduction defect, a *phoP*:Cam^{R} mutant (created via lambda red recombineering, Chapter 1) was treated with citrate, the metal chelator used during transduction. Both a *phoP* mutant (KCEC4032) and the wildtype (MG1655) were exposed to varying concentrations of citrate, ranging from 0-800mM, over a period of 19 hours. While normally in transduction the cells are exposed to a concentration of 100mM citrate, it was advantageous to investigate whether or not there was an upper or lower limit to the hypothesised citrate intolerance.

The wildtype maintains a growth rate consistent with that seen in the absence of citrate with concentrations of citrate up 100mM. However, the growth rate drops significantly once the concentration reaches 300mM. The wildtype is able to

continue to grow, even in the presence of the highest concentrations of citrate. In contrast, the *phoP* mutant is unable to tolerate many of the citrate concentrations, which are well below those used in transduction (Figure 4.1).

These data suggest a causal link between metal ion starvation and the transduction defect associated with *phoP*. There are two curious anomalies in the data in which *phoP* mutants suddenly show an increased growth rate, at 300mM of citrate after 500 minutes and 50mM of citrate after 300 minutes. One possible explanation for this phenomenon is that the strain has the ability to develop an adaptive phenotype in the presence of a low concentration of citrate (50mM) and a relatively high one (300mM), which could be the result of activating a genetic circuit that functions as an SOS response in the case of extreme metal ion starvation. Beyond the 300mM of citrate, this circuit cannot function and the cells become unviable. Yet, there is no evidence in the literature to support the idea of such a robust and specific chelation stress response. However, Hemm et al. found a set of four small proteins (<50 amino acids) that were up regulated in response to chelation and membrane permeabilisation[209]. Among these four proteins, the expression of two was found to increase in response to chelation alone. One of these proteins the aforementioned PhoPQ regulated YneM, and is therefore unlikely to be active under these conditions[209]. While the other protein, YkgO, is a putative ribosomal protein[209]. However, the most likely explanation for the growth anomalies is that it is evidence of *phoP* mutant populations containing additional suppressor mutations outcompeting the original *phoP* mutant isolate. This assertion is supported further by variations in the growth rate of *phoP* mutants at 50mM and 300mM observed between replicates (Figure 4.1), indicated by large error bars. This variability is consistent with the generation of suppressor mutations, which are likely a result of stress-induced mutagenesis.

The variation in growth rates between experiments is likely a result of fluctuating fixation probability, i.e. the probability that the frequency of a mutation in a population will reach 100%[210]. Under conditions of stress, when the mutation rate increases, beneficial lineages do not necessarily thrive. Fixation probability is dependent on a number of different factors, such as the prevalence of deleterious

mutations associated with adaptive ones, clonal interference, the competition that occurs between different subpopulations carrying independently arising adaptive mutations, the presence of phenotypically heterogeneous subpopulations, etc[211]. Therefore, it would be reasonable to expect some variation in the frequency with which adapted populations arise and are subsequently able to outcompete the original isolate successfully.

The data suggest that *phoP* mutants are sensitive to citrate, but citrate sensitivity does not seem to be dependent on MgtA expression (4.2), which would suggest that citrate has a deleterious effect on *phoP* mutants that is not associated with the uptake of extracellular Mg^{2+} into the cytoplasm. Together, these data support the hypothesis that the transduction defect is mediated by extracytoplasmic divalent cation starvation.

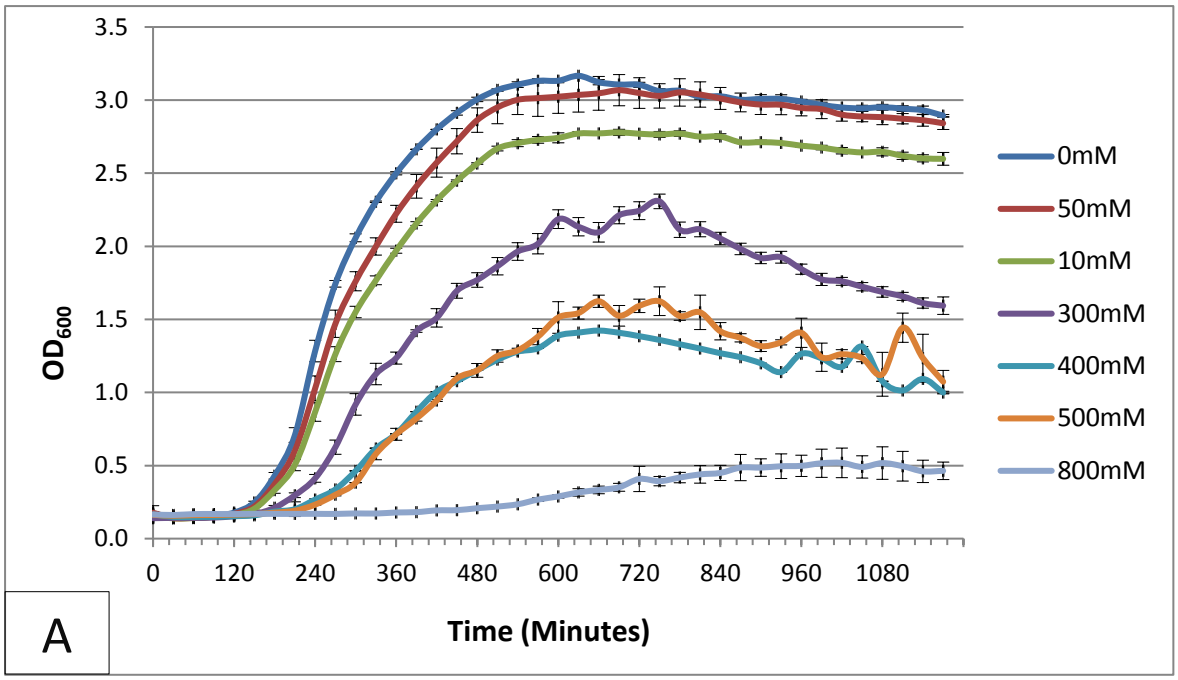
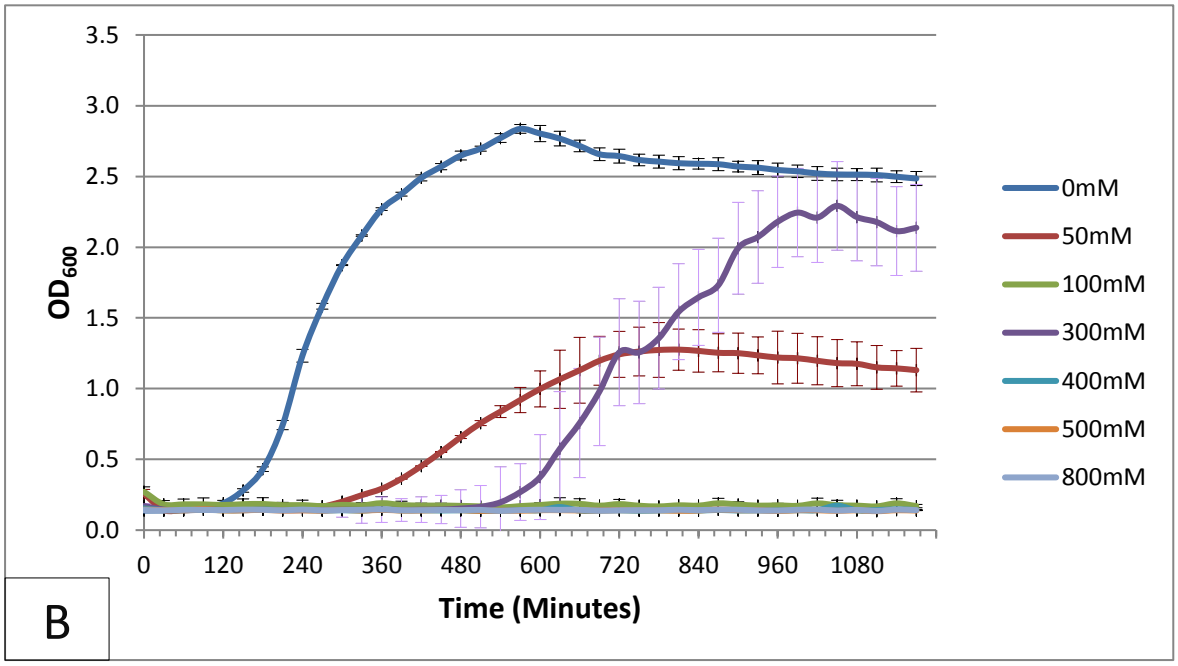


Figure 4.1 Growth curves of A) wildtype (MG1655) and B) *phoP*:Cam^R mutant (KCEC4032) in different concentrations of citrate(0-800mM) over 21 hours. Shown are average cell densities for four biological replicates from readings taken every thirty minutes. Error bars show standard error (n=4).



4.2.2 Testing the effect of citrate on the viability of *phoP* mutants

In the previous section it was found that *phoP* mutants are sensitive to citrate. It was prudent to investigate if this citrate intolerance remained under the conditions

of transduction. To test this hypothesis, the *phoP* mutant strain (KCEC4302) was incubated overnight. The *phoP* mutation used in this assay was created using an alternative methodology (transduced at 28°C, section 4.4).

200µl of the culture was inoculated into LB-citrate (100mM), the same concentration used during transduction. The culture was then incubated for 2.5 hours in a static incubator. Conditions and protocols were maintained similar to those for P1 transduction, albeit lacking the other reagents. The culture was then diluted down to -5, and 100µl was plated onto LB-Kan agar – this process was repeated twice, to produce a total of two replicates. A no-citrate control was also prepared. The wildtype (AAEC261A) was grown in the same conditions (+/- citrate) and compared to the *phoP* mutant, this procedure was repeated three times (Table 4.1).

	Wildtype control (1)	Wildtype control (2)	Average wildtype Control	Average wildtype difference (control- citrate)
Repeat 1	29	32	30.5	22
Repeat 2	33	30	31.5	21
Repeat 3	28	37	32.5	26
	wildtype citrate (1)	Wildtype citrate (2)`	Average wildtype citrate	
Repeat 1	8	9	8.5	
Repeat 2	11	10	10.5	
Repeat 3	6	7	6.5	
	phoP control (1)	*phoP* control` (2)	Average *phoP* control	Average *phoP* difference (control- citrate)
Repeat 1	16	17	16.5	10.5
Repeat 2	18	16	17	11.5
Repeat 3	17	15	16	12.5
	phoP citrate (1)	*phoP* citrate (2)	Average *phoP* citrate	
Repeat 1	6	6	6	
Repeat 2	5	6	5.5	
Repeat 3	3	4	3.5	

Table 4.1. Average number of wildtype(AAEC261A*)* and *phoP(*KCEC4302) colonies after being incubated with 100mM of LB-citrate diluted to -5 and then incubated overnight. Two replicates were done for both the wildtype and *phoP* mutant, denoted by the numbers 1 and 2 in parenthesis.

The results suggest that *phoP* mutants are sensitive to citrate when compared with the wildtype, as the latter has ~50% more colonies compared to the *phoP* mutants when exposed to citrate. These results were analysed using a t-test (Table 4.2) that was used to investigation the effect of citrate on the viability of *phoP* mutants and the wildtype, which was done by comparing the means of their average differences (control-experimental). The alpha level chosen was 0.05. The P-value calculated by the test was 0.006. The null hypothesis, namely that the average difference in the number of colonies of the wildtype, with and without citrate, would not be significantly different when compared to those of *phoP* mutant, was rejected.

Although *phoP* mutants were sensitive to citrate, they were still able to tolerate exposure to concentrations used in transduction, which provides an interesting juxtaposition with *phoP* mutant generated via the lambda red system. Those mutants (4.2.1) could not survive citrate concentrates strains were subjected to in this experiment. This suggests that transduction causes *phoP* mutants to genetically or physiologically adapt.

Although these data seemed to support the observation that *phoP* mutants are sensitive to citrate, it was logical to conduct a similar test that would be more representative of transduction.

	Average differences (control-experimental)	
	Wildtype	*PhoP*
Mean	23	11.5
Variance	7	1
Observations	3	3
Hypothesized Mean Difference	0	
Df	3	
t Stat	7.04	
P value	0.006	
t Critical two-tail	3.18	

Table 4.2. t-Test: Two-sample assuming unequal variances, alpha level 0.05, comparing the average differences (control-experimental) of the number of colonies of wildtype and *phoP* mutants, using data from Table 4.1. These data were generated using Microsoft Excel data analysis tools.

4.2.3 Testing the effect of transduction reagents on the viability of *phoP* mutants

The previous experiments suggest that *phoP* mutants were intolerant to citrate. To further test this hypothesis, a mock transduction was undertaken. Pre-made *phoP* mutants (KCEC4302) and laboratory wildtype (AAEC261A) were subjected to the P1*vir* transduction methodology, but no P1 lysate was used, so that the effect of the reagents could be tested.

Isolates subjected to the transduction reagents (experimental strains) were compared to controls, which were not subjected to any of the reagents used in P1*vir* transduction. The control strains were grown for 2.5 hours in LB. Due to the lack of bacteriophage killing and the selection of antibiotic-resistant colonies, all samples were diluted by factors of 10^{-5} and 10^{-6}, before being plated onto LB agar. For each dilution, two replicates were produced. Three technical replicates were conducted for this experiment.

***PhoP* (Experimental)**	**(1)**	**(2)**	**Average**		***PhoP* (control)**	**(1)**	**(2)**	**Average**	**Average difference (control-experimental)**
Repeat 1	13	19	16		Repeat 1	60	54	57	41.00
Repeat 2	23	24	23.5		Repeat 2	53	55	54	30.50
Repeat 3	17	18	17.5		Repeat 3	47	51	49	31.50
Total Average			19		Total Average			53.33	34.33
Wildtype (Experimental)	**(1)**	**(2)**	**Average**		**Wildtype (control)**	**(1)**	**(2)**	**Average**	**Average difference (control-experimental)**
Repeat 1	101	104	102.5		Repeat 1	118	127	122.5	20.00
Repeat 2	105	91	98		Repeat 2	121	119	120	22.00
Repeat 3	93	98	95.5		Repeat 3	116	113	114.5	19.00
Total Average			98.67		Total Average			119	20.33

Table 4.3. Average number of wildtype (AAEC261A*)* and *phoP (*KCEC4302) colonies after being subjected to transduction reagents (experimental) and controls, incubated overnight and then diluted to -5. One of the wildtype -6 values in the dataset was abnormally high (Appendix Table C.1), which was most likely due to contamination. As a result of this value, the -6 values were discounted and not used in the analysis of the results.

The data suggest that under normal transduction conditions, the *phoP* mutant is less viable than the wildtype, as the total average difference between control and experimental strains was almost 60% higher for *phoP* than the wildtype (Table 4.3). However, it was prudent to conduct several statistical tests, in order to provide a more robust analysis. First, a student's t-test was performed on the average number of colonies generated after exposure to transduction reagents and the control, for *phoP* mutants (Table 4.4A) and the wildtype (Table 4.4B). The alpha value chosen was 0.05, in order to test if transduction reagents had a statistically significant effect on viability.

phoP	Experimental	Control
Mean	19	53.33
Variance	15.75	16.33
Observations	3	3
Hypothesized Mean Difference	0	
Df	4	
t Stat	-10.50	
P value	<0.001	
t Critical two-tail	2.78	

Table 4.4A. t-Test: Two-sample assuming unequal variances, alpha level 0.05, comparing the average numbers of *phoP* 'control' colonies with average numbers of wildtype 'experimental' colonies from Table 4.3. These data were generated using Microsoft Excel data analysis tools.

wildtype	Experimental	Control
Mean	98.67	119
Variance	12.58	16.75
Observations	3	3
Hypothesized Mean Difference	0	
Df	4	
t Stat	-6.50	
P value	0.003	
t Critical two-tail	2.78	

Table 4.4B. t-Test: Two-sample assuming unequal variances, alpha level 0.05, comparing the average numbers of wildtype 'control' colonies with average numbers of wildtype 'experimental' colonies from Table 4.3. These data were generated using Microsoft Excel data analysis tools.

Analysis (Table 4.4A and Table 4.4B, *phoP* mutants and the wildtype, respectively) showed a significant difference between the experimental and the control strains for the *phoP* mutants and the wildtype (<0.001 and 0.003, respectively). The null hypothesis, namely that transduction reagents do not affect the number of colonies observed, can be rejected for both *phoP* and wildtype strains. The analysis conducted above suggests that transduction reagents did indeed influence the viability of both the wildtype and the *phoP* mutants. However, it did not test whether that effect was significantly different for both strains. This was done by performing a second t-test, by analysing the significance between the average differences (control- experimental) of *phoP* and the wildtype (Table 4.5). Once again, the alpha level chosen was 0.05.

	***phoP* average difference (control-experimental)**	**wildtype average difference (control-experimental)**
Mean	34.32	20.33
Variance	33.57	2.33
Observations	3	3
Hypothesized Mean Difference	0	
Df	2	
t Stat	4.05	
P value	0.056	
t Critical two-tail	4.30	

Table 4.5. t-Test: Two-sample assuming unequal variances, alpha level 0.05, comparing the average difference (control-experimental) in numbers of colonies of wildtype and *phoP* mutants, using data from Table 4.3. These data were generated using Microsoft Excel data analysis tools.

The P-value was 0.056 while the alpha value was 0.05. Although the results approached the borderline of significance, the null hypothesis, namely that transduction does not have a significantly different effect on the viability of *phoP* mutants and the wildtype, could not be rejected.

This was a surprising outcome, as previous experiments (4.2.1 and 4.2.2) suggested that *phoP* mutants were significantly more sensitive to citrate than the wildtype, so presumably this should also be true for exposure to transduction reagents, which included citrate. It is possible that some of the transduction reagents used in this experiment, probably the addition of 10mM $CaCl_2$ and 20mM $MgSO_4$, mitigated the effect of citrate on the *phoP* mutants. However, the possibility still remains that these *phoP* mutants adapted genetically by accumulating suppressor mutations and so were less sensitive to transduction reagents.

4.3 *phoP* mutants can adapt to the conditions of P1*vir* transduction

Genetic adaptation, through selection of advantageous mutations, will lead to greater fitness under altered conditions. The process of genetic adaption allows *E. coli* to thrive in conditions that would normally be sub-optimal. Many studies have

shown how *E. coli* can adapt quickly to changing environmental conditions. Isolates from older cultures of, 10 or 20 days will out-compete isolates from younger cultures, by increasing relative fitness via genetic adaption[212]. This is referred to as "growth advantage in the stationary phase" (GASP) phenotype[212], which is associated with a number of mutations, such as *rpoS*819 which leads to attenuated sigmaS activity[212].

Genetic adaption usually comes about through spontaneous mutation, where genes can be silenced, attenuated, duplicated, etc. Some cells are more likely to undergo mutation. Within populations of *E. coli* cells a sub-population of "mutators'' is usually maintained. The mutation rate of these cells is increased due to disruptions in DNA replication or repair[213]. Under certain circumstances, such as population bottlenecks, these cells can help beneficial mutations become fixed in a population.

However, it has also been found that environmental factors can cause genetic instability, thereby increasing the mutation rate and chances of increased fitness, in a process termed ''stress-induced mutagenesis"[214]. There are many different mechanisms through which stress-induced mutagenesis is carried out, but they are usually linked to stress responses – the aforementioned *rpoS*819 mutation, found in ageing colonies, is one such example. Along with sigmaS, sigmaE and subsequently envelope stress have also been implicated as important mediators for stress-induced mutation. Gibson et al. for instance, found that defective stress induced mutagenesis is associated with an inactivated *rpoE* promoter (P2)[214], suggesting that envelope stress can increase mutation rate.

It has been demonstrated that *phoP* mutants made via transduction have a different phenotype than those that have not gone through the same process. *phoP* mutants constructed via transduction have a greater tolerance to citrate than those constructed via lambda red. One possible hypothesis is genetic adaption. Cells lacking *phoP* undergo genetic adaption, additional suppressor mutations are selected for, during transduction dispensing with the need for *phoP*. To explore this hypothesis, *phoP* mutants were generated at 28°C (KCEC4302) and cured

(KCEC4425), in an attempt to test if the mutant could remain viable during transduction under normal conditions.

The *phoP* mutation used in this assay was created using an alternative methodology (transduced at 28°C, section 4.4). The mutant was then cured using the standard procedure, which leaves a scar between the FRT sites but does not affect the reading frame. This mutant was then subjected to P1*vir* transduction under normal conditions, using a lysate generated from a *phoP*::Kan^R mutant (JW1116-1).

It was possible to transduce a Kan^R cassette into the *phoP* locus (of strain KCEC4302) under standard transduction conditions (37°C incubated overnight) (data not shown). These conditions would normally prohibit the transduction of a *phoP*-Kan^R cassette into the wildtype background. An average of 13 kanamycin resistant colonies was yielded after three experiments. PCR (data not shown) was used to confirm the absence of the *phoP* gene, which demonstrates that the *phoP* mutants generated via P1*vir* transduction were pre-adapted to transduction.

Indeed, these data provide confidence in the primary hypothesis that PhoP is indispensable during transduction under standard conditions. The data suggests that *phoP* mutants cannot tolerate the specific conditions of P1*vir* transduction, and *phoP* transductants are forced to overcome and adapt to the loss of *phoP* to remain viable. Nonetheless, once the cells have adapted accordingly, the P1*vir* transduction of *phoP* is possible.

4.3.1 *phoP* mutants can be transduced if the standard incubation times are changed

The previous experiment showed that *phoP* mutants adapted after being subjected to transduction, making further attempts of transduction possible and providing further evidence that *phoP* is indispensible during transduction. However, the effects of loss or inactivation of an essential gene can vary. Herring et al. in this regard, found that mutation of the *murA* gene caused almost immediate inviability,

while a *map* gene mutation took several generations to cause growth arrest[215]. It was thought prudent to explore the hypothesis that a *phoP* mutation was growth-inhibitory rather than lethal. It was hypothesised that by extending the incubation time the *phoP* transduction defect could be better characterised.

The final incubation time was extended from overnight (18 hours) to 42 hours (24 hours longer than the standard time). *phoP (*JW1116-1) and *phoQ (*JW1115-1) mutations were transduced into the laboratory wildtype (AAEC261A), and once again, *phoQ* transduction was used as a positive control. All strains were incubated for 2.5 hours at 37°C. After the first incubation period, 100µl of each sample was plated onto two LB-Kan agar plates. Once plated, the control samples were incubated overnight at 37°C and then counted. The experimental strains were incubated further for an additional 24 hours. This procedure was repeated three times, giving a total of three pairs of plates for each condition.

Under these conditions it was possible to obtain *phoP* transductants at 37°C after 42 hours (Table 4.6). The data suggest that the transduction defect is not lethal but growth-inhibitory. The slow growth observed here may be evidence of suppressor mutations. It is also possible that *phoP* is indispensable during transduction, but under these conditions, the *phoP* transductants are able to remain viable for a sufficient amount of time to allow for the selection of suppressor mutations. These mutations, presumably, compensate for the loss of function of PhoP and allow for the generation of transductants.

Alternatively, it is possible that *phoP* is not essential during P1*vir* transduction but is important for viability, where transductants are generated but impaired, starved of divalent metal ions which are necessary for a host of biological processes[110]. However, slow growth was not associated with the *mgtA* mutation, suggesting slow growth is associated with loss of *phoP*, rather than Mg^{2+} starvation.

Time of transduction	Number of *phoP* transductants				Number of *phoQ* transductants			
	Repeat 1	Repeat 2	Repeat 3	Average	Repeat 1	Repeat 2	Repeat 3	Average
16 hours	0,0 0,0 0,0	0,0 0,0 0,0	0,0 0,0 0,0	0	10,18 6,7 8,12 (61)	13,7 9,10 22,18 (79)	11,14 8,11 9,6 (59)	8.5
40 hours	10,7 11,6 5,8 (47)	9,2 14,6 5,9 (45)	4,6 14,11 6,9 (48)	7.75	1,6 16,10 10,21 (64)	10,16 9,13 9,18 (75)	8,14 7,10 14,6 (59)	11

Table 4.6. Average number of AAEC261A *phoP* and *phoQ* transductants generated in the wildtype (AAEC261A), using 50µl of lysate, after 42 and 18 hours of incubation. Averages generated from three repeated experiments. Averages of each set of replicates are shown in brackets.

4.3.2 The inability to create suppressor-free *phoP* transductants supports the notion of genetic compensation for loss of the gene

The data presented in this chapter suggest that *phoP* mutants adapt during transduction. The simplest hypothesis is that transduction under altered conditions allows for the selection of mutants with secondary suppressor mutations. A simple assay was developed to test this hypothesis, the aim of which was to confirm the presence of expression heterogeneity in *phoP* isolates. To avoid expensive sequencing, an assay based on the expression of *rpoH*P3-*lacZ* in *phoP* mutants was developed, which was done by analysing the expression of *rpoH* from its third promoter. *rpoH* has four promoters, the third of which, P3, is regulated by sigmaE[216]. *rpoH*P3 expression correlates with sigmaE activity and is widely used as an indication of sigmaE activity[217]. *phoP* was transduced into the reporter

strain *rpoH*P3-*lacZ*, at 28°C (4.4), which yielded the strain *rpoH*P3-*lacZ*,*phoP*. Ten different *rpoH*P3-*lacZ*,*phoP* isolates were used to test for heterogeneity and increase confidence in the result.

A β-galactosidase assay on several different isolates of the strain *rpoH*P3-*lacZ*,*phoP* was carried out (Figure 4.2). The β-galactosidase produced by the mutants ranged from 10-fold (0.69)(KCEC5124) to 1.3-fold (4.88)(KECE5129) decreases compared to the wildtype (6.74)(CAG45114). All the *phoP* mutants showed a decrease in β-galactosidase production and thus *rpoH*P3 expression. These data show that isolated *phoP* mutants, even under non-standard conditions, acquire mutations.

It is unlikely that these wildly fluctuating expression patterns are indications of errors, as there seems to be little standard between the mutants. It is more likely that these patterns are emblematic of disparate, heterogeneous, compensatory mechanisms.

Indeed, cells could have any number of suppressors that might compensate for the loss of *phoP* functionality under the conditions of transduction. Finally, these data suggest that when *phoP* mutants can be isolated via P1*vir* transduction under non-standard conditions, this is the result of additional mutations that suppress the need for *phoP*. Moreover, this result casts doubt on the validity of any *phoP* mutants created by transduction.

Finally, all the *phoP* mutants tested exhibited a lower sigmaE activity than wildtype. Considering the loss of divalent cations and the envelope stress of the outer membrane that must exist during transduction, it would be expected that sigmaE activity would be increased. The possible implications of sigmaE activity being depressed in a *phoP* mutant are discussed in greater detail later in this work.

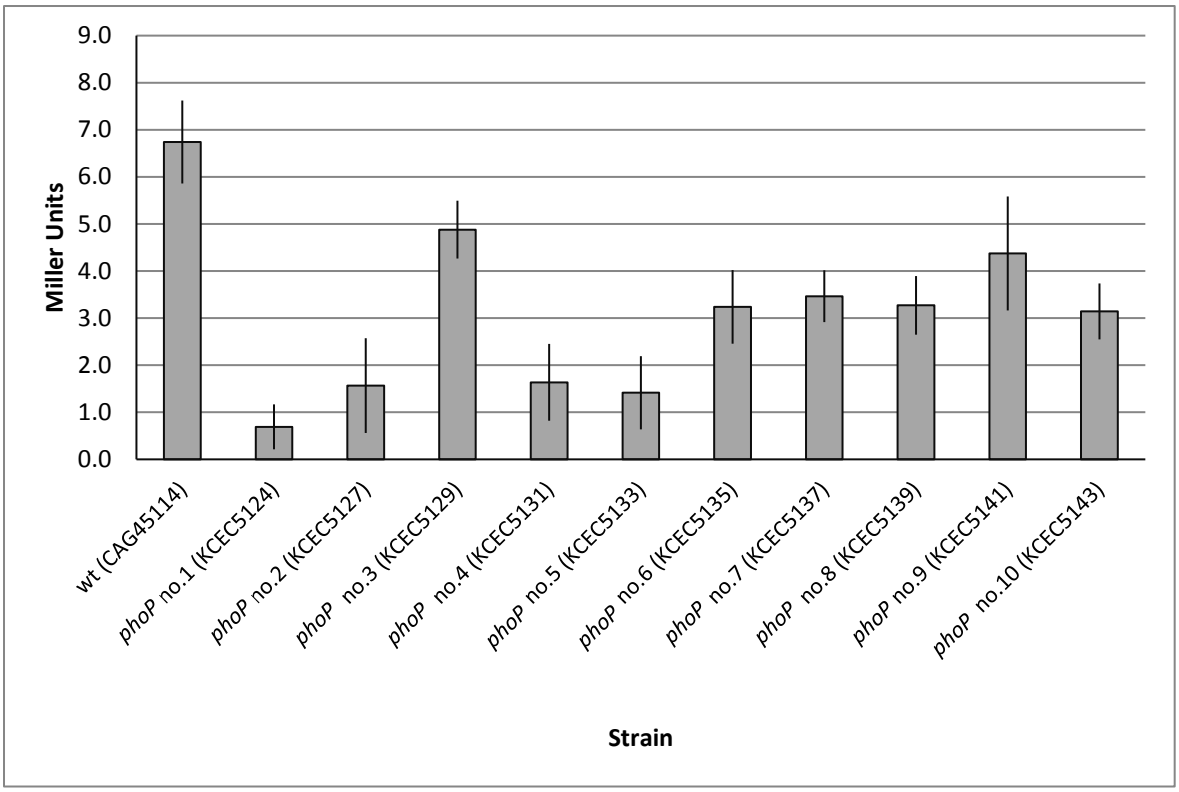


Figure 4.2. β-Galactosidase assay in LB, the heterogeneity of expression of sigmaE activity of *phoP* mutants and the wildtype *rpoH*P3-*lacZ*(CAG45114). Error bars show 95% confidence intervals (n=4).

4.4 *phoP* mutants can be transduced, if the standard temperature is changed

Temperature has a profound impact on the physical properties of bacterial membranes, and it is mediated by the constituents of the membrane. Phospholipids can transition between liquid, disordered, and gel-ordered phases[218]. The point at which the transition takes place is dictated by the ratio of constituents of the OM, LPS, OMP and phospholipids[23].

Phospholipids contain both unsaturated and saturated fatty acid tails. Most phospholipids contain fatty acid chains that are saturated or monounsaturated (polyunsaturated fatty acid tails appear at a much low frequency)[12]. Unsaturated fatty acid tails contain a 30 degree bend. In contrast, when fatty acids are saturated they have straighter tails lacking the characteristic bend of unsaturated fatty acids[24]. Under physiological conditions, the lipid bilayers are largely fluidic, but at

lower temperatures, the membrane undergoes a reversible transition[96], fatty acids are compressed together[96]. Straight-tailed saturated fatty acids press together and create a dense membrane; however, the bend in unsaturated fatty acid tails prevents these molecules from compressing as much as their saturated counterparts, which makes the membrane more fluid and helps protect against rupturing[25].

As the cultivation temperature is lowered, *E. coli* will incorporate more unsaturated fatty into its membrane[98], and in the same way as lipids, LPS also exhibits a form of thermal dichotomy. It is conceivable that such changes that occur in the membrane at lower temperatures may mitigate the effect of the *phoP* transduction defect, presuming the defect is associated with OM instability. Moreover, sigmaE also shows enhanced activity at lower temperatures[219].

The evidence presented herein supports a model in which the transduction defect arises from extracytoplasmic divalent cation stress. By changing the temperature, OM fluidity and constituents, it was hypothesised that *phoP* mutants could be isolated via transduction.

To determine if *phoP* transductants can be isolated at low temperatures, transduction was carried out at 28°C rather than the standard 37°C. It was possible to transduce *phoQ* at 37°C (Chapter Three), so the transduction of a *phoQ* mutation was used as a positive control. *phoP* and *phoQ* were transduced into the laboratory wildtype (AAEC261A). P1*vir* transduction was carried out using 50µl of *phoP* (JW1116-1) and *phoQ (JW1115-1*) lysate at 37°C and 28°C. The experimental strains were incubated for 2.5 hours at 28°C. After this incubation period, 100µl of each sample was plated onto two plates of LB-Kan agar, forming two replicates for each sample, and each replicate pair had two repeats, so there were three replicate pairs in total. Once plated, the samples were incubated overnight at 28°C and 37°C, respectively. This procedure was repeated in triplicate.

Using the procedure outlined above, it was possible to isolate kanamycin-resistant colonies at 28°C but not at 37°C (Figure 4.3). PCR was used to demonstrate the loss of *phoP* and the presence of kanamycin-resistant colonies (Data not shown) at 28°C.

This result demonstrates that the *phoP* transduction defect can apparently be suppressed under modified conditions. The change in the OM composition and general fluidity are well documented in response to temperature [217], [220], these data further supports a model in which the transduction defect is mediated by membrane stress this provide a cogent explanation for the *phoP* transduction defect.

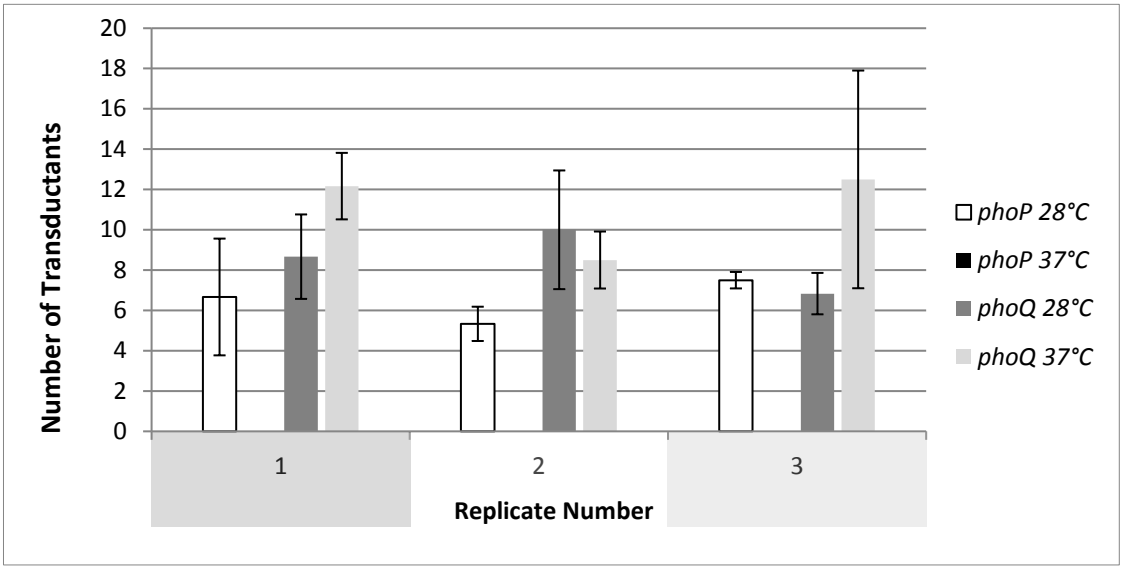


Figure 4.3. The number of *phoP (*JW1116-1) and *phoQ (*JW1115-1) transductants generated in the wildtype (AAEC261A), using 50µl of lysate at 28°C and 37°C, represented as a bar graph. Three repeats were performed and each had three technical replicates (replicates 1-3) for both genes (*phoP* and *phoQ*) at both temperatures (28°C and 37°C). Each bar represents the average number of transductants from three experiments. Error bars show standard deviation (n=3). Full data are shown in Appendix Table C.2.

4.4.1 *phoP* transductants can be isolated under standard conditions in suppressor mutant backgrounds

It was hypothesised that the *phoP* transduction defect was not associated with cytoplasmic divalent cation starvation, as there was no transduction defect associated with MgtA (4.3). However, because *phoP* mutants seemed to show intolerance to citrate, and considering PhoPQ's position in regulating and modifying the genes associated with the LPS, OM stress and disruption seemed logical avenues to explore as possible determinants of the transduction defect. To do this

phoP was transduced into a number of strains with known mutations (Table 4.7). This was known to be possible because previous work in the Blomfield laboratory had found an initial suppressor mutation.

In the initial stages of the project, fortuitously it was found that *ΔphoP* could be transduced successfully into a *gshA* mutant (Dr Alex Moores, personal communication), which is necessary for the synthesis of the low molecular weight thiol, glutathione (L-γ-glutamylcysteinylglycine)(GSH)[221]. GSH helps protect the cell from free radicals and electrophilic compounds, and with the correct folding of proteins containing disulphide bonds, covalent bonds formed by oxidation of the thiol groups between cysteine residues[221]. GSH is synthesised in the cytoplasm by two enzymes, γ-Glutamate-cysteine ligase (*gshA)* and glutathione synthase (*gshB)*[221], and is transported into the periplasm by the ATP-dependent transporter CydDC[221]. Once there, GSH works in concert with the Dsb systems DsbAB and DsbCD, which are integral for to disulphide bond formation and isomerisation. DsbA and DsbC are periplasmic proteins, DsbA is the primary disulphide bond donor in the periplasm introducing disulphide bonds into proteins or polypeptides that are transported into the periplasm[221], while DsbC performs isomerisation of disulphide bonds. Both proteins, DsbA and DsbC, work in concert with the integral membrane proteins DsbB and DsbD, respectively[221]. The system relies on GSH to help maintain its redox state, either directly in the case of DsbC or indirectly by reducing protein thiols, that are then oxidised by DsbA, during which GSH is oxidised into glutathione disulphide[221]. PhoPQ activation is dependent on the redox state of the periplasm[222]. DsbA activates MgrB, which inhibits PhoQ and substantially PhoP activation, via oxidising the several cysteine residues found on the protein[222]. It is possible that by inhibiting the synthesis of GSH, by deleting *gshA*, the periplasm is made into a more reducing environment, which in the wildtype cell would cause the inhibition of PhoPQ via MgrB. It is possible that under these conditions the cell employs another mechanism to fulfil the role of PhoP, so under these conditions it is dispensable.

To find further suppressor mutations, a number of genes that were associated with PhoPQ and the OM were assayed. *eptB, ompT* and *treR* were chosen as initial

candidates. *eptB* modifies the LPS in a manner that decreases the electrostatic charge between these molecules[39], and it is repressed by PhoP via MgrR. It was hypothesised that the transduction defect may have arisen from an overexpression of EptB and the subsequent effect on the LPS; however, no *eptB,phoP* transductants were generated. To rule out EptB as the cause of the transduction defect, an attempt was made to transduce Δ*phoP* into an *mgrR* mutant, but no transductants were generated.

OmpT is an outer membrane protease and targets a broad variety of extracellular, cationic peptides such as protamine[105]. Attempts were made to transduce *phoP* into a *ompT* mutant, but no *ompT,phoP* transductants were generated.

treR negatively regulates trehalose transport and catabolism. Trehalose is a non-reducing sugar, and when synthesised it can be incorporated into the membrane and offer protection to the cell from osmotic stress and desiccation[223]. Without *treR* it was hypothesised that the increased presence of trehalose may offer more protection to the cell during transduction, but no *phoP*,*treR* transductants were generated.

Next, sigmaE-controlled sRNAs were assayed because of sigmaE's position as a key regulator of extracytoplasmic stress. While a *rybB*, *phoP* mutant could not be made successfully, it was possible to create a *micA*,*phoP* mutant, the result of which was initially surprising. MicA has many targets but none seemed directly relevant to the *phoP* transduction defect, except *phoP* itself. Nonetheless, as subsequent experiments were done, a cogent hypothesis was formed as to why *micA* mutants suppressed the transduction defect. This is discussed later on in this work.

SigmaE induction was explored as a possible – direct or indirect – suppressor for the lethal phenotype. Next, Δ*phoP* was transduced successfully into an *rseA* mutant. sigmaE is negatively regulated by the anti-sigma factor RseA, while σE is held by *RseA*, in the inner membrane, until it is degraded by DegS[134]. In an *rseA* mutant, sigmaE is not sequestered by RseA, and so the expression of the sigmaE-controlled genes increases[136], which suggests that the *phoP* transduction defect is overcome by the hyper-induction of sigmaE.

To test further the hypothesis that sigmaE induction suppresses the *phoP* transduction defect, *phoP* was transduced successfully into a *surA* mutant background. SurA is a chaperone protein that escorts properly folded outer membrane porins[203]. The Blomfield Lab has found the loss of *surA* to be an inducer of sigmaE (data not shown). This hypothesis was further tested by successfully making a *rfaH*,*phoP* mutant. Mutation of *rfaH* also causes sigmaE induction (Dr Blomfield, personal communication).

While sigmaE is essential some suppressors, mutations have been found[109]. *ydcQ* encodes a putative DNA-binding protein[135]. In *ydcQ* mutant strains, *rpoE* is dispensable[135], hypothetically because essential products downregulated by sigmaE are regulated by YdcQ[135]. Alternately, it has also been suggested that upon losing sigmaE, the cell triggers a lethal response in the cell, which YdcQ may disrupt[135]. It was hypothesised that *ydcQ* loss may suppress the *phoP* transduction defect, but no *ydcQ,phoP* transductants could be generated. However, because it is not understood as to how *ydcQ* mutants suppress sigmaE essentiality little information can be discerned from this assay.

These data support model in which *phoP* transduction defect is associated with the OM rather than the cytoplasm. Moreover, the data shows that induction of sigmaE suppresses the transduction defect, inferring that loss of *phoP* leads to disruption of the outer membrane. Conversely, it is also possible that PhoP positively regulates the essential protein sigmaE, the induction of which, via genetic mutation, allows the cell to remain viable.

successful suppressors	Function of genes	unsuccessful suppressors	Function of genes
micA	Encodes an σE regulated sRNA.	*rybB*	Encodes an σE regulated sRNA.
rfaH	Encodes a transcriptional antiterminator.	*rraB*	Encodes an antagonist of RNase E.
surA	Encodes an isomerase which helps OMPS to form the correct conformation.	*gadE*	Encodes operons involved in acid resistance and pH homoeostasis.
cydD	Encodes part of the CydDC protein that transports glutathione from the cytoplasm to periplasm.	*ompT*	Encodes an outer membrane protease.
rseA	Encodes an antisimga factor that inhibits the action of σE.	*eptB*	Encodes a phosphoethanolamine transferase, which modifies LPS in response to changes in Ca^{2+}.
gshA	Encodes glutamate-cysteine ligase, which catalyses the first steps in the synthesis of glutathione	*treR*	Encodes the Trehalose repressor, that regulates operons involved in trehalose degradation.
		ydcQ	Encodes a sigmaE essentiality suppressor
		mgrR	sRNA, negative regulator of *eptB*
		ydcQ	Encodes a sigmaE essentiality suppressor

Table 4.7. Genes that failed to suppress the *phoP* phenotype and genes that successfully suppressed the *phoP* phenotype along with their functions.

The initial suppressor, *gshA*, suggest that the redox state of the periplasm is also associated with the transduction defect. It is difficult to marry these two observations. It is possible that -lack of GSH – and thus proper disulphide bond formation – may affect the folding of the outer membrane protein, and subsequently cause the induction of sigmaE. However, many OMPs lack cysteine, and so their folding is not dependent on the Dsb system. Moreover, *gshA* mutation is not associated with increased sigmaE activity (data not shown). It is likely that *gshA* mutation and sigmaE induction represents different mechanisms of suppression.

4.4.2 The effect of Procaine on *phoP* transduction

SigmaE is active under the range of stimuli, with its primary function being the maintenance and protection of the cell envelope. The core set of sigmaE-regulated genes is associated with the assembly, synthesis and maintenance of the LPS and OMPs, however sigmaE regulates a number of other targets including negatively impacting *phoP* expression, discussed previously[112]. It is possible that during P1*vir* transduction, when sigmaE is hyper-induced, *phoP* is no longer conditionally essential.

To support the hypothesis that sigmaE induction suppresses the *phoP* transduction defect, transduction was carried out using procaine, a drug that induces sigmaE in *E. coli* via the OmpR/EnvZ two-component system[85], [173]. Transduction was done in accordance with the standard protocol used for P1*vir* transduction, whereby *phoP* and *phoQ* were transduced into the laboratory wildtype (AAEC261A). The samples were plated onto LB-Kan agar containing procaine. The concentration of procaine was chosen in accordance with work done previously in the Blomfield Lab that showed the drug's ability to enhance sigmaE expression (data not shown). A *phoQ* transduction was carried out in parallel to act as a positive control. Both *phoQ* and *phoP* transductants were plated onto LB-Kan agar with and without procaine, the latter acting as a negative control. Two biological replicates were done for each gene under each condition. The plates were then incubated at 37°C

overnight. Several *phoP* transductants were subjected to PCR analysis to verify the loss of the *phoP* gene (Appendix Figure A.3).

Using procaine, it was possible to transduce *phoP* successfully under otherwise standard conditions (Table 4.8), which shows that sigmaE induction via the OmpR/EnvZ pathway is sufficient to overcome the *phoP* transduction defect. These data further corroborating the previous findings and show inducing sigmaE, chemically, is sufficient to suppress the transduction defect. This may be because the loss of *phoP* induces membrane stress, and induction of sigmaE is required to correct it. Alternatively, PhoP may have a positive regulatory effect on the induction of sigmaE; thus, sigmaE hyper-induction compensates for the loss of activity.

	***phoP* transductants**				***phoQ* transductants**			
	Repeat 1	Repeat 2	Repeat 3	Average	Repeat 1	Repeat 2	Repeat 3	Average
Procaine	14,13	15,16	17,15	15	20,23	19,17	15,14	18
Control	0,0	0,0	0,0	0	14,13	15,12	12,15	13.5

Table 4.8. The average number of *phoP (*JW1116-1) and *phoQ (*JW1115-1) transductants yielded over three experiments when grown with the addition of 100mM procaine, with two replicates for each mutant and condition per repeat.

4.5 Discussion

Essential genes are a balance between compensation and conservation. Under some circumstances, essential genes can lose their importance if selection pressure is applied[28]. In this chapter, it was observed that altering the standard conditions of P1*vir* transduction (both time and temperature) allows for the mitigation of the observed transduction defect (4.3.1 and 4.4, respectively). Moreover, in this chapter, evidence was provided that supported the hypothesis that the *phoP* transduction defect is associated with extracytoplasmic stress, rather than as a result of cytoplasmic divalent cation starvation.

It was hypothesised that both time and temperature allow for the selection of suppressor mutations that compensate for the loss of *phoP*. Extending this growth time allows slow-growing colonies containing secondary suppressor mutations sufficient time to adapt to the loss of *phoP*. While growth at 28°C increases sigmaE activity and helps mitigate the effect of losing *phoP*, presumably under these conditions cells with fewer or less extreme secondary suppressor mutations are selected.

To confirm the hypothesis that *phoP* mutants genetically adapt to transduction, a P1*vir* transduction was carried out successfully on a cured *phoP* mutant under standard conditions (4.3). This assay demonstrated that once *phoP* mutants had successfully undergone transduction, they were no longer subject to the defect, which strongly supports the hypothesis that genetic adaption was taking place. This notion was supported further by *phoP* mutants generated by P1*vir* transduction being tolerant to citrate, while *phoP* mutants created via lambda red where not so. It was surprising that *phoP* mutants were not significantly less tolerant to transduction reagents than the wildtype. However, because *phoP* mutants created via transduction are thought to be genetically heterogeneous, it is expected that different isolates will have different tolerances to both citrate and the other transduction reagents.

It is expected that *phoP* mutant would be more susceptible disruption of the outer membrane, due to divalent metal ion starvation and subsequent instability of the LPS. Therefore, it is likely that *phoP* mutants are hypersensitive to OM disruptions, an idea supported by the observation that induction of sigmaE facilities the transduction of *phoP* under otherwise standard conditions.

Consequently, it was unsurprising that strains in which sigmaE was induced*, rfaH, surA, rseA*, were able to act as a recipient strain for the transduction of the *phoP* allele. However, with regards to the data presented in 4.4.1, one possible hypothesis that cannot be rejected completely is that the curing procedure rescues the cells from the *phoP* transduction defect. All of the genes assayed (Table 4.7) were cured, excluding the *micA* mutant strain. This hypothesis could be further

supported by the data in 4.3, which show a cured *phoP* mutant is able to undergo P1*vir* transduction successfully. Nonetheless, this is unlikely, as there was a direct correlation between genes able to suppress the transduction defect and the induction of sigmaE.

Moreover, sigmaE was shown to be associated with the *phoP* transduction defect, in that the transductants created at 28°C all showed heterogeneity of expression from *rpoH*P3 (inferring that sigmaE activity fluctuates between *ΔphoP* isolates). Yet, it is important to note that from the data generated in section 4.3.2, it is not possible to confirm whether it is a lack or overabundance of sigmaE activity that is associated with the *phoP* transduction defect.

The simplest hypothesis is that PhoP is a negative regulator of sigmaE, and transductants acquire suppressor mutations to lower the concentration of sigmaE in the cell. This suggestion is supported by the observation that all *rpoH*P3-*lacZ*, *phoP* isolates showed diminished expression of *rpoH*P3 compared to the wildtype, although this hypothesis is not at all supported by the observation that inducing sigmaE (using genetic induction or chemical induction via procaine) allows for the transduction of *phoP*, nor is it consistent with the idea that transduction causes loss of divalent cations and subsequent envelope. SigmaE would be necessary to mitigate damage to the OM, so it is more likely that PhoP positively regulates sigmaE, either directly or indirectly, and that if a suppressor-free *phoP* mutant could be generated, the expression of *rpoH*P3 would be uniformly low, compared to the wildtype.

The hypothesis that the *phoP* transduction defect is a result of divalent cation starvation still remains a valid one. Transduction requires adding divalent metal ions to the reaction mix, which negates the negative charge on both surfaces of the bacteriophage and bacterium and is necessary for efficient binding to occur. However, cytoplasmic divalent ion starvation is unlikely to be cause, as it has been demonstrated that MgtA does not mediate the PhoP transduction defect. It is most likely that, in *phoP* mutants, sigmaE activity is depressed and is not sufficient to compensate for the OM damage caused by divalent cation starvation during transduction. Therefore it is probable that the *phoP* transduction defect is a

combined effect between increased extracytoplasmic stress (caused by divalent cation starvation), decreased sigmaE activity.

Chapter 5: The regulation of sigmaE by PhoPQ

5.1 Introduction

There is a well-established link between the cell envelope and PhoPQ [112]. For example, PhoPQ directly regulates the expression of HdeD, a transmembrane protein, which is responsible for acid resistance at high cell density, and several lipoproteins, such as SlyB and VboR[96], [42]. Moreover, PhoP indirectly and directly regulates the expression of OMPs TolC and OmpT, respectively[42]. TolC is an outer membrane protein that acts as a channel for a number different molecules, and it enables the efflux of antibiotics, the export of haemolysins and bacteriophage absorption[105].

Furthermore, PhoPQ plays a prominent role in modifying the LPS, by regulating genes such as *eptB* and *crcA*, the latter of which functions to catalyse the transfer of palmate to lipid A. Moreover, *eptB*, which encodes phosphethanolamine transferase, modifies the LPS core in a Ca^{2+}-dependent manner and decreases electrostatic repulsion in the presence of divalent cations [206], which in turn increases resistance to antimicrobial peptides such as polymyxin [224]. Finally, PhoPQ is negatively regulated by the sigmaE-controlled sRNA MicA [112].

Interestingly, it has been found that MicA is regulated in an Mg^{2+}-dependent manner [225]. Above certain concentrations, MicA will dimerise in the presence of Mg^{2+}, the effect of which is hypothesised to inactivate MicA, either by RNase E degradation or by preventing normal MicA-Hfq binding. Mg^{2+} also contributes to the regulation of sigmaS, and the general stress response. DsrA is a positive regulator of sigmaS[193], Mg^{2+} -dependent dimerisation increases the activity of DsrA by increasing binding efficiency[193]. Perhaps it is not surprising, then, that in the absence of the stimulatory effect of Mg^{2+} on *rpoS* expression, under low Mg^{2+} conditions, PhoPQ acts as a positive regulator of sigmaS by preventing RssB meditated proteolysis [147].

In the previous chapter, the importance of PhoP during P1*vir* transduction was established and strong evidence of its essentiality under these conditions was demonstrated. Furthermore, it was demonstrated that the transduction defect can

be suppressed by the induction of sigmaE and membrane perturbation. SigmaE mediated suppression was established in numerous different ways: genetically via knockout mutants, chemically via procaine and environmentally via lowering the temperature.

In this chapter, the details of PhoPQ's regulatory relationship with sigmaE are elaborated. A novel, positive regulatory link between PhoPQ and sigmaE activity, independent of sigmaE's chief regulator *rseA*, is demonstrated. Conversely, evidence is also presented that demonstrates a second negative regulatory relationship between PhoPQ and *rpoE* expression, adding further complexity to the relationship that exists between PhoP and sigmaE.

5.2 PhoPQ positively regulates sigmaE activity

It was observed that the induction of sigmaE suppresses the *phoP* transduction defect (Chapter 4). In addition, it was shown that *phoP* mutants, transduced at 28°C, exhibit a lower incidence of sigmaE activity when compared to the wildtype (Chapter 4.3.2), and it was postulated that these data may have been an indication of a positive regulatory relationship between PhoP and sigmaE. To test this hypothesis, both *phoQ* (JW1115-1) and *phoP* (JW1116-1) were transduced into the *rpoH*P3-*lacZ* (CAG45114) strain, using P1*vir* transduction. Both *phoP* and *phoQ* transductions were carried out at 28°C, using the method outlined in Chapter 4. A β-galactosidase assay was done to investigate the expression of *rpoH*P3 in both mutants (Figure 5.1).

There was a 0.47-fold and a 0.3-fold decrease in expression from the *rpoH*P3 promoter in the *phoP* and *phoQ* mutants, respectively. These data suggest that PhoPQ positively regulates sigmaE activity, which would support the hypothesis that the transduction defect is a result of accumulated envelope stress and loss of *phoP* that leads to the insufficient induction of sigmaE, thereby causing inviability. Interestingly, the assay shows that the absence of *phoQ* has a greater effect on sigmaE activity than the loss of *phoP*. The most palatable explanation for this is that the *phoP* mutants contain suppressor mutations and that the level of sigmaE

activity shown (Figure 5.1) for *phoP* is not a true representation of a *phoP* mutant's effect on sigmaE activity. While the *phoQ* mutant may contain additional suppressor mutations, the fact that there are no associated transduction defects makes it less likely. Therefore, it is probably that the levels of β-galactosidase measured in the *phoQ* mutant are a more accurate expression of PhoPQ's effect on sigmaE activity. Alternatively, it is possible that in the absence of phoQ, PhoP retains activity, which would be supported by the absence of the *phoQ* transduction defect.

It is likely that this observation, although marred by the appearance of suppressor mutations, does suggest a real regulatory effect between PhoPQ and sigmaE.

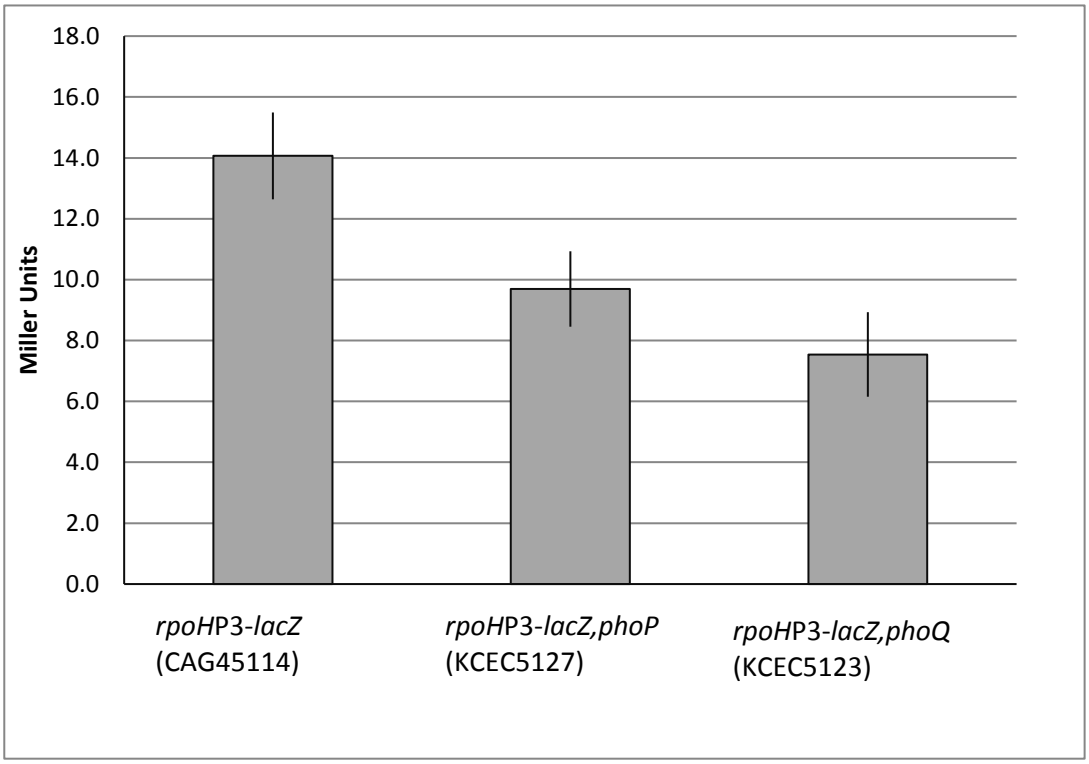


Figure 5.1. β-galactosidase assay showing the effect of *phoP* and *phoQ* mutants on *rpoH*P3 expression. Error bars show 95% confidence intervals (n=4). Strain wildtype - CAG45114. *rpoH*P3-*lacZ,phoP* - KCEC5127, *rpoH*P3-*lacZ,phoQ* - KCEC5123.

5.2.1 The effect of *mgrB* mutation on sigmaE activity

As it was not possible to make *rpoH*P3-*lacZ*, *phoP* or *rpoH*P3-*lacZ*, *phoQ* strains, that lacked suppressor mutations with confidence an *mgrB* mutant was used to investigate the effect of PhoP on sigmaE activity. MgrB is a PhoP-regulated protein that inhibits PhoP activity by interacting with the periplasmic domain of PhoQ, in response to the changing redox state of the periplasm[222], deletion of *mgrB* is known to effect the expression of PhoP regulated genes [222]. The transduction defect, and associated suppressor mutations, were observed in strains that lack PhoP, so it was hypothesised that an *mgrB* mutant would not accumulate suppressor mutations, therefore, it was inferred that an *mgrB* mutant would give a more accurate insight into how altering the expression of *phoP* would affect sigmaE activity.

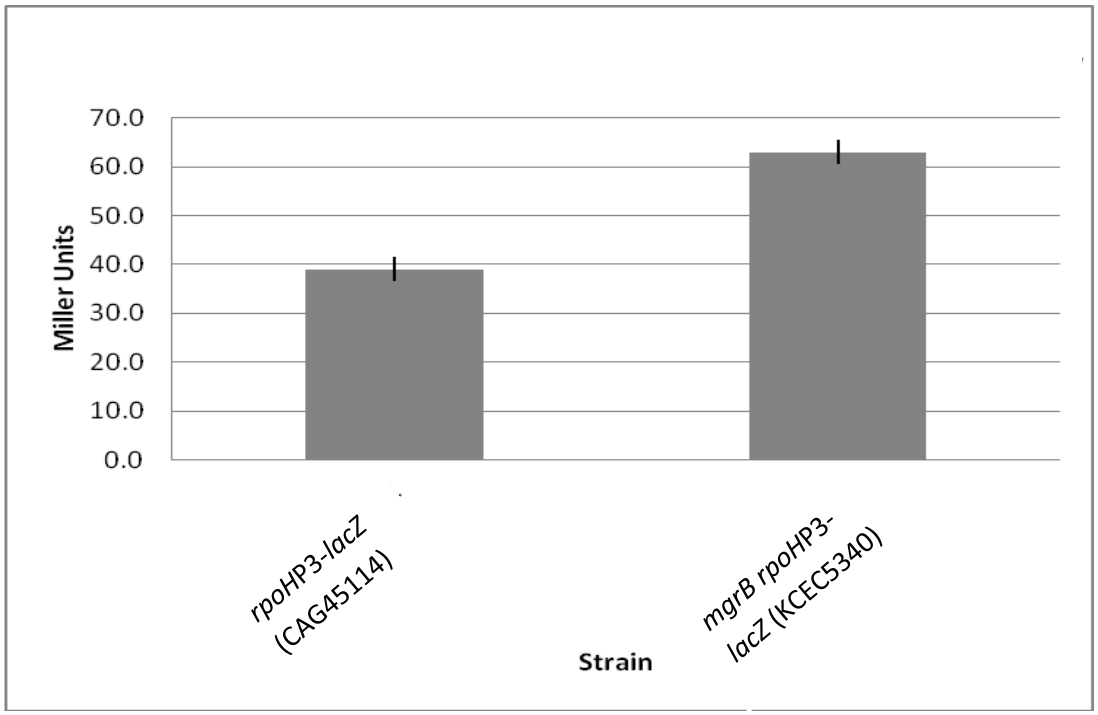


Figure 5.2. β-galactosidase assay showing the effect of the *mgrB* mutant on *rpoH*P3 expression. Error bars show 95% confidence intervals (n=4). Strain wildtype - CAG45114. *rpoH*P3-LacZ *mgrB*; Kan^R -KCEC5340.

An *mgrB* mutation (JW1815-1) was transduced into the *rpoH*P3-lacZ strain (CAG45114), without any discernible defect, and the effect of the mutation on sigmaE activity was explored (Figure 5.2). These data demonstrate an increase in sigmaE activity when PhoP is no longer inhibited by MgrB. Importantly, an increase in sigmaE activity is observed in a strain where suppressor mutations are not expected. Together, Figures 5.1 and 5.2 establish that PhoPQ has a stimulatory effect on sigmaE activity.

5.2.2 The effect of *phoP* loss on sigmaE, demonstrated via complementation

To increase confidence in the previous observations (5.2&5.2.1), the complementation strain *ΔphoP lac*UV5-*phoP* was used with the aim of inducing an increase in sigmaE activity upon adding IPTG, which would establish that increasing PhoP expression has a positive effect in this regard.

MicA is a well-known inhibitor of PhoP [112], and previously it has been demonstrated that the lack of this sRNA acts as a suppressor of the lethal *phoP* phenotype. A *ΔphoP lac*UV5-*phoP, micA-lacZ* strain was constructed via allelic exchange. To achieve this a plasmid, (pAM012), containing *micA-lacZYA, ΔlacZYA* was transformed into a *ΔphoP lacUV5-phoP* strain (KCEC5409). A *ΔphoP lac*UV5-*phoP rpoH*P3-*lacZ* strain was not constructed, due to time and resource constraints; however, *micA* expression is positively regulated by sigmaE, so it provides an appropriate surrogate for monitoring changing sigmaE activity in a way similar to the *rpoH*P3-lacZ reporter fusion[226]. The necessary tools to create a *ΔphoP lac*UV5-*phoP*, *micA-lacZ* strain, the *micA-lacZYA, ΔlacZYA* (pAM012), was easily attainable and already constructed, which meant *micA-lacZYA, ΔlacZYA* could be transformed into the complementation strain. Creating an *rpoH*P3-*lacZ, ΔphoP lacUV5-phoP* strain would mean reconstructing the complementation strain in an *rpoH*P3-*lacZ* background. Once the *ΔphoP lac*UV5-*phoP*, *micA-lacZ* strain was constructed, it was tested +/- IPTG. It was hypothesised that in the IPTG+ strains, *phoP* expression would be induced and subsequently sigmaE activity would increase, which in turn would increase the expression of *micA-lacZ* (Figure 5.3).

These data show a clear increase in β-galactosidase, and therefore *micA* expression, upon the induction of *lac*UV5-*phoP*, while the results show a lowered expression of *micA* in the strains lacking IPTG. These data support the hypothesis that PhoP has a positive regulatory effect on sigmaE activity. Unlike the *rpoH*P3-*lacZ*,*phoP* mutants, is it unlikely that the *micA*- *lac*UV5-*phoP* strains have suppressor mutations, because they were not transduced under conditions in which *phoP* was absent or not expressed. This data provides further supports of the hypothesis that PhoP positively affects sigmaE activity. Together with Figures 5.1 and 5.2, a clear link can be seen between PhoPQ and enhancing not only sigmaE activity, but also a member of the sigmaE regulon.

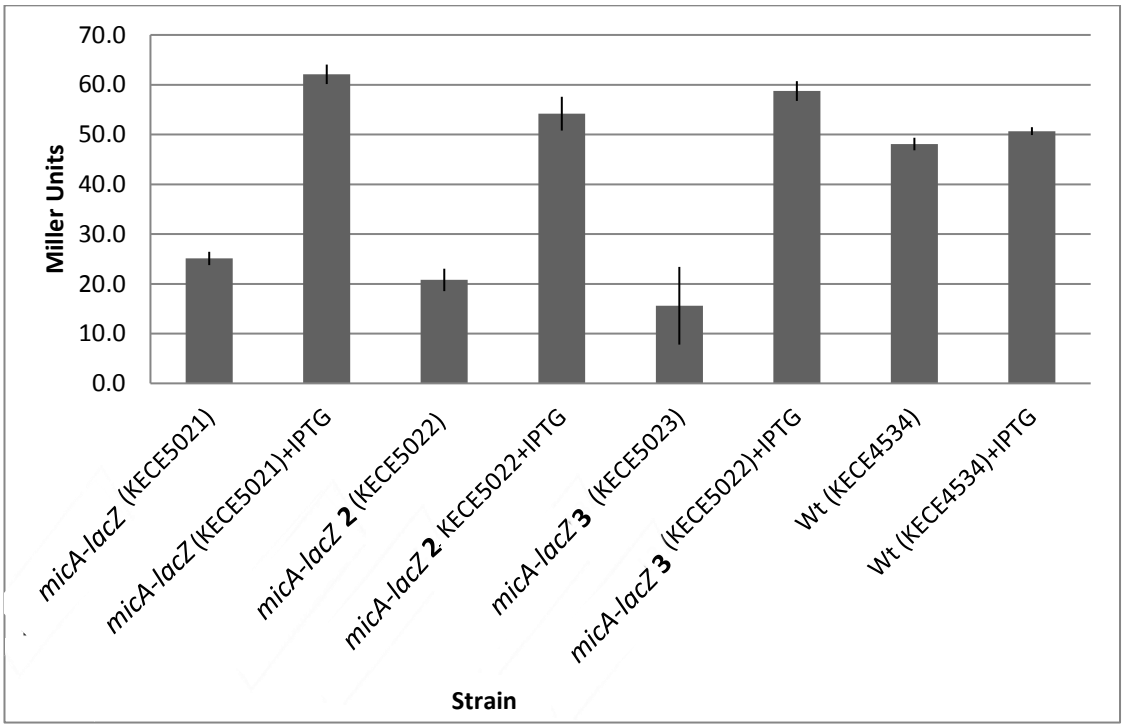


Figure 5.3. β-galactosidase assay showing the effect of IPTG on the expression of *micA* in *a lac*UV5-*phoP* complementation strain . The wildtype is *micA-lacZYA, ΔlacZYA* (KCEC4534). Strains 1-3 are *micA-lacZYA, ΔlacZYA* Ω, *lac*UV5-*phoP* (i- KCEC5021 ii- KCEC5021, iii- KCEC5021). Error bars show 95% confidence intervals (n=4).

5.3 Directly observing the effect of loss of *phoP* on cellular sigmaE concentration

It was possible to observe a drop in expression of the *rpoH* P3 promoter (Figure 5.1) in both the *phoP* and the *phoQ* mutants. Although using the *rpoH*P3 expression as a measure for sigmaE activity is a well-established practice [227], [228], it only allows for measuring of activity rather than total protein concentration. This presents a number of different hypotheses. PhoP could be activating sigmaE indirectly via an unknown FX factor; alternately, it is possible that PhoPQ affects the amount of sigmaE, either transcriptionally by regulating the expression of *rpoE* or the post-transcriptionally, by regulating proteolysis of sigmaE, in the same way that PhoPQ regulates sigmaS[72]

Therefore, it was advantageous to attempt to look at PhoP's direct effect on sigmaE protein concentration *in vivo*. In order to measure levels of sigmaE directly, a semi-dry colorimetric Western blot was used with primary sigmaE polyclonal antibody (donated by the Gross lab). Four strains were chosen for this assay: a sigmaE protein control 10ug/ml (donated by the Ades lab), used as a control for the antibody used, the wildtype (CAG45115), an *mgrB* mutant (KECE5340) (in which PhoP activity is enhanced) and an *rseA* mutant (KCEC4420). The *rseA* mutant was used as a control for this experiment, because without RseA, sigmaE is no longer sequestered to the inner membrane. It has been shown that *rpoE* is autoregulated, so it would be expected that in an *rseA* mutant *rpoE* expression would be enhanced[229].

The strains outlined above, excluding the pure sample of sigmaE protein, were first prepared in accordance with the lysis protocol. Once lysed, the samples were assayed for their protein concentration, using a standard Markwell assay, and then diluted down to the same total protein concentration of 1.5ug/ml. The colorimetric Western blot was then carried out on these samples (Figure 5.4A).

These data were then subjected to an image enhancement the aim of which was to increase clarity. The location in the original image (Figure 5.4A) where the sigmaE

protein was expected to be, just above the 22kDa band, was coloured corrected by a 400% increase in saturation(Figure 5.4B).

Once data for the Western blot was completed, the results were quantified using a readily available protocol, utilising the software ImageJ[230]. The ImageJ plugin "subtract background'' was used to mitigate the effect of the high background signal on the analysis. This plugin uses a "rolling ball" algorithm to correct for uneven background signal[231]. Bands were selected manually and their intensities converted to peaks. The area under these peaks was then given an arbitrary value, with higher intensity bands corresponding to larger peaks and a greater area. These values were termed "relative intensity," because they were expressed in arbitrary units and so were only directly comparable to each other. The value for each band peak area was as follows: protein control-29643.25, *rseA*-25172.76, wildtype-15725.46 and *mgrB*- 9402.054. These values were then normalised using the standardise function of the Microsoft Excel software package. This function returns a Z-score for each intensity value by subtracting the population mean from the value and then dividing this number by the population standard deviation (Figure 5.4C).

sigmaE is a 24kDa protein[114], the protein bands in that position suggest the protein control and *rseA* mutant have the highest concentration of sigmaE. This is not unexpected, as *rpoE*P6 responds to sigmaE and allows autoinduction[229]. While the lowest concentration of sigmaE is associated with the *mgrB* mutant in which PhoP is uninhibited(Figures 5.4A&B). These data were surprising in this respect, as it was hypothesised that higher sigmaE activity observed from previous experiments (Chapter 4&5) would correlate with increased sigmaE expression.

The data suggests sigmaE expression, in the *mgrB* mutant, is diminished rather than epistatic with the wildtype, which suggests that PhoP is a negative regulator of sigmaE expression – an unanticipated result.

These data and the previous observations suggest that PhoP regulates sigmaE both negatively and positively, increasing activity and diminishing expression. However, it would seem that the effect of decreasing expression is subsitory, since PhoP

causes overall sigmaE activity to increase. Therefore, PhoP should be considered a positive regulator of sigmaE. However, this conclusion is contingent on the cogency of the Western blot data presented herein. Several weaknesses, such as a high background signal or lack of a protein loading controls, increase the difficulty in making accurate assertions based on these data. These weaknesses are further discussed at the end of this chapter.

If PhoP is a key regulator of sigmaE activity then a mechanism to inhibit excess expression might be expected. As PhoP is a transcription factor that indirectly regulates a large number of genes, the simplest hypothesis to explain this observation is that PhoP utilises two mechanisms of action to regulate sigmaE. These mechanisms may be direct and indirect, inhibiting expression but increasing activity indirectly through other means.

Figure 5.4A. Original gel. A colorimetric Western blot image using a sigmaE polyclonal antibody (donated by the Gross lab). Run in a Bis-Tris 10% MOPS buffer. Showing the sigmaE concentration in four samples, lane 1-10ug/ml of sigmaE protein (donated by the Ades lab), 3-(KCEC4420), 5-wildtype (CAG45115), 7-*mgrB* (KECE5340) and M denotes the pre-stained molecular weight marker.

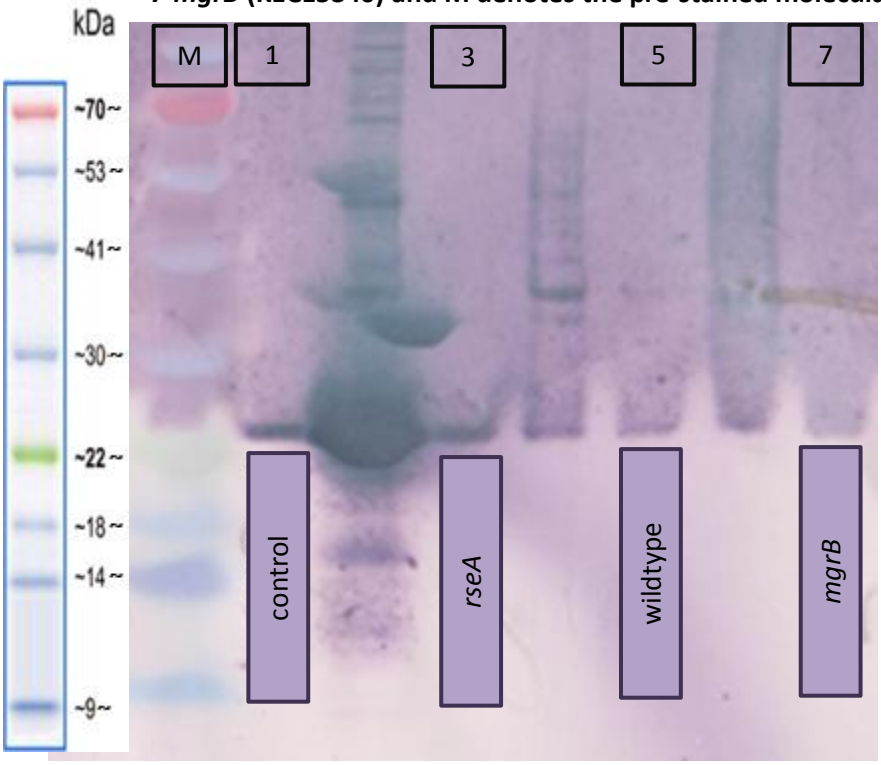
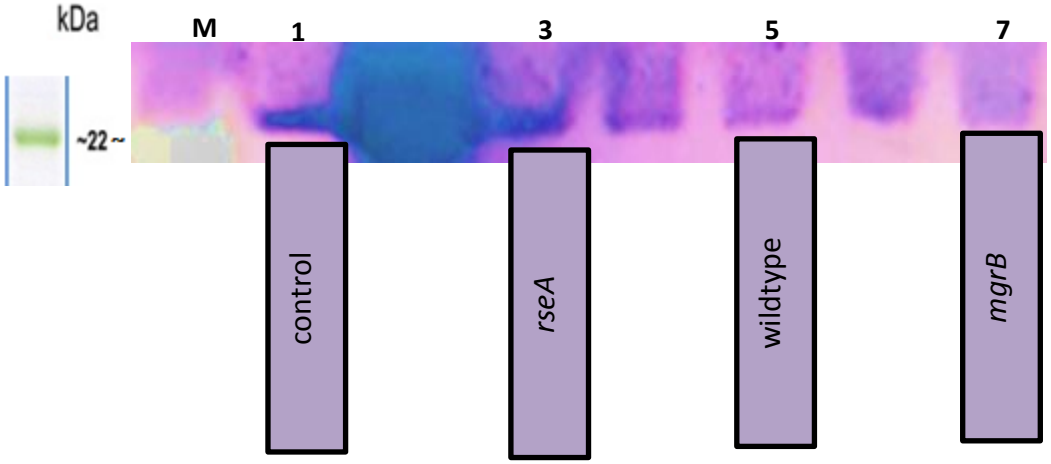


Figure 5.4B. Colour-corrected gel (400% saturation) of 24kDa band of Figure 5.4A. A colorimetric Western blot image using a sigmaE polyclonal antibody (donated by the Gross lab). Run in a Bis-Tris 10% MOPS buffer. Showing the sigmaE concentration in four samples, lane 1-10ug/ml of sigmaE protein(donated by the Ades lab), 3- (KCEC4420), 5-wildtype (CAG45115), 7- *mgrB* (KECE5340) and and M denotes the pre-stained molecular weight marker.



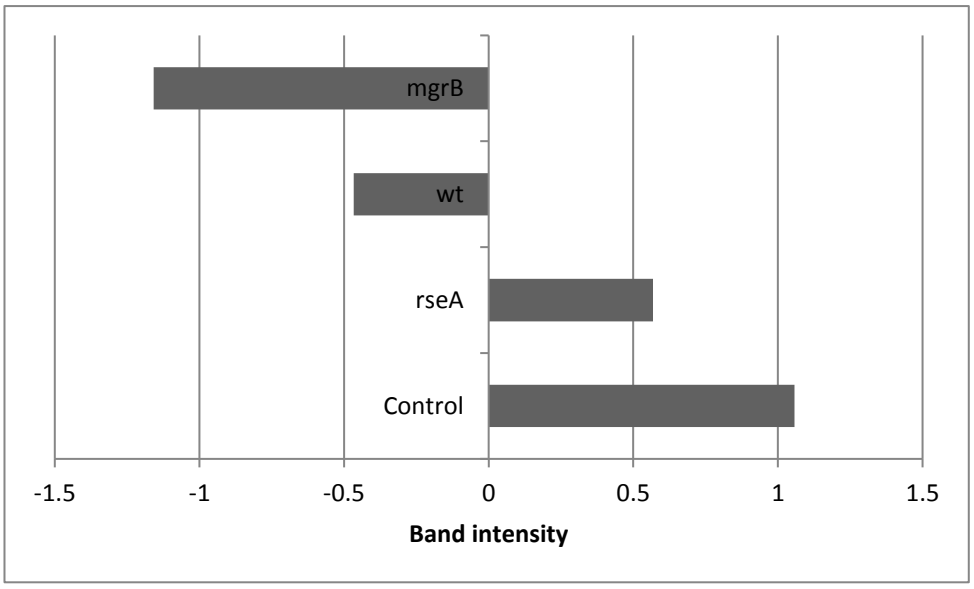


Figure 5.4C. Z-score of the relative intensity of chemiluminescent bands observed from blot, Figure 5.3A, generated using the standardise function of the Excel software package. The population mean and standard deviation of these values are 19985.88 and 9134.86, respectively .

5.4 PhoP influences sigmaE activity independent of RseA

Unexpectedly, the increase in sigmaE activity was not corroborated when sigmaE protein concentration was measured directly (5.3). To further characterise PhoP's regulatory relationship with sigmaE, sigmaE activity was measured in an *rseA*,*phoP* double mutant. The aim of this experiment was to try and elucidate the mechanism by which PhoP's regulation of sigmaE occurs. RseA is the chief regulator of sigmaE activity, and so it would be expected that sigmaE activity in an *rseA*,*phoP* double mutant would be epistatic with an *rseA* mutant, which would help classify PhoP's place in sigmaE's regulatory hierarchy.

A *phoP* mutation (JW1116-1) was transduced into an *rseA*, *rpoH*P3-*lacZ* strain (KCEC4418) under standard conditions. This was possible because *rseA* was previously found to be a suppressor of the *phoP* transduction defect. A β-galactosidase assay was carried out using the created *rpoHP*3-*lac*Z,*rseA,phoP* double mutant (Figure 5.5).

These data suggest that PhoP regulates sigmaE activity regardless of whether or not it is sequestered by RseA, which might suggest that PhoP is a key regulator of sigmaE activity. However, the previous experiment suggests that PhoP causes sigmaE protein concentration to fall. It is possible that PhoP directly causes an increase in sigmaE activity by increasing *rpoE* expression, while indirectly causing sigmaE total protein concentration to fall via an unknown FX factor.

The assay shows that there is a slightly higher level of sigmaE activity in the *rseA,phoP* mutant, compared with the *phoP* mutant. This may be due to suppressor mutations, in a clean mutant the difference in sigmaE activity between *rseA,phoP* and *phoP* mutants could be insignificant. Suppressor mutations are expected to arise from increased envelope stress and lack of sigmaE activity in a PhoP mutant. However, it is unknown at what level of sigmaE activity suppressor mutations become prevalent, so they may or may not be expected in an *rseA,phoP* mutant. If there are no suppressor mutations in the *rseA,phoP* strain, the increase in sigmaE activity compared to the *phoP* mutant would be the result of another regulator that activates sigmaE in independent of *phoP* and *rseA*.

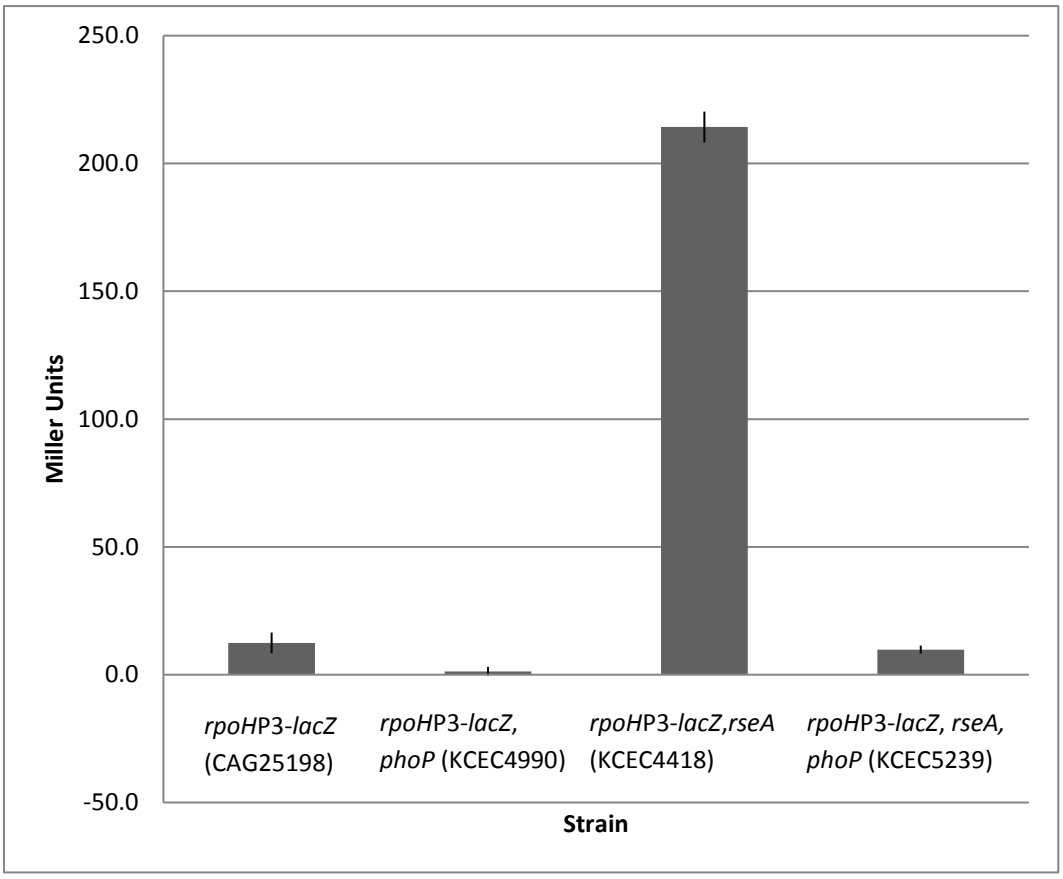


Figure 5.5. β-galactosidase assay showing the effect of a *rseA, phoP* mutants on *rpoH*P3 expression. Error bars show 95% confidence intervals (n=4). Wildtype- CAG45114. *rpoH*P3 –*lacZ,phoP* (transduced at 28°C) - KCEC5127, *rpoH*P3 –*lacZ*,*rseA*- KCEC4418, *rpoH*P3 –*lacZ rseA,phoP* -KCEC5239.

5.5. Searching for a PhoP-Box in the *rpoE* promoter region, using *in silico* analysis

It has been suggested that PhoP regulates sigmaE both positively and negatively, albeit through different mechanisms. One of these proposed mechanisms is direct transcriptional regulation of *rpoE* expression. Many *in vitro* and *in vivo* assays are used to detect transcription factor-binding sites, such as an electro-Mobility Shift Assay, ChIP-on-chip or promoter analysis. The advantages and disadvantages of these experiments depend on several factors, such as if the TF binding motif is known and the scope of the analysis. *in silico* experiments can be used to supplement *in vitro* and *in vivo* motif discovery assays by analysis and extrapolation.

PhoP binds to the consensus sequence (T/G)GTTTAnnnnn (T/G)GTTTA)[96]. Twenty-six genes have been found that contain this hexanucletoide repeat with a 5-nucleotide spacer, within 500bp upstream of the start codon of the gene[96]. It was thought that by searching and finding possible PhoP binding sites associated with *rpoE* it would support the hypothesis that PhoP is a regulator of *rpoE* expression and help validate the existence of at least one regulatory arm of PhoP's proposed duel regulatory relationship with sigmaE. To further test the hypothesis that PhoP regulates *rpoE* expression, a comprehensive search was done for PhoP-box sequences, using a position weight matrix (PWM).

In a consensus sequence, each position in the sequence is represented by a single character, which has the advantage of expressing a good deal of information as a simple string. However, although specific characters can be used to represent degeneracy, i.e. two or more bases, the sequence cannot be truly representative of the motifs used to create it, because information such as the frequency of nucleotides in each position is lost. A readily used alternative is to represent a set of aligned transcription factor binding motifs using a position weight matrix that shows which nucleotides are present – and most prevalent – in each position of a sequence and which positions are conserved.

A PWM is a probabilistic model used to characterise and predict transcription factor binding motifs[232]. PWMs are generated from aligned sequences that are determined experimentally. Initially, aligned sequences from the PhoP Mg^{2+} stimulon were used as defined by Minagawa et al. (2003) (Appendix Table D.1)[42]. First, a position frequency matrix (PFM) was constructed, which is simply a count of each nucleotide in each position in the sequence alignment (Figure 5.6A). These data can also be represented graphically as a sequence logo (Figure 5.6B).

	Nucleotide frequency			
position	A	C	G	T
1	0	0	2	7
2	1	0	8	0
3	0	0	1	8
4	0	0	0	9
5	0	0	0	9
6	8	0	1	0
7	1	0	4	4
8	2	2	4	1
9	3	1	3	2
10	3	0	2	4
11	1	3	1	4
12	0	0	4	5
13	2	0	7	0
14	0	0	0	9
15	0	0	1	8
16	0	0	2	7
17	9	0	0	0

Figure 5.6A. Frequency matrix generated from the PhoP-box consensus sequences of the genes in the Mg^{2+} stimulon, using UGENE software, based on Minagawa et al. (2003)[42].

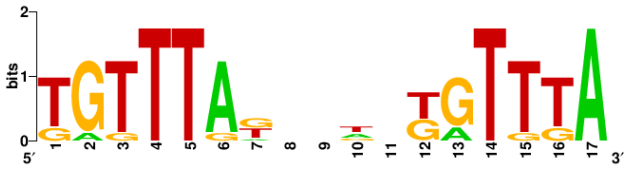


Figure 5.6B. Position weight matrix represented graphically as a sequence logo, created using the PFM (Figure 5.6A) and UGENE software.

Converting a PFM to a PWM is calculated by using the following equation[233]:

$$W_{b,i} = \log_2 \frac{p(b,i)}{p(b)}$$

$p(b)$ = background probability of base b; $p(b,i)$ = corrected probability of base b in position i; $W_{b,i}$ = PWM value of base b in position i

This equation generates normalised nucleotide frequencies, which are then converted to log values. A PWM was constructed by inputting the PhoP-box consensus sequences of members of the Mg^{2+} stimulon into Unipro UGENE software (Figure 5.7)[234].

	Nucleotide frequency			
Position	A	C	G	T
1	-2.64297	0.360729	-0.00391	0.329859
2	-1.03353	-1.8365	1.13316	-1.67162
3	-1.03353	0.360729	-1.22769	0.525604
4	-0.44575	-1.8365	-2.83713	0.689233
5	-2.64297	0.360729	-1.22769	0.610762
6	1.248849	-0.22706	-0.6399	-3.28106
7	-1.03353	-1.8365	0.65938	-0.06218
8	0.190242	0.996718	0.207395	-1.08383
9	0.401551	0.360729	-0.00391	-0.44785
10	0.190242	-0.22706	0.207395	-0.23654
11	-1.03353	1.20803	0.207395	-0.44785
12	-1.03353	0.728454	-0.00391	0.086237
13	0.190242	-0.22706	0.876445	-3.28106
14	-0.44575	-1.8365	-2.83713	0.689233
15	-0.07802	-1.8365	-1.22769	0.525604
16	-2.64297	0.360729	0.207395	0.215449
17	1.4679	-1.8365	-2.83713	-3.28106

Figure 5.7. Position weight matrix generated using UGENE software from the Mg^{2+} stimulon PhoP-box consensus sequences (Minagawa et al. 2003[42]), represented by PWM

5.5.1 Optimising the PFM threshold for consensus sequence prediction

Many bioinformatics tools predict TF binding sites from known, experimentally, derived binding motifs that can be inputted in the form of a number of aligned sequences or scoring matrices such as PWMs. For the analysis below, 'find individual motif occurrences' (FIMO), which is part of the MEME suite, was chosen because it was evaluated by Jayaram et al. and found to be one of the most comprehensive tools for identifying individual TF binding sites[235]. The MEME (Multiple EM for Motif Elicitation) suite is a software toolkit offering a number of different analyses, such as motif-sequence database searching, motif discovery and motif database searching. Many of these tools and analyses, such as FIMO, utilise the MEME algorithm which uses a probabilistic method with expectation-maximization[235]. One or several motifs can be input into FIMO as either a consensus sequence or a scoring matrix, following which the program calculates a log-likelihood ratio score for each motif with respect to each sequence position and background frequency and then convert them to P-values[236]. FIMO will detect all motifs with a P-value smaller than a chosen value, because P-values directly correlated to scores, they function as a scoring threshold.

FIMO offers a number of different thresholds that can be chosen(Table 5.1). To choose a reasonable threshold for PhoP binding site discovery, the performance of a number of different P-value thresholds, <1.0E+00, <1.0E-01, <1.0E-02, <1.0E-03, <1.0E-04, and <1.0E-05, was evaluated using a confusion matrix, which uses test data to the evaluate a classifier by a number of different criteria. The test datasets used were experimentally derived binding sites to which the *in silico* predications could be compared. However, few PhoP binding sites have been determined using *in vivo* assays, and many are based on *in silico* predictions[96]. Discounting the original eight motifs, from genes constituting the Mg^{2+} stimulon, three others were found in the published literature, namely *crcA*- TATTAAggttaTGTTAA[98], *treR*-

GGTTTATcgttGGTTTAG[237] and *ybjG*- TCTTTAagtttTATTTA[98], which were found via either a promoter or ChIP analysis. To generate a test dataset, the gene sequence 600bp – upstream of the ATG startcodon of the genes *crcA, treR* and *ybjG* – was analysed with FIMO, using the least specific threshold P-value <1.0. Any motif that matched at least 75%, of experimentally derived, PhoP-Box was considered a correct match. For each of the three genes, *crcA*,*treR* and *ybjG* the analysis correctly identified the experimentally derived PhoP binding motif and a number of false ones(Appendix Tables D.2-D.4). Using this, a confusion matrix was plotted to evaluate the performance of the six P-value thresholds.

Confusion matrices allow the comparison of the rate of true positives, true negatives, false positives and false negatives to evaluate the error rate, accuracy, sensitivity, specificity, precision and false positives rate of a classifier, in this case different scoring thresholds[238]. True positives (TPs) are instances in which the confusion matrix predicts a positive case correctly . True negatives (TNs) are instances in which a negative predicts correctly, while false positives (FPs) are predicted positive cases that are negative. Finally, false negatives (FNs) are predicted negative cases that are positive. To generate TP, TN, FP and FN predictions were made for each threshold using the test data (Table 5.1). For example, the PhoP binding site for *ybjG* was identified with a P-value of 2.02E-05, which would fall into the threshold of a P-value <1.00E-4. At a threshold of a P-value <1.00E-5, this motif would not have been found making it a FN at <1.00E-5 but a TP for all other thresholds lower than a P-value <1.00E-5.

P-value	TP	TN	FP	FN
<1.0E+00	3	0	30	0
<1.0E-01	3	16	14	0
<1.0E-02	3	17	13	0
<1.0E-03	3	26	4	0
<1.0E-04	2	30	0	1
<1.0E-05	0	30	0	3

Table 5.1. A summary of a confusion matrix plotted showing the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) for each threshold value, using data generated from *crcA, treR* and *ybjG* (full data is available in Appendix Table D.5).

Using these predictions (Table 5.1), the error rate, accuracy, sensitivity, specificity, precision and false positives rate can be defined for each threshold value. The error rate is a number between zero and one, which represents the proportion of incorrect predictions at each threshold. It can be calculated by dividing the number of false predictions (FN and FP) by the total number of predictions (1):

$$\text{Error Rate} = \frac{\text{FP} + \text{FN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \quad (1)$$

The accuracy is the proportion of correct predictions. It is the inverse of the error rate and is represented by a number between zero and one, where the higher numbers represent greater accuracy. It can be calculated by either dividing the total positive predictions (TP and TN) by the total number of predictions (2) or by simply subtracting the error rate from 1:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \quad (2)$$

Sensitivity, also known as recall or the true positive rate, is the proportion of TP that has been identified correctly. It is represented by a number between zero and one, wherein numbers closer to one represent higher sensitivity. It is calculated by dividing the number of correctly predicted positives (TP) by the total number of positives (TP and FN) (3):

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (3)$$

Specificity, also called the true negative rate, is the proportion of the number of correctly identified negative results. The most desirable specificity is one, while the least desirable is zero. It is calculated by dividing the number of TN by total predicted negatives (TN+FP) (4).

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \qquad (4)$$

Precision, also called the positive predictive value, is the proportion of predicted positives that are true positives. It is represented by a number between zero and one, wherein numbers closer to one signify greater precision. It is calculated by dividing the TP by the total number of positive predictions (TP+FP) (5):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \qquad (5)$$

The false positive rate is the inverse of specificity. It is the proportion of negatives incorrectly identified as positives and can be calculated by dividing the number of FP by total number of negatives (TN+FP) (6), or more simply by subtracting 1 from specificity:

$$\text{False positive rate} = \frac{FP}{TN + FP} = \quad (6)$$

The <1.0E-00 and <1.0E-05 thresholds were the weakest, with the former having high accuracy but a large false positive rate, while <1.0E-05 detected none of the experimentally derived binding motifs. The <1.0E-01 and <1.0E-02 thresholds are marginally better than the <1.0E-00 threshold, due to a lower false positive rate, but it is still higher than the three remaining thresholds (Figure 5.8).

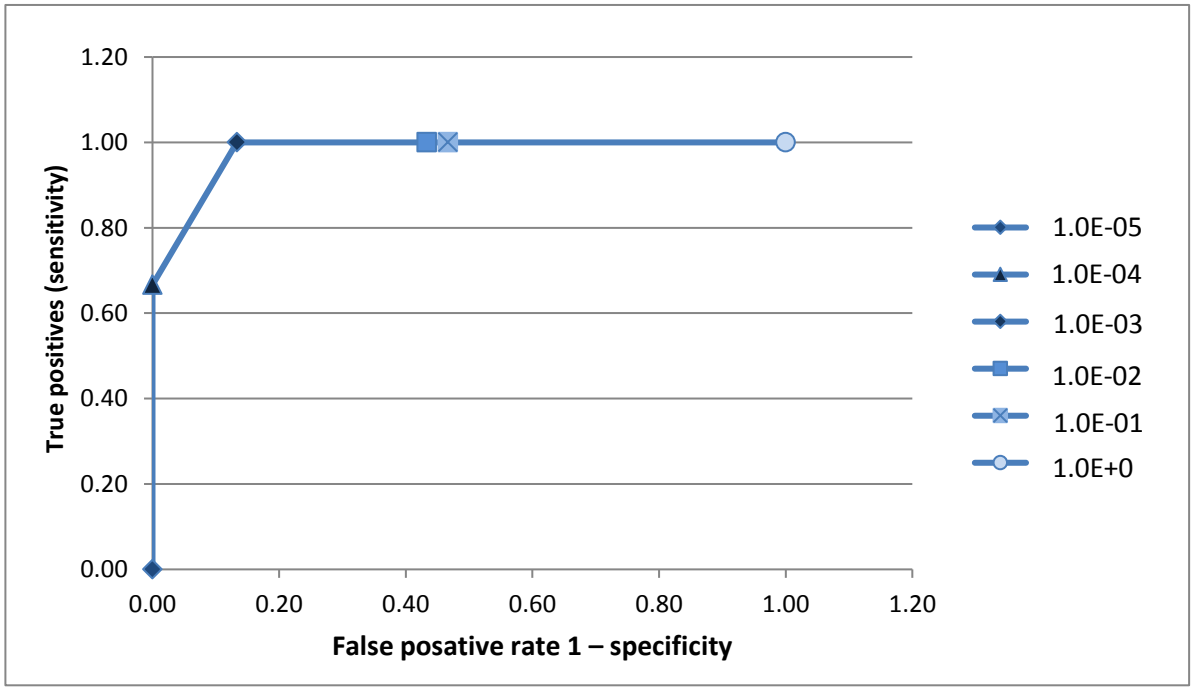


Figure 5.8. Receiver Operating Characteristic (ROC) curve analysis showing the false positive rate and the true positive rate for each threshold value.

The <1.0E-04 and <1.0E-03 thresholds are the strongest, but each offers distinct advantages and disadvantages. For instance, <1.0E-03 is more likely to return a correct motif but many false ones, while <1.0E-04 is less likely to find either a true or a false motif. However, the <1.0E-04 threshold was chosen, as it seemed to offer the most advantages over the others, namely that although it had a comparatively low sensitivity, it was more accurate and had much higher precision than the other thresholds (Figure 5.9).

P-value	Error Rate	Accuracy	Sensitivity	Specificity	False positive rate	Precision
<1.0E+00	0.91	0.09	1	0	1	0.09
<1.0E-01	0.42	0.58	1	0.53	0.47	0.18
<1.0E-02	0.39	0.61	1	0.57	0.43	0.19
<1.0E-03	0.12	0.88	1	0.87	0.13	0.43
<1.0E-04	0.03	0.97	0.67	1	0	1
<1.0E-05	0.09	0.91	0	1	0	

Figure 5.9. Error rate, accuracy, sensitivity, specificity, precision and false positives rate for each threshold value expressed as a heat map.

5.5.2 Searching for PhoP binding sites in the *rpoE* promoter region

Using the PWM, generated from PhoP binding site motifs, and the P-value <1.0E-04 threshold, FIMO was used to search the promoter region (600bp upstream of the ATG startcodon) of *rpoE*, in order to test the hypothesis that PhoP regulates *rpoE* transcription. This analysis returned three potential binding sites (Table 5.2). These motif scores had a range of 3.77477-1.04505, and although these scores seemed high, inferring high sequence similarity with known PhoP binding sites, it was prudent to analyse their significance.

#	Score	P-value	
1	3.77477	0.000197	Tgttcagattctgtaga
2	2.01802	0.000339	ggtttggggagacttta
3	-1.04505	0.000829	tgggcataaaatgttga

Table 5.2. PhoP binding motifs in the *rpoE* promoter predicted FIMO analysis showing the P-value and score.

To determine the statistical significance of these scores, 1,000 random sequences 600bp in length were generated with the same distribution of nucleotides as the *rpoE* promoter region (600bp upstream from ATP start codon), i.e. A-28.4% T-27.1% G-24.5% C-20%, using the software rMotifGen[239]. These 1,000 sequences were then analysed with FIMO, using the lowest threshold, namely P<1.0, to generate the largest number of possible scores. This analysis yielded 72,603 possible motif occurrences, with scores ranging between 17.3604 and -22.4955(Data not shown).

The probability distribution of these data was then calculated using the probability distribution function in the Minitab software package, assuming the data had a largest extreme value distribution, this was then plotted graphically (Figure 5.10). From this data a confidence interval could be derived. An α-value of 0.05 requires a score of at least 2.9279 to be significant, which suggests the motif with a score of 3.77477 was significant.

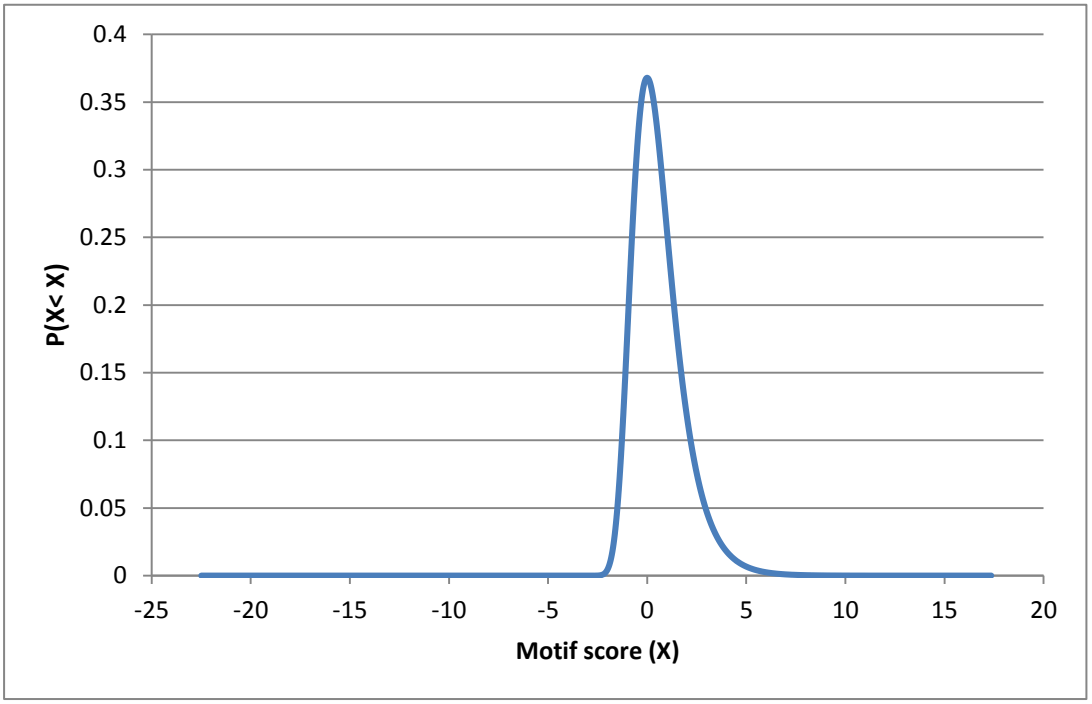


Figure 5.10. Probability distribution plot F(x), assuming a largest extreme value, of 72603 FIMO-generated scores and the associated P-value.

Although only one motif had a significant score, it was thought practical to access the biological significance of all three, with regard to the *rpoE* gene. Six promoters have been associated with *rpoE*, (P1 (-381), P2 (-327), P3 (-327/-326), P4 (-218), P5 (-153) and P6(-78))[229].

These promoters respond to different exogenous and endogenous signals, such as LPS defects, starvation (entry into stationary phase), osmolality stress, nitrogen starvation and overexpression of lipoproteins[229]. P2 and P3 share the same transcription start site. P2 has a sigmaN consensus sequence and a GlrR binding site. Both of these act as positive regulators of P2. GlrR is part of the GlrK TCS, which is associated with glucosamine-6-phosphate metabolism, and it positively regulates GlmY, which is a positive regulator of *glmS*. P3 has a sigma70 consensus sequence and is also thought to be regulated by RcsB and CRP and activated in response to the mislocalisation of lipoproteins such as YhdV, YghG, Spr and YceB[229]. RcsB is a response regulator and part of a multi-component phosphorelay system along with RcsF/RcsC/RcsD/RcsA[240]. P4 has a sigmaS consensus sequence, thereby suggesting an *rpoE* transcription increase upon entry into the stationary phase; however, this sequence is also consistent with many sigma70 consensuses[229]. P6 is autoregulated by sigmaE[229]. P1 and P5 have not been associated with any transcription factors and no consensus sequences have been predicated[229].

The potential PhoP binding sites that were predicted were compared to the *rpoE* promoters (Figure 5.11). Motif 2 is downstream of P2 and P3, and motif 3 is close to the ATG start codon. However, one possible PhoP binding site (motif 1) is within the P5 region – this motif had the highest score and the only one of statistical significance based on the analysis conducted above. The regulator of P5 is unknown, but these data implicate PhoP as a regulator for this promoter.

Figure 5.11. A summary of the *rpoE* promoter region -499bp upstream of the ATG start codon. *rpoE* promoters (P1-P6)[229] **are highlighted along with three potential PhoP binding sites (green) found by *in silico* analysis.**

5.6 SigmaS positively regulates sigmaE activity, independent of RseA

A possible mechanism by which PhoP directly regulates *rpoE* expression was suggested in the previous experiments. However, it has been hypothesised that PhoP must regulate sigmaE through a number of different mechanisms, as it has been observed that PhoP increase sigmaE activity while leaving protein concentration reduced. It is likely that these mechanisms are separate, i.e. one direct and one indirect. A possible candidate for indirect regulation is sigmaS, as it has been suggested by Klein et al. that it regulates *rpoE*P4 expression [229],[102]. To explore this association, an *rpoH*P3-*lacZ rpoS,rseA* mutant (KCEC5354) was constructed. This strain was subject to a β-galactosidase assay (Figure 5.12).

The data show that the wildtype and *rpoS* mutant are epistatic, *rpoS* only affects sigmaE activity in the absence of *rseA*. The effect of an *rseA* mutation on sigmaE activity was diminished in the absence of sigmaS, which mirrors the results found above (Figure 5.5). These data suggest that *rseA*'s effect on sigmaE activity requires sigmaS, inferring that sigmaS enhances the expression of *rpoE*, a novel observation when this assay was carried out. However, there is no transduction defect associated with *rpoS*, which suggests that sigmaS is not the only mechanism by which PhoP regulates sigmaE.

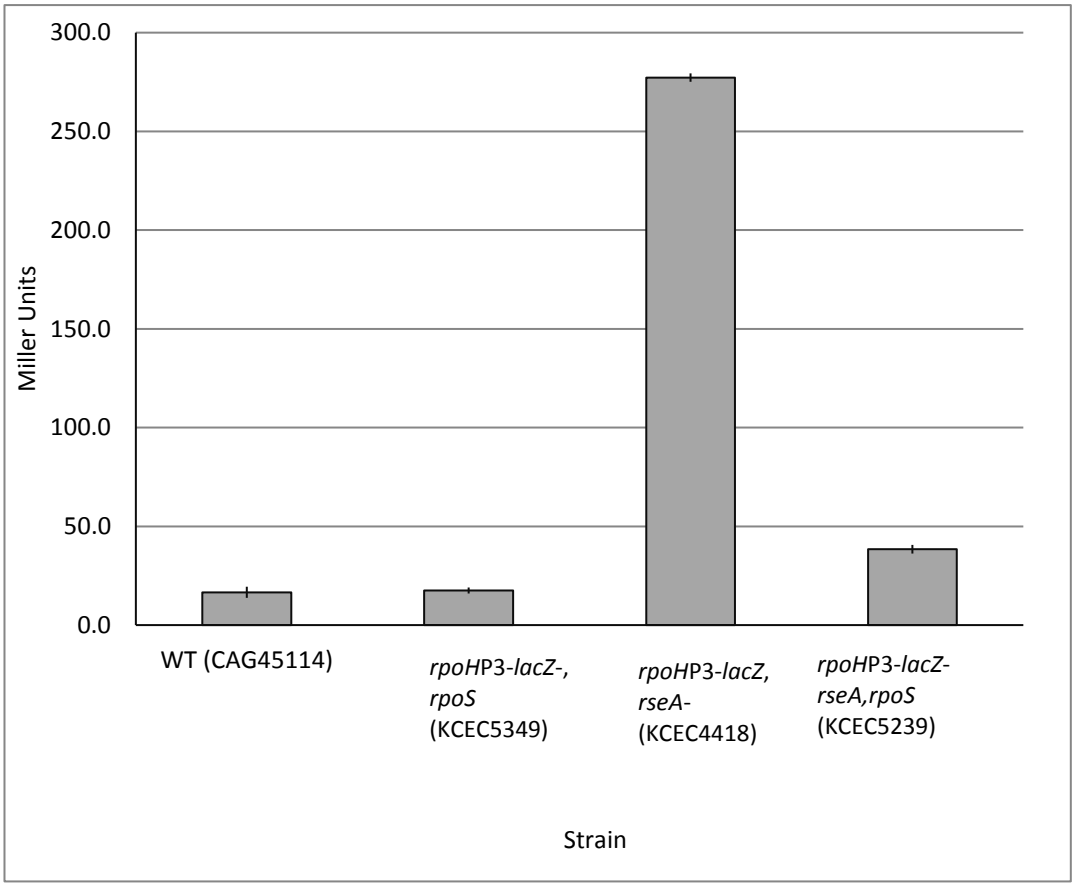


Figure 5.12. β-galactosidase assay showing the effect of *rpoS* mutants on rpoHP3 expression. Error bars show 95% confidence intervals (n=4). Strain wildtype *rpoH*P3-*lacZ*- CAG45114, *rpoH*P3-*lacZ*-, *rpoS* - KCEC5349, *rpoH*P3-*lacZ*, *rseA*- KCEC4418, *rpoH*P3-*lacZ*- *rseA,rpoS* KCEC5239

5.6.1 *rpoS* mutation acts as an anti-suppressor for the *phoP* Transduction defect

To further explore the extent to which sigmaS contributes to the PhoPs, regulation of sigmaE an attempt was made to construct an *rpoS,rseA,phoP* triple mutant via P1 transduction under standard conditions.

Previously, it had been found that *rseA* was a suppressor of the *phoP* transduction defect (chapter 4). It was hypothesised that if *rpoS* was part of the regulatory mechanism by which PhoP enhanced sigmaE activity, an *rseA,rpoS* mutant would not suppress the *phoP* transduction defect. An *rseA* mutation (CAG25198) was transduced into an *rpoS* mutant under standard conditions (KCEC5354), following which an attempt was made to transduce a *phoP* (JW1116-1) mutation into this background, using *phoQ* (JW1115-1) as a control. Transduction into an *rseA* mutant was used as a secondary control, and two biological repeats were done for each condition. This was repeated twice (Table 5.3).

It was not possible to transduce the *phoP* mutation successfully into an *rseA,rpoS* background, which suggests that without sigmaS, sigmaE can no longer reach activities high enough to suppress the transduction defect. This supports the findings of the previous assay, sigmaS enhances sigmaE activity in an *rseA* mutant. Moreover, these data infer that sigmaS is not the only route by which PhoP regulates sigmaE, as it would be expected that if it were then transduction of a *phoP* mutation would have been possible as it is with an *rseA* mutant. Although, this assertion cannot be made with confidence as PhoP is not the only activator of sigmaE. It may be possible that in the absence of *phoP* sigmaS still retains sufficient activity to allow transduction of a *phoP* mutation into an *rseA* mutant. Additionally, sigmaS and sigmaE are important for cellular function, and in the *rseA,rpoS* mutant, the function of both of these alternative sigma factors is disrupted, it is possible that the observation is the result of an accumulation of deleterious effects that are independent of the *phoP* transduction defect.

Additionally, this experiment strengthens the initial hypothesis that sigmaE suppresses the transduction defect, as it was possible to transduce *phoP* into an *rseA* mutant, wherein sigmaE is hyper-induced, but no longer possible in an *rseA*,*rpoS* strain where sigmaE activity is inhibited.

Strain	**Average number of *phoP* transductants**			**Average number of *phoQ* transductants**		
	Repeat 1	Repeat 2	Average	Repeat 1	Repeat 2	Average
rseA(KCEC4418)	11,14	14,13	13	7,11	8,10	9
rseA,rpoS (KCEC5354)	0,0	0,0	0	2,3	1,2	2
Wildtype (CAG45114)	0,0	0,0	0	7,8	6,7	7

Table 5.3. Showing the average number of *phoPQ* transductants transduced into the wildtype (CAG45114), *rseA, rpoS* (KCEC5354) and *rseA* (KCEC4418) over two experiments.

5.7 Discussion

It was hypothesised that PhoPQ is a positive regulator of sigmaE activity, this was strongly supported by several assays that showed a strong correlation between PhoP induction and sigmaE activity. It was observed that the loss of the *phoP* inhibitor MgrB led to an increased expression of sigmaE activity (5.2). Moreover, *micA* expression increased when *phoP* expression was induced from the *lac*UV5 promoter using IPTG. Presumably, this increase in activity was caused by elevated sigmaE activity, dependent on PhoP. Furthermore, a model in which PhoP is a key regulator of sigmaE provides a cogent explanation for the observation that a *micA* mutant could act as a recipient strain for the transduction of *phoP* under otherwise standard conditions (4.4.1). It is likely the strain was pre-adapted because in a *micA* mutant, sigmaE activity would be expected to be higher due to a defective negative feedback loop between sigmaE, MicA and PhoP.

Interestingly *phoQ* showed a lower incidence of sigmaE activity compared with *phoP*(5.2). However, no transduction defect associated with *phoQ* was found, which may suggest that the transduction defect is not associated with sigmaE, albeit this seems unlikely. As discussed previously, it is more likely that differences in sigmaE activity in the *phoP* and *phoQ* mutants are due to secondary suppressor mutations in the *phoP* mutant strain. However, if this is true and secondary suppressor mutations are a result of losing sigmaE, then it should logically follow that *phoQ* mutants must also contain secondary suppressor mutations. Alternatively, lack of suppressor mutations in *phoQ* mutant would support the hypothesis that PhoP regulates sigmaE independently of PhoQ, by either having activity in its unphosphorylated state or being activated by another sensor kinase. It is possible that in the absence of PhoQ, PhoP is still able to maintain limited regulatory control of sigmaE. This small concentration of active PhoP, in *phoQ* mutants, might be sufficient to yield transductants with fewer secondary suppressor mutations, which may also explain why there is no transduction defect in *phoQ* mutants.

In this chapter two possible mechanisms by which PhoP may regulate sigmaE were explored, directly by regulating *rpoE* expression and indirectly via sigmaE. Considering it was shown that PhoPQ has a stimulatory effect on sigmaE activity (5.2, 5.2.1 and 5.2.2), it was surprising that increased *phoP* expression seemed to have a negative effect on sigmaE concentration (5.3). It is possible that sigmaE was affected by feedback inhibition; however, this seems unlikely. Firstly the Western blot assay showed that in an *rseA* mutant sigmaE protein concentration was increased compared to the wildtype. Secondly, It has been shown (5.2.1) how sigmaE activity rises, seemingly, without evidence for a feedback mechanism. However, because of several weakness in the Western blot data, the conclusions discussed above, namely that PhoP has a negative effect on total sigmaE protein concentration, may be questionable.

One of the most apparent weaknesses of the Western blot data presented in this chapter is the lack of a protein loading control. Protein loading controls help demonstrate that the difference in protein concentration between samples is not due to any variance in the amount of sample loaded or in total protein

concentration between samples. Moreover, protein loading controls help ensure proteins have been transferred from gel to membrane efficiently and that signal detection is uniform across the gel. Proteins that are constitutively expressed at a high level and have a different molecular weight than the protein of interest are typically chosen[241].

Although care was taken to measure and standardise the total protein concentration of each sample, as no protein loading control was used, the possibility that the apparent variation in sigmaE concentration between samples was an artefact of the experimental procedure cannot be definitively rejected.

Moreover, it is evident that there was a high background signal present across the Western blot, for which there are a number of different possible causes, ranging from procedural errors, such as the incomplete blocking of the membrane or insufficient washing, to contamination of the reagents or the membrane itself, although the most likely explanation is the use of a polyclonal antibody. Unlike monoclonal antibodies, polyclonal antibodies bind to multiple epitopes, which may be shared by related proteins and in turn may lead to cross-reactivity and a high background signal, the latter of which may mask small differences in protein concentration, thereby increasing the difficulty in making accurate assertions regarding variability between samples[242].

In an attempt to mitigate the effect of the background signal, the ImageJ plugin "subtract background" was used. This plugin uses a "rolling ball" algorithm, which can help to compensate for irregular intensity distribution within an image[231]. The workings of the algorithm can be understood by envisioning a three-dimensional representation of the image, with the third dimension corresponding to pixel intensity. The background is calculated by "rolling" a sphere – of a user-defined radius – across the underside of the surface. The sphere may invade peaks with larger radii than itself. The area of the peak that the sphere touches is determined as the background and removed from the image[231], [243]. Since a sphere cannot penetrate peaks with smaller or narrower radii than itself, a radius was chosen that was approximately equal to that of the bands to be analysed.

While using this process helped mitigate the effect of the high background signal, several minor bands were detected in between the sample lanes, at 35kDa and above, and one major band detected at 23kDa-28kDa could not be subtracted, because their radii and intensity were similar to those of the bands to be analysed.

While multiple bands of varying molecular weights might be expected when using a polyclonal antibody, a large band in between lanes is likely to be indicative of sample leakage into adjacent lanes. It is possible that this may have led to *rseA* and protein control, in between which the major band is located, not being representative of sigmaE protein concentrations.

In an attempt to circumvent leakage from outside of the sample and subsequent overlapping bands during quantification, the region of analysis was carefully selected. For all bands, a rectangle with the same dimensions was drawn, thus constituting the region of analysis. As stated in the protocol[230], the region of analysis should be larger than the band to be analysed, as this ensures that the band signal can be analysed in its entirety and a consistent border of quantification is maintained for all bands, even if they might be somewhat ill-defined.

When choosing the region of analysis for the protein control and *resA* bands, care was taken to avoid any major and minor bands outside of the sample lanes. To achieve this aim, a region of analysis was chosen that omitted the outermost left and right borders of the protein control and *resA* bands but still extended slightly vertically into the sample lane. Although doing this removed approximately 5% of the band from the region of analysis, it nevertheless increased the likelihood that sample leakage would not be included in the analysis of the sigmaE protein control and *rseA* bands. This was done for all bands, in order to maintain consistency.

However, if band leakage had occurred, then saturation may be expected. Band intensity is expected to be proportional to protein concentration, but at very high protein concentrations it is no longer linearly proportional to protein abundance, and accurate quantification is therefore not possible[241]. If saturation had occurred, very little variation between the protein control and the *rseA* bands would be expected. Nevertheless, the relative intensities of the protein control and

rseA bands are dissimilar, which would suggest that the samples were not saturated and increases the likelihood that the protein control and *rseA* samples are representative of protein concentration.

The data presented in this chapter infers PhoP's regulation over sigmaE is complex and multi-factorial. The results suggest a duel mechanism of action for PhoP, both lowering the protein concentration of the cell and increasing activity. It is likely that these mechanisms are both direct and indirect. Direct regulatory control of expression was suggested by the observation that the effect of PhoP on sigmaE is independent of RseA. Further evidence of direct transcriptional regulation was inferred by the discovery of a PhoP binding site inside *rpoE*P5. Indirect regulation was observed when measuring sigmaE activity in an *rpoS,rseA* mutant, suggesting that sigmaS is one way in which PhoP exerts control of sigmaE. These ideas are explored and expanded on in the final chapter of this thesis.

Chapter 6: Discussion

6.1 Background Summary

The PhoPQ two-component system assists *E. coli* in proliferating in an Mg^{2+} starved environment. This is facilitated by PhoPQ's ability to sense and respond to falling concentrations of extracytoplasmic Mg^{2+} [76][102]. Moreover, proteins regulated indirectly and directly by PhoPQ, such as OmpT, EptB, CrcA and SigmaS, reveal a secondary role of PhoPQ, distinct from (and yet related) to its ability to modulate divalent cation, which places it as a regulatory interlinkage between OM modification and stabilisation and the general stress response. SigmaE is a key regulator of extracytoplasmic stress, and the majority of sigmaE-regulated genes function to regulate and respond to OM damage. Furthermore, sigmaE regulates PhoP post-transcriptionally via the sRNA MicA. The primary hypothesis of this work is that PhoPQ is required for full sigmaE induction, under Mg^{2+} limited conditions, which would solidify its position as a key regulator of stress.

In the initial stages of this project, it was found that *phoP* was associated with a transduction defect. With regard to PhoP's most well studied role in cellular metabolism, responding to Mg^{2+} starvation, it was initially hypothesised that this defect was the result of cytoplasmic Mg^{2+} starvation. However, it was demonstrated that no transduction defect was linked to *mgtA*, the PhoP-regulated divalent ion transporter. A body of evidence was presented that suggested that *phoP* mutants suffer from extracytoplasmic stress, which is the probable cause of the inviability of *phoP* transductants, this led to the observation that PhoP positively regulates sigmaE activity independently of the sigmaE's chief regulator RseA.

The proposed mechanism by which PhoP regulates sigmaE is complex, and multi-factorial (Figure 6.1), but evidence suggests a dual role in which PhoP has a negative effect on sigmaE protein concentration but promotes sigmaE activity. Moreover, evidence was presented to support the hypothesis of both indirect and direct regulatory effects. Although the precise mechanism by which PhoP regulates sigmaE could not be identified completely with confidence, the data present a number of possibilities.

The data suggest that PhoP positively regulates sigmaE via at least two different mechanisms. It was shown that sigmaS, a known regulator of *rpoE* expression[229], enhances sigmaE activity in the absence of *rseA*. Furthermore, it remains a possibility that it is via sigmaS that PhoP exerts *rseA*-independent control over sigmaE activity; however, because *rpoS* is not associated with a transduction defect, it is probable that PhoP exploits an alternative mechanism by which to promote sigmaE activity.

An *in silico* analysis revealed that a possible PhoP box exists inside *rpoE*P5. Support for the notion that this is a mechanism of positive regulation was provided by data demonstrating that sigmaE activity falls dramatically in both *phoP* and *phoP*,*rseA* mutants, thereby suggesting transcriptional regulation. However, it is possible that RseA-sigmaE proteolysis is dependent on PhoP, and, if true, PhoP may negativity effect *rpoE* expression. Regardless of the mechanism this work has shown that PhoP is a primary regulator of sigmaE, independent of RseA.

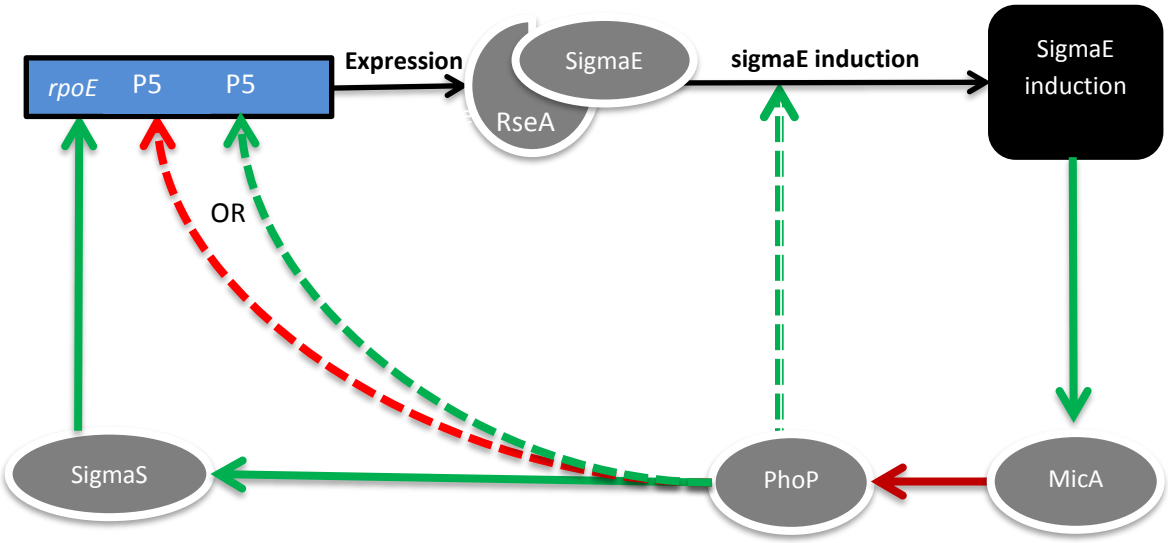


Figure 6.1, showing a proposed model in which PhoP regulates sigmaE. Red arrows represent negative regulatory relationships, green represent positive regulatory relationships. Solid arrows represent regulatory relationships from published literature, while dashed arrows represent regulatory relationships proposed by this work. Grey ovals represent proteins or mature sRNAs. The *rpoE* gene, along with promoter 5 (P5), is depicted as a blue rectangle.

6.2 *phoP* is necessary for successful P1*vir* transduction under standard conditions

It was demonstrated that *phoP* cannot be transduced efficiently or with confidence under standard conditions using P1*vir* transduction. This was surprising as no such incapacity had been reported previously. The transduction defect was found to be non-specific to the K-12 wildtype used, with both MG1655 and W3110 being susceptible. Neither was this defect confined to either of the two donor strain used to create the *phoP* lysate, suggesting it not affected by the method of creating the *phoP* deletion. In addition, the complementation of *phoP*, placing an ectopic copy of *phoP* under the control of the *lac*UV5-promoter, provided confidence that transduction of the native *phoP* gene could only be achieved when the ectopic allele was being expressed.

The most pertinent question to ask is why has no *phoP* transduction defect been reported? It is possible, as stated previously, that transduction is seldom the tool of choice for creating mutants when other, more modern, methods are available. Another reason for this may be that many *phoP* mutants are not created but obtained from other laboratories, and several highly cited papers have used the same *phoP* mutants (Appendix Table E.1). Although this is not unusual, it does preclude the possibility of encountering the *phoP* transduction defect and may be one reason why the defect is not widely reported. Nevertheless *phoP* transductants are cited in the published literature (Appendix Table E.1). This work, however, does not argue the impossibility of creating *phoP* mutants, as it has been observed that they have been created under a variety of conditions; instead, this work argues that they cannot be made without acquiring additional suppressor mutations under standard conditions.

The existence of *phoP* mutants created via P1*vir* transduction, cited in the literature, suggests a number of possibilities:

I) The *phoP* transduction defect, as defined by this work, is specious.

II) *phoP* mutants reported in the published literature were not constructed under standard conditions.

III) *phoP* mutants reported in the published literature contain suppressor mutations

The first hypothesis outlined above can be dismissed with some confidence, because although there are a number of factors that could lead to the fallacious characterisation of a transduction defect, several observations and assays conducted in this work make this possibility unlikely.

In nature, there are many bacteria that are able to resist infection by bacteriophage. Bacteriophages can be blocked from absorption into the cell, by modifying the host-cell receptor or by utilising a competitive inhibitor[244]. For example, *E. coli* can synthesise a lipoprotein, Llp, which block the T5 phage receptor, FhuA[244]. Moreover, molecules that are present or added to the host-cell environment can also be used to inhibit bacteriophage entry; for example, the antibacterial peptide microcin J25 can also be used to block FhuA and subsequent bacteriophage entry[244]. Further, some genes have a naturally low frequency of transduction, due to their locus. Some DNA sequences are poor substrates for recombination, and this can cause the frequency of gene transduction to be reduced up to 25-fold[245].

Two K-12 wildtype, MG1655 and W3310 were used as recipient strains for the transduction of a number mutations, including *phoQ*, so it is unlikely that they were both defective. Similarly, it is unlikely that the strain of P1 used herein was defective, because it was employed extensively in the creation of many strains in this work. Furthermore, it is unlikely that *phoP* has a naturally low transduction frequency, because *phoP* transductants were obtained in strains with a number of different genetic backgrounds. However, the observation that the transduction defect can be suppressed by sigmaE induction suggests that loss of *phoP* is the cause of rather than a defect in recipient, donor or viral vector.

Strain name	Origin of strain	recipient strain	recipient strain genotype
EG12976 *phoP*::Kan^R	Dissecting the PhoP regulatory network of *Escherichia coli* and *Salmonella enterica*[96]	MG1655	F$^-$, lambda$^-$, rph-1
WP3022 *phoP*::Cam^R	Identification and Molecular Characterization of the Mg^{2+} Stimulon of *Escherichia coli*[42]	W3110	F- λ- rph-1 INV(*rrnD*, *rrnE*)
MG1446 *phoP*::Kan^R	MicA sRNA links the PhoP regulon to cell envelope stress[112]	MG1173	MG1655 Δ*lac*X174 λRS*ompT-lacZ*
MP4022 *phoP*::Cam^R	Molecular Characterization of the PhoP-PhoQ Two-Component System in *Escherichia coli* K-12: Identification of Extracellular Mg^{2+} Responsive Promoters[88]	MC4100	F−Δ(argF-lac)U169 araD139 *rpsL*150 *ptsF*25 fibB5301 *rbsR deoC* relA1
FS1000 *phoP*::Kan^R	Molecular Genetic Analysis of the *Escherichia coli* phoP Locus[102]	MC1061	F- araD139 Del(araA-leu)7697 Del(lac)X74 *galK*16 *galE*15(GalS) lambda- e14- mcrA0 relA1 *rpsL*150(strR) *spoT*1 *mcrB*1 *hsdR*2
FS1002 *phoP*::Kan^R	Molecular Genetic Analysis of the *Escherichia coli* phoP Locus[102]	MC4100	F- araDJ39 A(lac)U169 *rpsL*150 relAl thiflbBS301 deoCi ptsF25 *rbsR*
EG13711 *phoP*::Cam^R	Phenotypic differences between *Salmonella* and *Escherichia coli* resulting from the disparate regulation of homologous genes[103]	EG13709	*pmrD*$^+$-*lacZY*$^+$
EG13729 *phoP*::Kan^R	Phenotypic differences between *Salmonella* and *Escherichia coli* resulting from the disparate regulation of homologous genes[103]	EG13709	*pmrD*$^+$-*lacZY*$^+$
TIM8 *phoPQ*::Cam^R	Stimulus-dependent differential regulation in the *Escherichia coli* PhoQ–PhoP system[246]	TIM80	MG1655 Δ*phoQ attHK*::[pTM27 Δ*Kan*]
KMP1 *phoP*::Cam^R	Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*[247]	KMY1	F$^-$ Δ(*argF-lac*)U169 *araD139 rpsL150 ptsF25 fibB5301 rbsR* lRS45[Φ(emrK9-lacZ)]
KMP2001 *phoP*::Cam^R	Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*[247]	KMY2001	F$^-$ Δ(*argF-lac*)U169 *araD139 rpsL150 ptsF25 fibB5301 rbsR*, lRS45[W(emrK9-lacZ)], *evgS*

Table 6.1. *phoP* transductants generated in published articles, including the recipient strain. Strain name and genome transcribed from the papers in which they were originally cited.

To investigate the validity of *phoP* mutants created by other groups, a comprehensive search was undertaken of 20 of the most highly cited papers reporting the creation of *phoP* mutants (Appendix Table E.1), and of these, a total of 28 *phoP* mutants were cited. Of these 28, 18 were created via P1*vir* transduction. However, only 11 of them were originally generated in the study in which they were cited (Table 6.1).

Nine out of eleven *phoP* transductants were transduced into derivative strains containing mutations (Table 6.1), which is especially important because in this work it was shown that *phoP* transductants can be created if the recipient strain has adapted accordingly (Chapter 4), and sigmaE induced either chemically or via mutations that cause extracytoplasmic stress. However, there are no obvious mutations within the recipient strains that would increase sigmaE activity and allow for the transduction of *phoP*.

Alternatively, it is possible that the *phoP* transductants citied in the published literature were not created under standard conditions. When searching for examples of *phoP* transductants in the published literature, the transduction methodology was seldom referenced, with only a minority of articles directly referencing the original procedure (Miller 1972)[183]. This may suggest a procedural drift between laboratories; indeed, subtle changes in procedures can be seen when looking at P1*vir* transduction protocols from different sources. For example, the procedure published by the Roa Research Group uses 10mM $CaCl_2$ and 5mM $MgSO_4$[248] to enhance bacteriophage infectivity, while a procedure published by Kranz Lab suggests using 5mM $CaCl_2$ and 100mM $MgSO_4$[249] and yet another source suggests using 50mM $CaCl_2$ only[250]. Although these differences are minor, it was shown in this work that the addition of transduction reagents, presumably $CaCl_2$ and $MgSO_4$, had a statistically significantly impact on the survivability of *phoP* mutants (Chapter 4). Therefore, it is convincible that small changes in the published methodology that may go unreported will influence the ability to create *phoP* mutants.

However, without data pertaining to the precise conditions under which *phoP* transductants, citied in the published literature, were constructed, it is difficult to expound upon their creation. Correspondingly, the hypothesis that *phoP* transductants cited in the literature contain suppressor mutations cannot be confirmed or rejected without the genomic data of these strains.

Finally, it is difficult to reconcile the observations of this work with the fact that *phoP* transductants have been generated in the past by many different groups. Nonetheless, the idea that transduction is not the favoured method of creating *phoP* mutants, coupled with the variations in procedures cited above, provided a possible means by which the *phoP* transduction defect may have gone unreported until this point.

6.3 The mechanism of the *phoP* transduction defect

PhoPQ helps mitigate Mg^{2+} starvation by positively regulating the expression of *mgtA*. Because P1*vir* transduction requires Mg^{2+} chelation it was initially thought that the *phoP* transduction defect was due to cytoplasmic Mg^{2+} starvation. It was hypothesised that *mgtA* mutants would be epistatic with *phoP*, exhibiting the same transduction defect as *phoP*. However it was found that *mgtA* could be transduced normally under standard conditions, which suggests that cytoplasmic Mg^{2+} starvation is unlikely to be the cause of the *phoP* transduction defect. Despite this finding, *phoP* mutants were found to be hypersensitive to citrate (4.2.1 & 4.2.2). Moreover, it has been shown that the transduction defect can be mitigated by sigmaE induction (4.4.1 & 4.4.2). Together these data suggests that the *phoP* transduction defect could be caused by extracytoplasmic Mg^{2+} starvation which leads to OM disruption and consequently inviability.

SigmaE is an essential protein in *E. coli*[134]. SigmaE inhibition or inactivation will lead to cell lysis and OM deformation such as 'blebbing', i.e. large bulges in the cell envelope containing cytoplasmic materials[138]. Considering that sigmaE activity is reduced in the *phoP* mutant, sigmaE activity deficiency is an appealing explanation for the transduction defect. However, extremely low SigmaE activity levels, as seen

in an *rseA,rpoS* mutant, lead to neither death nor a transduction defect (chapter 5.6). This suggests that low sigmaE activity alone is insufficient to explain the transduction defect. It is likely that *phoP* mutants are uniquely sensitive to conditions of transduction, particularly chelation and permeabilisation of the OM, due to PhoP mediated LPS modification [194], [251]. This sensitivity, presumably, results in a transduction defect when sigmaE cannot be appropriately induced to respond to the damage.

Negatively charged phosphate and carboxylate groups attached to glucosamines of lipid A, and the Kdo residue found in the inner core, are important for OM stability[36]. These negative charges are bound by divalent cations and allow LPS molecules to form tight cross-links[224]. Removing divalent cations can cause large disruptions in the LPS and the permeabilisation of the OM[207]. However, under the divalent cation-limited conditions of transduction, excessive negative charges on the LPS would presumably cause increased electrostatic repulsion between the molecules and greater OM instability. Furthermore, it has been found that treatment with EDTA can result in LPS shedding and the formation of large holes in the inner and outer leaflets of the OM[207]. It has been identified that the effects of EDTA on the OM can be abated by adding divalent cations in a concentration-dependent manner[252]. Presumably, then, increased LPS shedding would be observed if the electrostatic repulsion between LPS molecules were to increase by removing divalent cations and increasing the negative charges associated with the LPS.

Yet, permeabilisers, namely compounds that disrupt the LPS and OM, are not known to be bactericidal in the concentrations used in transduction[207]. Further, it is true that cells are often exposed to citrate and other divalent metal ion chelators and remain viable. Moreover, it has also been shown in this work that an *eptB* mutant is not able to suppress the *phoP* transduction defect, nor is there a transduction defect associated with *mgrR*, *eptB*'s negative regulator. This suggests that LPS electrostatic repulsion caused by unregulated EptB, alone, is not sufficient to cause the transduction defect. However, it seems probable that in a *phoP* mutant, in which, EptB is not inhibited by MgrR, increased LPS shedding and OM

permeabilisation would be observed. Many OMPs rely on LPS for insertion and trimersation, LPS disruption will cause the accumulation of unfolded OMP and OMP intermediates[166]. Therefore it is unsurprising that the cell will respond to LPS shedding by inducing sigmaE. This will mitigate the effect of misfolded OMPs and increased LPS synthesis[252].

SigmaE regulates approximately 90 genes[166], which can be divided into core and extended regulons. The core regulon consists of 23 proteins, 20 of which have known functions, and the majority of these genes contribute to the synthesis and correct assembly of LPS and OMPs[166]. *lpxABD*, *plsB* and *bacA* are core sigmaE regulon genes that contribute to LPS synthesis and assembly[166]. So, while it would be expected that under normal conditions the cell is able to respond to the effect of LPS shedding and OM disruption by increasing sigmaE activity, this is not possible in a *phoP* mutant and it is likely that this is what causes the inability to transduce *phoP* under standard conditions.

The hypothesis, that a disruption of LPS assembly and synthesis causes *phoP* transduction defect is supported by the observation that *phoP* mutants can be cultivated at low temperatures. Both sigmaE activity and LPS synthesis is increases at lower temperatures[219]. Moreover, as the temperature drops the cell membrane will go through a process of homeoviscous adaption in which membrane composition and fluidity are altered in response to the environment[253]. Phosphatidylethanolamine constitutes the largest proportion of phospholipid in the OM, followed by phosphatidylglycerol and diphosphatidylglycerol[253]. The ratio of these phospholipid changes in response to lower temperatures, increasing the concentration of unsaturated fatty acid found in phospholipids, which in turn increases membrane fluidity[253]. The concentration of unsaturated fatty acid can affect the synthesis of LPS[254]. LpxK is an integral membrane protein and necessary for the synthesis of lipid A[254], and lower membrane fluidity can increase its activity[254].

Another hypothesis, which could not be rejected by this work, is that *phoP* mutants may be sensitive to kanamycin. Both chloramphenicol and kanamycin interfere with

protein synthesis; however, while chloramphenicol inhibits protein synthesis by inhibiting the action of peptidyl transferase[255], kanamycin interacts directly with 30s ribosome, causing an increase in mistranslated proteins[256]. It is possible that the diminished sigmaE activity found in *phoP* mutants causes the cell additional stress, as the cell is not able to respond to the increased concentration of mistranslated OMPs, thereby providing a cogent hypothesis for the hyposensitivity of *phoP* mutants to kanamycin. The creation of a *phoP*:Cam^R mutant, via lambda red, may support this hypothesis.

6.4 The Absence of a *phoQ* transduction defect

If the *phoP* transduction defect, as presented by this work, is genuine there can be two hypotheses to explain the lack of a *phoQ* transduction:

i) There is a non-apparent transduction defect associated with *phoQ*
ii) PhoP is active in the absence of *phoQ*

The observations of this work suggest that *phoQ* can be transduced normally under standard conditions (3.2). So, the former hypothesis can be dismissed with some confidence. Conversely, the literature supports a model in which *phoP* is active in its unphosphorylated state or that it can be activated by alternative non-cognate sensor kinase. Therefore, the latter hypothesis is anticipated by this work to be the most likely.

Several studies have shown that PhoP activity is independent of PhoQ, in *S. enterica*[189], [257]. Lejona et al. demonstrated that unphosphorylated PhoP could activate various targets in a concentration-dependent manner[257]. It is possible that this is also true for *E. coli*, but there is nothing in the published literature to support this assertion, aside from the similarity of the PhoP protein in both organisms.

Alternatively, it is possible that PhoP is phosphorylated by a non-cognate sensor kinase. Yamamoto et al. investigated incidence of the trans-phosphorylation for the majority of response regulators in *E. coli*, and evidence was found that PhoP could

be activated by a non-cognate sensor kinase[190]. The authors asserted that PhoP enhances the dephosphorylate of non-cognate sensor kinase EnvZ, which is important, because it may suggest that EnvZ has the ability to phosphorylate and activate PhoP *in vivo*. Finally, many papers have characterised the cross regulation of PhoPQ and EvgSA targets[98], [190]. Although there is no substantial evidence that EvgS can phosphorylate PhoP, the high correlation coefficient of their expression profiles might suggest this is possible.

6.5 Postulations on the mechanism by which PhoP regulates sigmaE

The observation was made that all *phoP* mutants, when assayed, exhibited a drop in sigmaE activity compared to the wildtype(CAG45114). Although this was initially thought to be a result of the accumulation of secondary suppressor mutations, the observation was confirmed by assaying several other backgrounds (*mgrB*, *lac*UV5-*phoP*), which would presumably be suppressor-free. These data assert that PhoP is a key promoter of sigmaE activity. This regulatory relationship is logical when considering the role Mg^{2+} has in stabilising the OM, divalent cation starvation would be an advantageous signal to be integrated into the sigmaE expression network.

Conversely, it was also found in this work that PhoP causes a decrease in total sigmaE protein concentration. Although surprising, this observation supports the assertion that PhoP is a positive regulator of sigmaE activity. While it is known that sigmaE inhibition causes cell lysis[192], it has been shown that excessive sigmaE activity can also be detrimental to the viability of the cell and that sigmaE induction must be appropriately modulated[57], [258]. Ades et al. suggest that this is because many sigmaE-regulated genes require translocation through the periplasm, and the increased expression of these genes may oversaturate and overwhelm the periplasmic chaperones, thereby leading to improperly folded OMP, which in turn increases sigmaE induction, forming a positive feedback loop[213]. Therefore, if PhoP is a key inducer of sigmaE activity, then a mechanism to constrain the hyper-induction of sigmaE would be expected.

A clear and unambiguous method by which PhoP regulates sigmaE was not explicated experimentally by this work, but assertions can be made from both the published literature and observations made herein. The discovery of three potential PhoP box consensus sequences in the *rpoE* promoter region offers a potential mechanism through which PhoP could positively regulate sigmaE, acting as transcriptional activator for *rpoE*. However, the observation that PhoP causes a decrease in total sigmaE protein concentration remains unexplained. An alternative hypothesis is that PhoP is a direct transcriptional repressor of *rpoE* expression and increases sigmaE activity indirectly. One way in which PhoP indirectly increases sigmaE activity is via sigmaS[229]. Recently, it was found that sigmaS has a positive effect on *rpoE* expression[229], which was supported by observations in section 5.6. However, if sigmaS was the only mechanism by which PhoP regulated sigmaE, then it would be expected that there would be a transduction defect associated with *rpoS* and none was observed. Therefore, it is likely that there is at least one other mechanism by which PhoP positively regulates sigmaE activity. It is posited by this work that PhoP increases the rate of sigmaE-RseA proteolysis by increasing the rate of LPS synthesis and concentration of LPS precursor molecules, which it has been suggested enhances sigmaE activity[259].

SigmaE-RseA proteolysis is a multi-step event involving several enzymes. Misfolded outer membrane proteins activate DegS[259], which cleaves the periplasmic domain of RseA. The cytoplasmic domain of RseA is cleaved by RseP, and then the free RseA-sigmaE complex is finally degraded by ClpXP, to release free sigmaE[259]. The action of DegS can be inhibited by RseB[259], which binds to the periplasmic domain of RseA[260]. BO et al. identified that RseB can bind LPS, or LPS derivatives such as lipid A or Kdo_2-lipid A, causing the subsequent displacement of RseB from RseA[260]. Further, they concluded that RseB-LPS binding only requires the phosphorylated *N*-acetylglucosamine (GlcNAC) disaccharide and N-linked acyl chains of the lipid-A moiety[260].

LPS synthesis is a complex process that takes place in the periplasm and the cytoplasm and involves more than 20 enzymatic reactions[37]. However, the starting molecule for the LPS precursor lipid A is UDP-N-acetylglucosamine, which

is important for other cell processes such as peptidoglycan synthesis[37]. The synthesis and catabolism of UDP-N-acetylglucosamine involves several pathways and many different enzymes. UDP-N-acetylglucosamine can be synthesised from several precursors such as glucosamine (GlcN),GlcNac or lactose[37]. The *glm* operon regulates the synthesis of UDP-N-acetylglucosamine from fructose-6-P, while the *nag* operon regulates the transport and catabolism of GlcNAc[29]. NagA (GlcNac-6-P deacetylase), negatively regulated by PhoP, converts GlcNAc-6-P to GlcN-6-P.

In a *phoP* mutant, there would presumably be increased concentrations of LPS precursors available, which would allow for the increased synthesis of LPS. Therefore, it is possible that PhoPQ may indirectly regulate sigmaE by regulating the availability of LPS or LPS precursors such as UDP-N-acetylglucosamine, by negatively regulating NagA and increasing the available concentration of GlcNac-6-P available for LPS synthesis, which causes inhibition of RseB and an increased rate of DegS mediated sigmaE-RseA proteolysis (Figure 6.2).

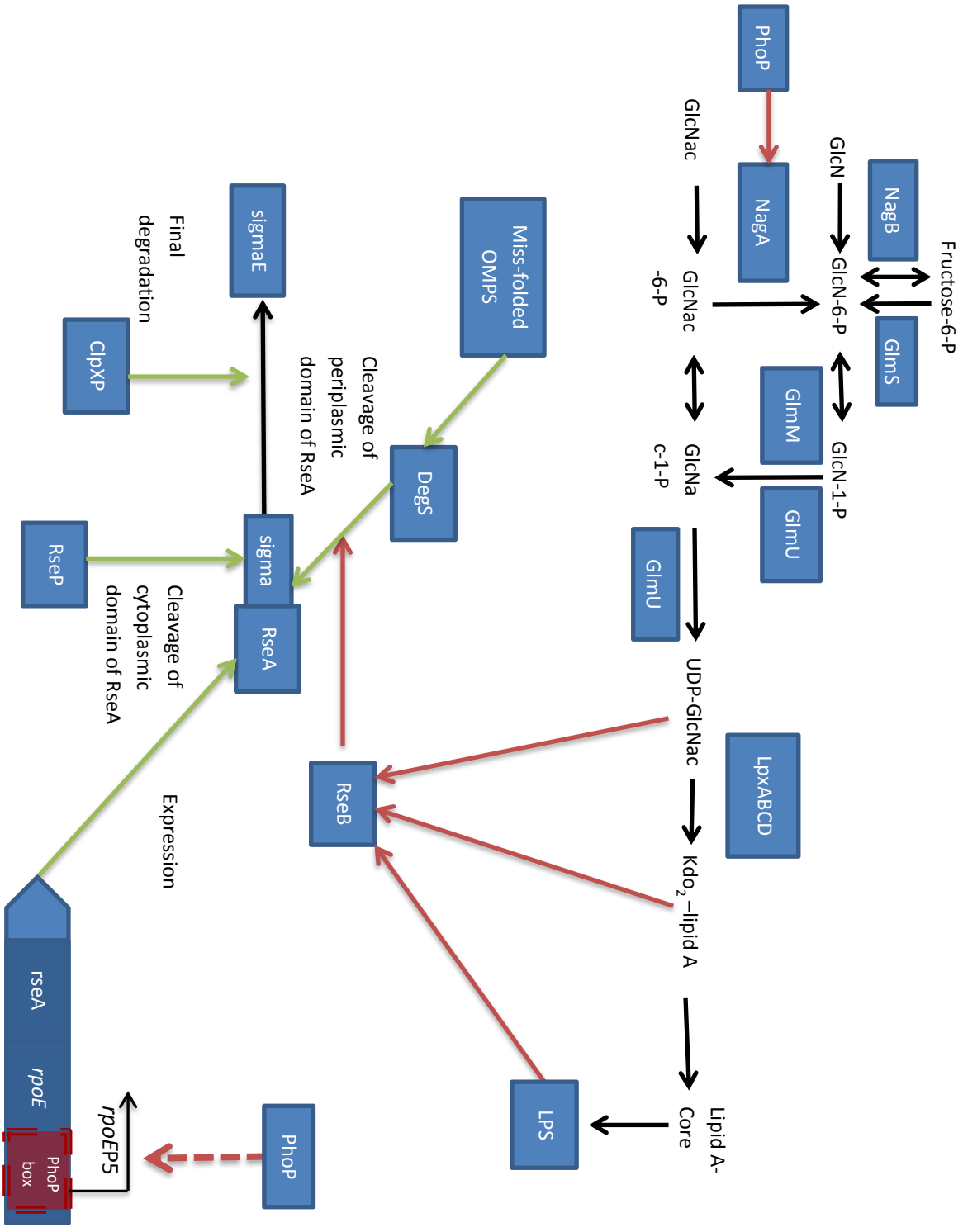


Figure 6.2. Showing a proposed model in which PhoP regulates sigmaE. Red arrows represent negative regulatory relationships, green represent positive regulatory relationships. Solid arrows represent regulatory relationships from the published literature, while the dashed arrow represent a regulatory relationship proposed by this work. The PhoP box, posited by this work to exist in the *rpoE* promoter region (P5), is highlighted in red.

6.6 Final model

This work has aimed to demonstrate the existence of, and then explain, a transduction defect associated with *phoP*. Characterisation of this defect has shown that it can be suppressed by the genetic (4.5), chemical (5.6) or environmental (4.4) induction of sigmaE, through adaption, ostensibly mediated via stress-induced mutation (4.3) and presumably by increasing the availability of divalent metal ions during transduction (4.2).

The most likely explanation for the primary observation of this work, i.e. that it is not possible to transduce *phoP* under the standard conditions used for P1*vir* transduction[153], is that PhoP is a primary activator of sigmaE and is indeed necessary for the full induction of sigmaE. However, it is clear that, unlike *rpoE*, *phoP* is not an essential gene. This suggests that the loss of *phoP* does not cause sigmaE activity to drop to a point at which the cell becomes inviable under optimum conditions. It is expected that the unique stresses of transduction, divalent metal ion chelation and treatment with SDS, cause levels of sigmaE activity in *phoP* mutants to become insufficient for the continued survival of the cell.

The method by which PhoP regulates sigmaE is complex and multifactorial. Recent work has highlighted new details of the transcriptional regulation of *rpoE*, such as the effect of sigmaS[229], providing one mechanism through which PhoP regulates sigmaE activity. Moreover, the discovery of a potential PhoP binding site, located inside *rpoE*P5, provides another convincing mechanism through which PhoP may regulate *rpoE* expression directly and causes a drop in the total sigmaE protein concentration. Finally, it is posited by this work, that there exists another mechanism by which PhoP positively regulates sigmaE activity, by increasing the availability of LPS precursor molecules thereby enhancing sigmaE-RseA proteolysis and substantially increasing sigmaE activity (Figure 6.3).

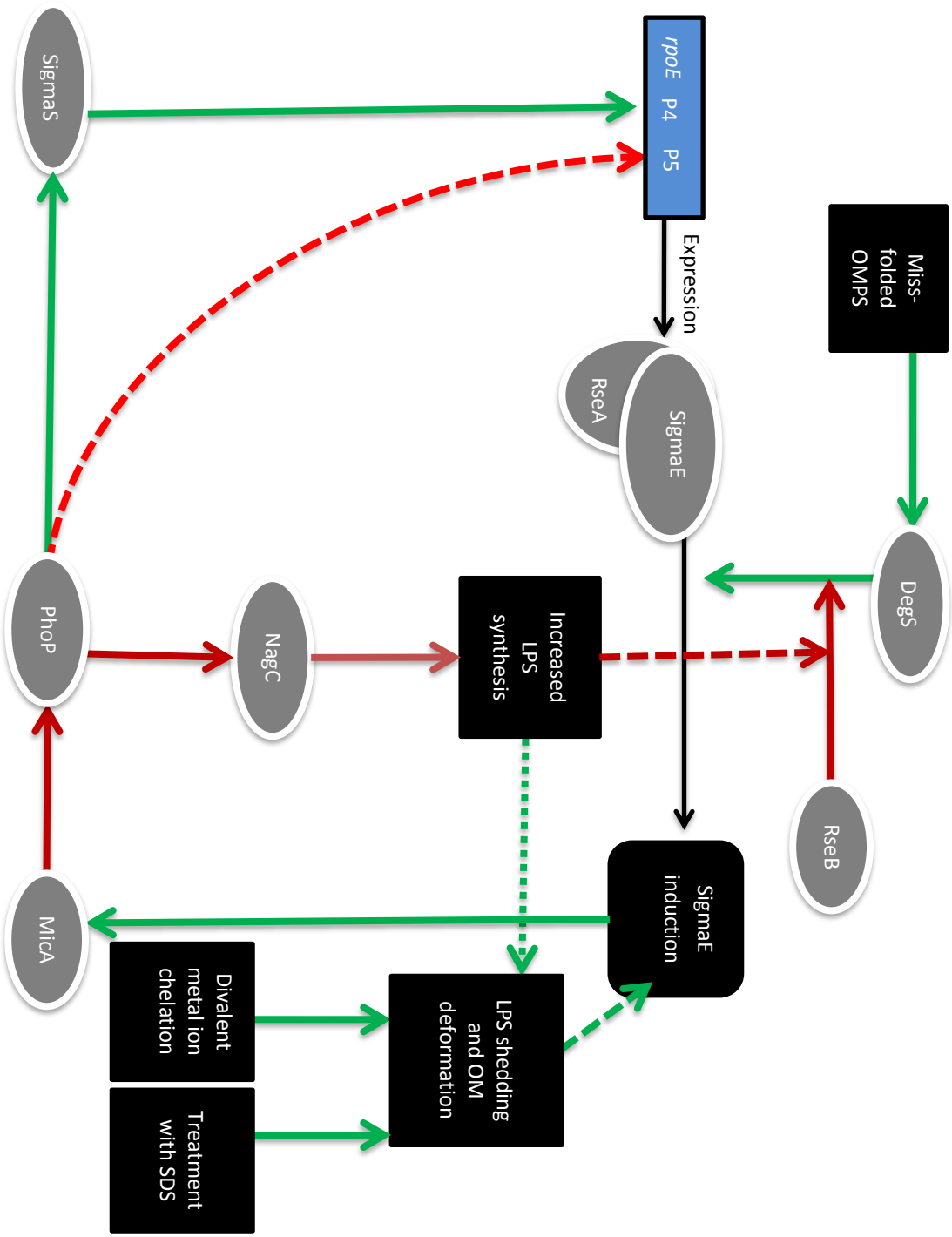


Figure 6.3. Proposed model in which PhoP regulates sigmaE and the effect on the OM. Red arrows represent negative regulatory relationships, and green represent positive regulatory relationships. Solid arrows represent regulatory relationships from the published literature, while dashed arrows represent regulatory relationships proposed by this work. Grey ovals represent proteins or mature sRNAs. The *rpoE* gene, along with promoter 4 (P4) and promoter 5 (P5), is depicted as a blue rectangle.

6.7 Future work

Although many experiments were carried out in this work, in an attempt to better understand the transduction defect associated with *phoP* and the underlying mechanism, there were still many more that were not done, due to either time constraints or resources. The assays detailed below aim to reinforce the validity of observations already made and perhaps reveal new avenues of study.

6.7.1 Characterising *phoP* mutants from the published literature

Various *phoP* mutants in the literature have been made via transduction (Appendix Table E.1). The primary thesis of this work is that no, suppressor free, *phoP* mutant can be created via transduction under normal conditions, and so testing the validity of *phoP* mutants created via transduction and procedures used from the published literature would be a primary aim of any future work. It would be hypothesised that any *phoP* mutants created via transduction would contain secondary suppressor mutations, or the recipient strain would have been prepared in such a way as to make transduction possible; for example, strains were physiologically adapted to accomplish the transduction of *phoP* in this work through exposure to procaine and transduction at 28°C.

Initially, several *phoP* mutants and their original parent strains would be sequenced (Table 6.1), and it would be expected that the parent strain or the *phoP* mutant would contain additional suppressor mutations. If this was not so, then it might suggest that either the recipient strains were predated or that some aspect of the precise procedure used allowed for the creation of a *phoP* mutant. Consequently, it would be advantages to attempt to recreate the *phoP* mutants by using the recipient and donor strains specified in the published literature. If any transductants could be created, then they would be sequenced to confirm that they lack any additional suppressor mutations. If no suppressor mutations could be found, it would suggest that creating *phoP* mutants is possible but something about the strains or the procedure used in this work prevented it.

This idea could be explored further by transducing different combinations of donor and recipient strains from the published literature and the Blomfield laboratory stock. This would allow for isolating the defective strain and permit the further characterisation of the transduction defect as either a pervasive defect effecting the creation of *phoP* mutants or the result of a genetic artefact in the laboratory strains.

6.7.2 Exploring the Absence of a *phoQ* transduction defect

The literature suggests that the most likely explanation as to why there is no transduction defect associated with *phoQ* is because PhoP is active in the absence of *phoQ*. It has been suggested previously in this chapter that the sensor kinases EnvZ or EvgS may have the ability to trans-phosphorylate PhoP. To test if either one of these response regulators has the ability to phosphorylate PhoP, both *evgS* and *envZ* could be transduced in *phoQ* mutants under standard conditions. If the transductions were possible, then it would rule out the most likely candidates for non-cognate trans-phosphorylation of PhoP and suggest that PhoP is not being activated by an alternative sensor kinase.

Second, it would be advantageous to test if PhoP is active in its unphosphorylated state, in *E. coli*. Lejona et al. definitely showed that PhoP is able to activate its target genes in a concentration-dependent manner, independently of phosphorylation by PhoQ in *S. enterica*[257]. The procedure that was used could be replicated in *E. coli*, wherein expression of PhoP targets was measured in a strain which PhoP had been modified to prevent phosphorylation.

Finally, testing for a non-apparent *phoQ* transduction defect could be done in a number of ways. It would be expected that a *phoQ* transduction defect would be associated with suppressor mutations, as observed in *phoP* mutants, so the assay in which *phoP* was transduced into the *rpoH*P3-lacZ reporter fusion could be repeated for *phoQ*. If the expression of *rpoH*P3 was heterogenous, then this would indicate the presence of suppressor mutations. However, a far more definitive way to test

for suppressor mutations would be simply to sequence a number of *phoQ* mutants made via transduction.

6.7.3 Further tests required to characterise the *phoP* transduction defect

The most expedient way to characterise the transduction defect would be to sequence a number of *phoP* mutants. Whole-genome sequencing would be carried out on a number of *phoP* transductants, and it would be hypothesised that a number of secondary mutations would be in systems that either increase sigmaE activity, through increased expression or induction or mutations that relive envelope stress, perhaps by modifying the LPS in such a way as to decrease electrostatic repulsion. Categorising these mutants and the systems they affect may help to elucidate the mechanism underlining the *phoP* transduction defect.

Moreover, sequencing could also be used to confirm the hypothesis that *phoP* mutants have secondary mutations because they undergo transduction – this seems a probable assertion, but it is possible that *phoP* mutants are generally more likely to undergo mutation, perhaps because of an impaired ability to respond to DNA damage. Indeed, the *phoP*-regulated HemL has been implicated in playing a role in SOS DNA damage response, which has been linked to mutational rate, the frequency at which new mutations arise[261]. To confirm the hypothesis that transduction causes *phoP* mutants to accumulate suppresser mutations, several *phoP* mutants created via lambda red could be subjected to whole-genome sequencing. These data could then be compared to those of the mutants that have been created via transduction. Mutations in different systems or at different rates would suggest that transduction is the cause of suppresser mutation accumulation rather than simply the loss of *phoP*.

P1's contribution to the transduction defect could be explored in greater detail. One possible cause of the transduction defect may have been a low titre of P1 in the lysates used. Although three different preparations of lysates were used when validating the initial observation (3.2) it is possible that a low titre of P1 resulted

from the particular method used to create the *phoP* P1 lysate. This seems unlikely, though, there were no transduction defects noted when the P1 lysates were prepared from other donor strains using the same strain of P1, albeit it does remain a possibility. Lysate concentration could be discerned by doing a plague assay, using *phoP* P1 lysate and *E. coli*. If the concentration of P1 was particularly low, this could be a cogent explanation for the observed transduction defect.

Further, it would also be advantageous to assess *phoP* mutants' sensitivity to P1 killing. Although P1 killing seemed unlikely to be the cause of the *phoP* transduction defect, it was not definitely dismissed. A simply assay using varying concentrations of P1 could be done to compare survival rates between the wildtype and *phoP* mutants created via both transduction and lambda red. It would be expected that *phoP* mutants would not be more sensitive to P1 killing than the wildtype, although because *phoP* mutants are expected to suffer from OM disruption, it is a possibility. If sensitivity to P1 was the cause of or at least a contributor to the *phoP* transduction defect, it would suggest that *phoP* mutants that contain additional suppressor mutations gained immunity to P1, as *phoP* mutants that had undergone P1 transduction could be transduced normally. If this were true, then it would be interesting to examine the LPS, P1 absorption receptor to look for any structural anomalies.

6.7.4 Further characterising the relationship between PhoP and sigmaE

The observation that PhoP induction caused a fall in sigmaE protein concentration was unexpected, so in any future work, importance should be placed on replicating this data along with a particular emphasis on clarity and quantifiability. A protein loading control, for example the constitutively expressed 60kDa chaperonin GroEL[262], could be used to increase confidence and reproducibility in subsequent experiments. Moreover, to increase confidence in this observation, several other assays could be done to quantify the sigmaE protein concentration in cells in which *phoP* expression is induced. Several variations of the original western blot could be done. In the original Western blot procedure, a polyclonal antibody was used, but

although inexpensive, using this method has the disadvantage of detecting multiple epitopes, due to a high chance of cross-reactivity. The lack of clarity was a weakness of the original assay, but this could be improved by using a sigmaE monoclonal antibody, which would be highly specific to the sigmaE epitope and therefore much easier to use for the quantification of protein levels. Alternatively, a recombinant histidine-tagged, a poly-histidine tail, sigmaE protein could be designed and used to perform another Western blot, with the aim of identifying his-tagged proteins. These proteins could also be easily purified and quantified by either measuring absorbance, via a BCA or Branford assay, or performing an immunoassay, such as ELISA.

Throughout this work, evidence has been put forward to support the hypothesis that PhoP regulates sigmaE. Specifically, observations were made that suggest a model in which PhoP has a negative effect on sigmaE concentration but a stimulatory effect on activity. It was hypothesised that this was due to two different mechanisms of regulation, namely direct and indirect, so it would be valuable to explore these hypothesised mechanism of regulation further.

To examine the hypothesised direct effect of PhoP on *rpoE* expression, several assays could be completed, one of which is a gel shift, electrophoretic mobility assay, used commonly to detect TF-DNA binding based on the principle that typically the electrophoretic mobility of a TF-DNA complex is less than the corresponding unbound DNA. If no binding occurs, a single band will be seen that is expected to migrate through the gel at the same speed to a DNA alone control. TF-DNA interaction would normally be indicated by a slower migration through the gel compared to DNA alone. This is visualised as either one or two bands, namely one band nearer the cathode signifying a DNA-TF complex, and possibly a second band nearer the anode representing excess unbound DNA. Alternatively, a promoter analysis could be done. Promoter analyses require incorporating a gene of interest, along with upstream regulatory sites, into a plasmid containing a reporter, such as a green fluorescent protein or luciferase, which is then transformed into the cell. If the gene of interest is regulated by a TF, the expression of the reporter will be dependent on the activity of the TF. Point mutations can be introduced into the

suspected TF binding site, affecting the TF-DNA binding, and in this way the binding site and crucial nucleotides can be interpreted. Either of these assays could demonstrate PhoP binding the *rpoE* promoter region, thus indicating a direct regulatory relationship.

It has been proposed that PhoP regulates sigmaE via direct transcriptional control and RseA-sigmaE proteolysis. The regulation of sigmaE via proteolysis has been hypothesised to be mediated via RseB, specifically its ability to bind LPS. If this is correct, then RseB may share a transduction defect; however, PhoP's effect on sigmaE has not been postulated to be entirely dependent on RseB, so measuring sigmaE activity in an *rseB* mutant would not provide sufficient evidence to support this hypothesis. However, PhoP regulation of sigmaE proteolysis, mediated via RseB is asserted to be dependent on the available concentration of lipid A precursors, specifically GlcN-6-P. Therefore, this hypothesis could be tested by creating an inducible *nagA* construct. Increasing the endogenous concentration of GlcN-6-P by controlling the expression of *nagA* and monitoring the effect of sigmaE activity would show that this is a valid mechanism through which PhoP may regulate sigmaE activity.

Appendix

Appendix A

A.1. To assess the success of allelic exchange between *gshA,phoP*Δ*lac*Ω*sacB*-Kan^R and pA001, two PCR assays were performed. The aim of the first assay was to amplify a *phoP* gene fragment using the same internal primers employed to create the original fragment used to create pA001 (Figure A.1A). This showed that the isolates contained a *phoP* allele. Although successful, this assay could not be used to differentiate between the wildtype and the ectopic *phoP* gene. A second PCR assay (data not shown) was done, using external primers (Figure A.1B), the aim of which was to amplify a gene fragment with primers that annealed to regions flanking the *phoP* open reading frames, 94Bp and 57Bp, up- and downstream of the ATG start codon and the TGA stop codon, respectively. However, the PCR assay failed to yield any results, which suggests that isolates contained the ectopic copy of *phoP* but did not have the same success with the wildtype.

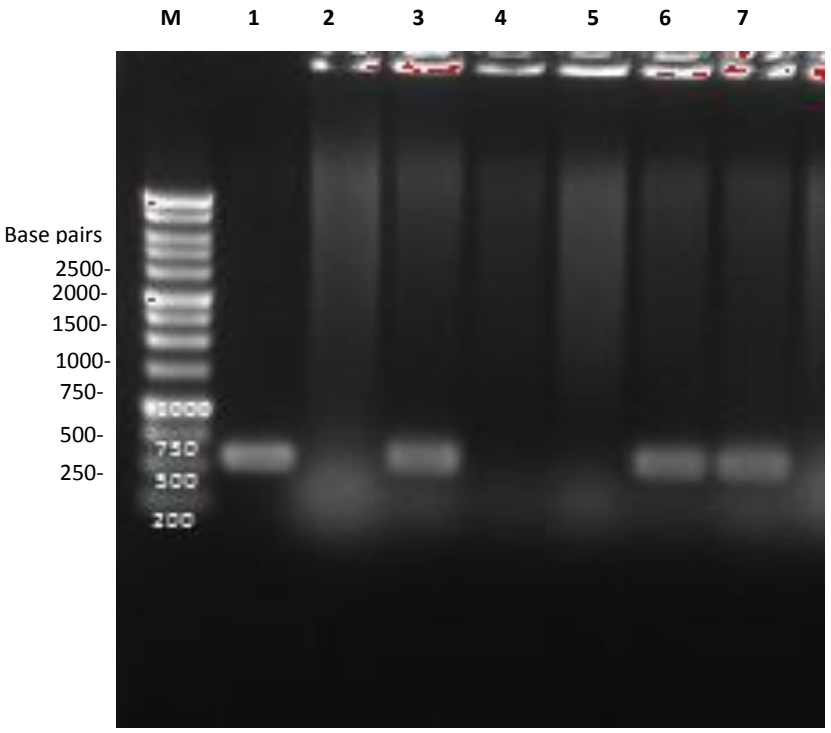


Figure A.1A. Gel electrophoresis. 1% agarose gel. Multiplex PCR assay performed on *lac*UV5-*phoP*, *phoP* isolates (section 3.2.1). Showing *phoP* fragment generated 687BP from internal primers in the first, third sixth and seventh lanes, while lanes two, four and five represent unsuccessful constructs. The legend represents the molecular ladder ranging from 250Bp to 2500Bp.

Gagctatcacgatggttgatgagctgaaataaacctcgtatcagtgccggatggcgatgctgtccggcctgcttattaa
gattatccgcttTttatttttttcactttacctcccctccccgctggtttatttaatgtttaccccataaccacataatcgcgttac
actattttaataattaagacagggagaaataaaa
atgcgcgtactggttgttgaagacaatgcgttgttacgtcaccaccttaaagttcagattcaggatgctggtcatcaggtcg
atgacgcagaagatgccaaagaagccgattattatctcaatgaacatataccggatattgcgattgtcgatctcggattg
ccagacgaggacggtctgtcactgattcgccgctggcgtagcaacgatgtttcactgccgattctggtattaaccgcccg
tgaaagctggcaggacaaagtcgaagtattaagtgccggtgctgatgattatgtgactaaaccgtttcatattgaagagg
tgatggcgcgaatgcaggcattaatgcggcgtaatagcggtctggcttcacaggtcatttcgctccccccgtttcaggttg
atctctctcgccgtgaattatctattaatgacgaagtgatcaaactgaccgcgttcgaatacactattatggaaacgttgat
acgcaataatggcaaagtggtcagcaaagattcgttaatgctccaactctatccggatgcggagctgcgggaaagcc
ataccattgatgtactgatgggacgtctgcgcaaaaaaattcaggcacaatatccccaagaagtgattaccaccgttcg
cggccagggctatctgttcgaattgcgctga

tgaaaaaattactgcgtctttttttcccgctctcgctgcgggtacgttttctgttggcaacggcagcggtagtactggtgctttcg
cttgcctacggaatggtcgcgctgatcggttatagcgtcagtttcgataaaactacgtttcggctgttacgtggcgagagca
atctgttctatacccttgcgaagtgggaaaac

PhoP internal primer forward (Vector_*phop*- forward)

PhoP internal primer reverse (Insert_*phop*2- reverse)

PhoP external primer forward (*phop*-forward)

PhoP external primer reverse (*phoP*-reverse)

Figure A.1B. Showing the *phoP* gene sequence and position of the internal (purple) and external (green) primers used in an attempt to amplify the ectopic *lac*UV5-*phoP* construct generated in section 3.2. The green highlighting indicates the transcription start site. The blue highlighting indicates the Shine-Dalgarno sequence[112]**. The red highlighting indicates the start codon. The grey highlighting indicates an area upstream of the ATG start codon and downstream of the TGA stop codon.**

A.2 Two PCR assays were done on the isolates generated in section 3.2.3.The aim of the first example was to amplify the wildtype *phoP* gene (using the primers *phoP*-forward and *phoP*-reverse (data not shown); however, no amplicon was detected by gel electrophoresis, which suggested that the wildtype *phoP* gene was not present. A second PCR assay (Figure A.2) was performed with the aim of amplifying a 616Bp fragment of the *cat* gene, which encodes chloramphenicol acetyltransferase, using the primers Cam^R-reverse and Cam^R-forward. pKD (lane 1) was used as a control, as it was from this plasmid that the *cat* gene was originally cloned. Lanes 2-5 show the successful amplification of the *cat* gene from several isolates, and these two assays demonstrate that *phoP* was successfully replaced with the *cat* gene.

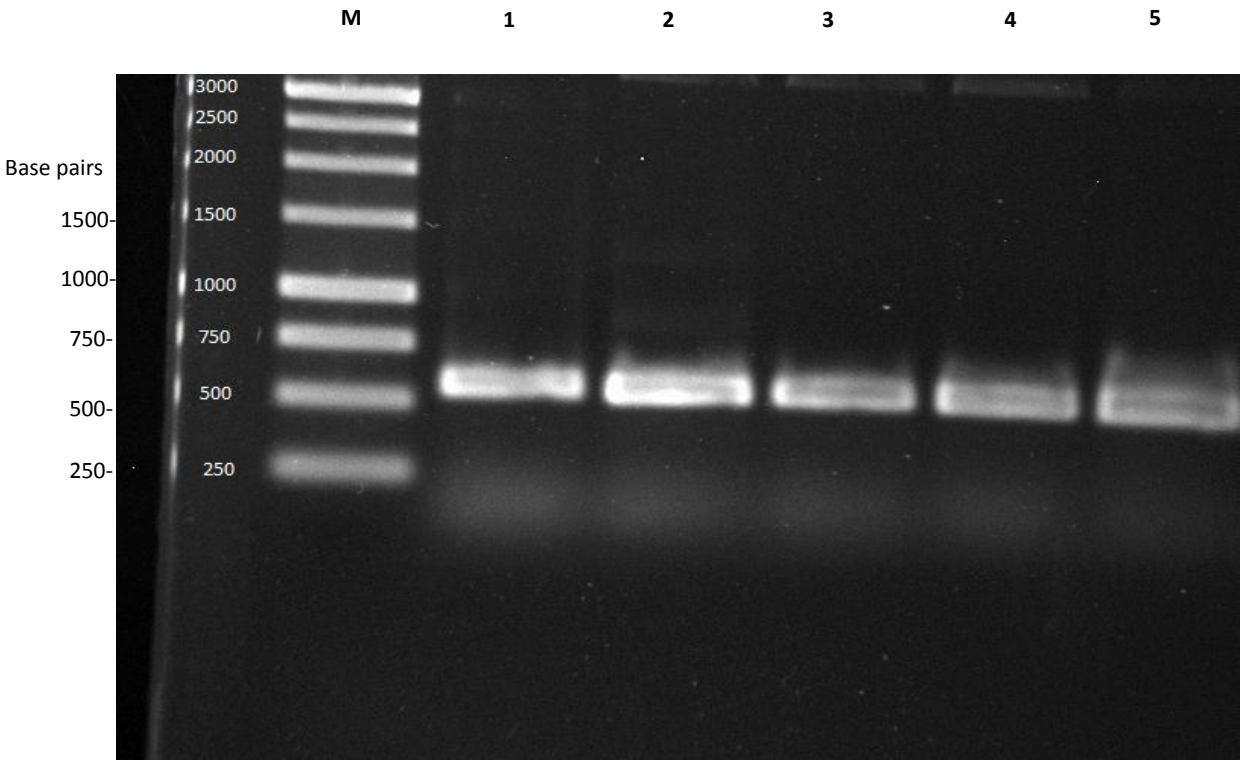


Figure A.2. Gel electrophoresis. 1% agarose gel. PCR assay performed one reference sample, control, and four experimental isolates. Lane 1, *cat* gene (530Bp) fragment amplified from pkD3, lane 2-5 *cat* fragment (530Bp) amplified *phoP*:Cam^R isolates, created using lambda red (section 3.2.3), using primers Cam^R-reverse and Cam^R-forward. The marker (M) represents the molecular ladder ranging from 250Bp to 3000Bp.

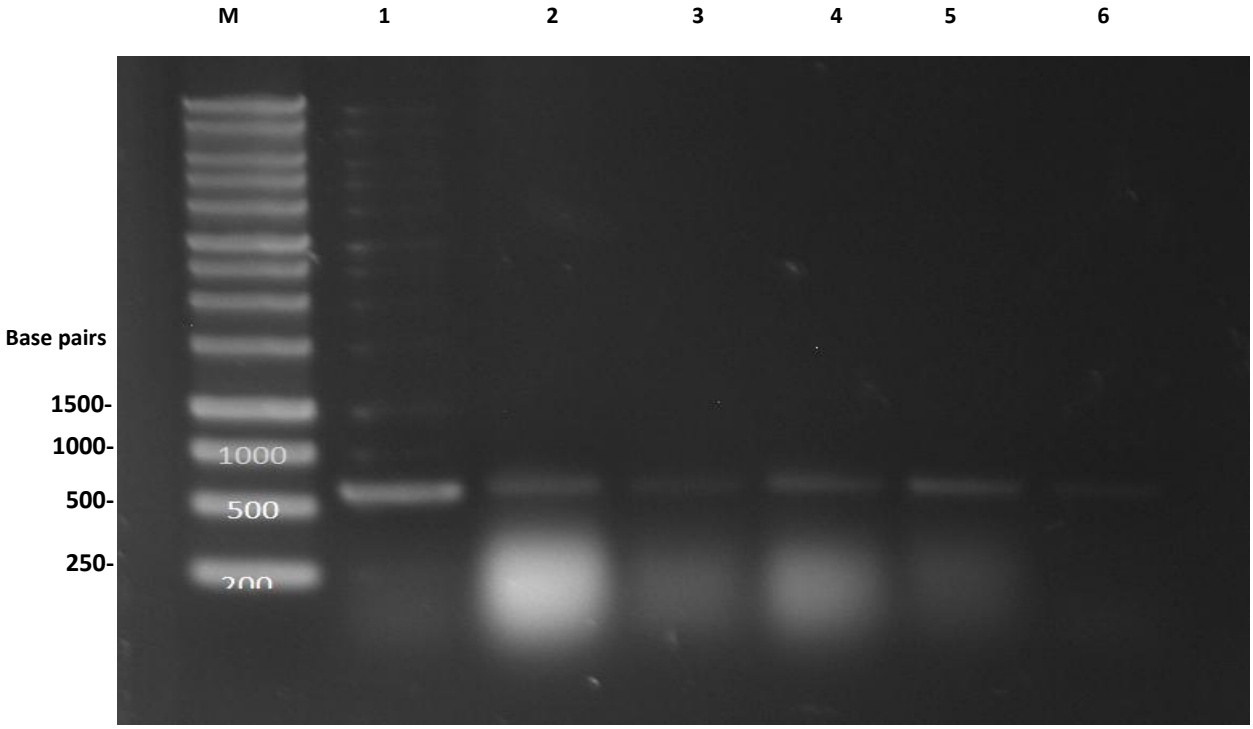


Figure A.3. Gel electrophoresis. 1% agarose gel. Showing *phoP* mutants generated in 4.4.2 using primers *phoP*-forward and *phoP*-reverse . The legend represents the molecular ladder ranging from 200Bp to 1000Bp. A 795Bp fragment containing kanamycin resistance cassette can be seen in lanes 2-6, while lane 1 contains a positive control, an amplified gene fragment containing the kanamycin resistance cassette.

Appendix B

B.1. The dataset (Appendix B.1A and B.1B) represents 25 paired transduction assays, in each of which (repeats 1-25) attempts were made to transduce *phoP* and *phoQ* at the same time for two different lysate concentrations, i.e. 10µl and 50µl. Thus, each repeat consisted of four transduction reactions. After the final incubation step of 2.5 hours, two replicates were plated out for each reaction. The numbers of transductants generated in each replicate for each lysate concentration, per repeat, are given below. In addition, the average of each replicate and the total average of all replicates and repeats for each lysate concentration are given.

Furthermore, the total average number of transductants per 50µl per repeat is given. This value represents the average number of transductants per repeat for both lysate volumes, normalised to 50µl. This was calculated by first dividing the average number of transductants, the average of each replicate pair, for each lysate concentration, 10µl and 50µl, by 10 and 50, respectively, which yielded the average number of transductants per microlitre for each lysate concentration. Next, the total number of transductants per microlitre, per repeat, was calculated. This was the averaged value of the number of transductants per microlitre of both lysate concentrations. Finally, the total average number of transductants per microlitre, per repeat, was multiplied by 50, which provided the total average number of transductants per 50µl.

During these experiments, several different preparations of *phoP* and *phoQ* lysates were used, which are represented in the figure by the colours blue, pink, and green.

Lysate	Repeat	*phoQ* 10µl (1)	*phoQ* 10µl (2)	Average	*PhoQ* 50µl (1)	*phoQ* 50µl (2)	Average	Total average number of *phoQ*(50µl&10µl) transductants/50µl of lysate
phoQ::Kan^R (1)	1	1	2	1.5	18	16	17	12.25
	2	0	2	1	11	10	10.5	7.75
	3	0	0	0	36	29	32.5	16.25
	4	11	12	11.5	45	38	41.5	49.5
	5	2	6	4	33	30	31.5	25.75
	6	0	2	1	30	34	32	18.5
	7	1	0	0.5	30	24	27	14.75
	8	1	2	1.5	10	7	8.5	8
	9	0	0	0	9	7	8	4
	10	1	0	0.5	28	30	29	15.75
	11	1	2	1.5	18	11	14.5	11
phoQ::Kan^R (2)	12	1	2	1.5	43	36	39.5	23.5
	13	1	1	1	24	30	27	16
	14	0	1	0.5	20	22	21	11.75
	15	0	1	0.5	45	40	42.5	22.5
	16	3	1	2	15	15	15	12.5
	17	2	2	2	8	10	9	9.5
	18	6	3	4.5	28	35	31.5	27
	19	0	0	0	7	5	6	3
phoQ::Kan^R (3)	20	0	0	0	9	5	7	3.5
	21	1	1	1	20	21	20.5	12.75
	22	8	1	4.5	26	29	27.5	25
	23	3	6	4.5	30	33	31.5	27
	24	5	3	4	28	26	27	23.5
	25	1	0	0.5	7	6	6.5	4.5
	Total Average			2.0			22.5	16.22
	StDev			2.49			11.67	10.32

Figure B.1A. The number of *phoQ* transductants isolated in the wildtype (MG1655) over 25 assays using both 10µl and 50µl of P1 lysate. The average of the two replicates done for each assay and total average for each lysate concentration is given. Additionally, the total average number of transductants, per 50µl, per repeat is shown which represents the average number of transductants per repeat, for both lysate volumes, normalised to 50µl of lysate.

Lysate	Repeat	*phoP* 10µl (1)	*phoP* 10µl (2)	Average	*PhoP* 50µl (1)	*phoP* 50µl (2)	Average	Total average number of *phoP*(50µl&10µl) transductants/50µl of lysate
phoP::Kan^R (1)	1	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0
	3	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0
	5	0	0	0	0	0	0	0
	6	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	0
	8	0	0	0	0	0	0	0
	9	0	0	0	0	0	0	0
	10	0	0	0	0	0	0	0
	11	0	0	0	0	0	0	0
phoP::Kan^R (2)	12	0	0	0	0	0	0	0
	13	0	0	0	0	0	0	0
	14	0	0	0	0	0	0	0
	15	0	0	0	0	0	0	0
	16	0	0	0	0	0	0	0
	17	0	0	0	0	0	0	0
	18	0	0	0	0	0	0	0
	19	0	0	0	0	0	0	0
phoP::Kan^R (3)	20	0	0	0	0	0	0	0
	21	0	0	0	0	0	0	0
	22	0	0	0	0	0	0	0
	23	0	0	0	0	0	0	0
	24	0	0	0	0	0	0	0
	25	0	0	0	0	0	0	0
	Total Average			0			0	0
	StDev			0			0	-

Figure B.1B. The number of *phoP* transductants isolated in the wildtype (MG1655) over 25 assays using both 10µl and 50µl of P1 lysate. The average of the two replicates done for each assay and total average for each lysate concentration is given. Additionally, the total average number of transductants, per 50µl, per repeat is shown which represents the average number of transductants per repeat, for both lysate volumes, normalised to 50µl of lysate.

Appendix C

***PhoP* experimental**	-5	-5	**Average**		***phoP* control**	-5	-5	**Average**	Av diff (control-experimental
Repeat 1	13	19	16		Repeat 1	60	54	57	41
Repeat 2	23	24	23.5		Repeat 2	53	55	54	30.5
Repeat 3	17	18	17.5		Repeat 3	47	51	49	31.5
Total Average			19		Total Average			53.33	34.33
Wt experimental	-5	-5	**Average**		**Wt control**	-5	-5	**Average**	Av diff (control-experimental
Repeat 1	101	104	102.5		Repeat 1	118	127	122.5	20
Repeat 2	105	91	98		Repeat 2	121	119	120	22
Repeat 3	93	98	95.5		Repeat 3	116	113	114.5	19
Total Average			98.67		Total Average			119	20.33

***PhoP* experimental**	-6	-6	**Average**		***phoP* control**	-6	-6	**Average**	Av diff (control-experimental
Repeat 1	2	3	2.5		Repeat 1	15	14	14.5	12
Repeat 2	8	5	6.5		Repeat 2	13	11	12	5.5
Repeat 3	2	2	2		Repeat 3	13	12	12.5	10.5
Total Average			3.67		Total Average			13	9.33
Wt experimental	-6	-6	**Average**		**Wt control**	-6	-6	**Average**	Av diff (control-experimental
Repeat 1	11	12	11.5		Repeat 1	39*	15	27	15.5
Repeat 2	11	10	10.5		Repeat 2	13	12	12.5	2
Repeat 3	12	14	13		Repeat 3	14	13	13.5	0.5
Total Average			11.67		Total Average			17.67	6

Table C.1 - Average number of (AAEC261A*)* and *phoP (*KCEC4302) colonies after being subjected to transduction reagents, incubated overnight then diluted to -5 and -6. Results denoted with an asterix (*) were assumed to be contamination and so all -6 results were rejected from the main body of work.

Temperature of transduction	**Number of *phoP* transductants**				**Number of *phoQ* transductants**			
	Repeat 1	Repeat 2	Repeat 3	Average	Repeat 1	Repeat 2	Repeat 3	Average
28°C	13,8 4,6 9,7 (47)	7,5 6,3 5,9 (35)	5,2 7,6 11,4 (35)	6.5	8,5 7,7 9,5 (42)	10,13 11,7 4,7 (52)	11,5 15,13 10,6 (60)	8.5
37°C	0,0 0,0 0,0	0,0 0,0 0,0	0,0 0,0 0,0	0	10,18 6,7 8,12 (61)	13,7 9,10 22,18 (79)	11,14 8,11 9,6 (59)	11

Table C.2. The number of *phoP (*JW1116-1) and *phoQ (*JW1115-1) transductants generated in the wildtype (AAEC261A), using 50µl of lysate. Replicate pairs are shown with the average of the three replicates in parenthesis, as well as total averages generated from the three repeated experiments.

Appendix D

Gene	PhoP binding sequence
PhoPQ	GGTTTAtttaaTGTTTA
mgtA	GGTTTAtcgttGGTTTA
mgrB	TAGTTAggcgcTGTTTA
hemL	TGTTTGacgagTATTTA
nagA	TGTTTAtgggcGGTGTA
rstA	TGTTTAgaaacGATTGA
slyB	TGTTTAtaattGGTTGA
yrbL	TGTTTAggtttTGTTTA
vboR	TGTTTAggattTGTTTA

Table D.1 – The PhoP binding motifs for the Mg^{2+} stimulon outlined by Minagawa et al. (2003)[42]

Gene	P-Value	Motif
treR 1	5.65E-05	GGTTTATCGTTGGTTTA
treR 2	0.254	GGGTAAAGTCTGGTTTA
treR 3	0.281	GGGTAAAAATTTCTTTA
treR 4	0.496	CATTTCTGTACTGTTTC
treR 5	0.496	TGTTTAAAGAAATTTTT
treR 6	0.496	TGTTGTTTAATTATTTG
treR 7	0.496	GGGTTAAATCAGGCGGA
treR 8	0.513	TGGGTAAAGTCTGGTTT
treR 9	0.513	GTTTAAAGAAATTTTTA
treR 10	0.591	GGCTCATTCGCCATTTA
treR 11	0.591	TGGGTGTCGAACGTTTT
treR 12	0.591	GTTTTAATCTCCGTCGA
treR 13	0.647	TGGTTTATCGTTGGTTT
treR 14	0.647	GCGTTACCGGATGCGTA
treR 15	0.647	GATATAATACCTGCTGA
treR 16	0.647	GTTTCAGACAGTGCGGA
treR 17	0.647	TGGTTCATCGTGATCCA

Table D.2 – List of *treR* PhoP binding motifs predicted by FIMO, experimentally validated motifs are highlighted in green.

Gene	P-Value	Motif
ybjG 1	2.02E-05	Tctttaagttttattta
ybjG 2	0.000347	TGCTTAACCGTGGTTTC
ybjG 3	0.00179	TGTTTACGGTGGGTGAT
ybjG 4	0.00231	ATGTTGGCCGTGGTGGA
ybjG 5	0.00315	TCGTTAGCTGCGCTTTT
ybjG 6	0.00322	GGGAGTGGTGTGGTTTA
ybjG 7	0.00407	Tagctacgctttcttta
ybjG 8	0.00475	Gctttctttaagtttta
ybjG 9	0.0062	AGGTTAAACTGGGTAAA
ybjG 10	0.00881	AGGTTAAATAAAACTTA
ybjG 11	0.00892	Agtatagctacgctttc
ybjG 12	0.01	GGTATGTTGGCCGTGGT

Table D.3 – List of *ybjG* PhoP binding motifs predicted by FIMO, experimentally validated motifs are highlighted in green.

Gene	P-Value	Motif
crcA	0.000133	GCTTTAGGAATTTTTTA
crcA	0.000157	Tctttatgttgggtcta
crcA	0.000202	GTTTTGTTATCTATGTA
crcA	0.00037	Tattaaggttatgttaa

Table D.4 – List of *crcA* PhoP binding motifs predicted by FIMO, experimentally validated motifs are highlighted in green.

Gene	P- value	Observed	1.00E-05	1.00E-04	1.00E-03	1.00E-02	1.00E-01	1.00E+00
treR 1	5.65E-05	p	N	P	P	P	p	p
ybjG 1	2.02E-05	p	N	P	P	P	p	p
crcA 1	0.000133	N	N	N	P	P	p	p
crcA 2	0.000157	N	N	N	P	P	p	p
crcA 3	0.000202	N	N	N	P	P	p	p
ybjG 2	0.000347	N	N	N	P	P	p	p
crcA 4	0.00037	p	N	N	P	P	p	p
ybjG 3	0.00179	N	N	N	N	P	p	p
ybjG 4	0.00231	N	N	N	N	p	p	p
ybjG 5	0.00315	N	N	N	N	p	p	p
ybjG 6	0.00322	N	N	N	N	p	p	p
ybjG 7	0.00407	N	N	N	N	p	p	p
ybjG 8	0.00475	N	N	N	N	p	p	p
ybjG 9	0.0062	N	N	N	N	p	p	p
ybjG 10	0.00881	N	N	N	N	p	p	p
ybjG 11	0.00892	N	N	N	N	p	p	P
ybjG 12	0.01	N	N	N	N	N	p	P
treR 2	0.254	N	N	N	N	N	N	P
treR 3	0.281	N	N	N	N	N	N	P
treR 4	0.496	N	N	N	N	N	N	P
treR 5	0.496	N	N	N	N	N	N	P
treR 6	0.496	N	N	N	N	N	N	P
treR 7	0.496	N	N	N	N	N	N	p
treR 8	0.513	N	N	N	N	N	N	p
treR 9	0.513	N	N	N	N	N	N	P
treR 10	0.591	N	N	N	N	N	N	P
treR 11	0.591	N	N	N	N	N	N	P
treR 12	0.591	N	N	N	N	N	N	P
treR 13	0.647	N	N	N	N	N	N	P
treR 14	0.647	N	N	N	N	N	N	P
treR 15	0.647	N	N	N	N	N	N	P
treR 16	0.647	N	N	N	N	N	N	P
treR 17	0.647	N	N	N	N	N	N	P

Table D.5 .Confusion matrix predictions for P-value thresholds. Motifs predicted by FIMO, testing the efficacy of P-value thresholds showing the number of correct (P) and incorrect (N) predictions.

Appendix E

Paper	Strain	Origin of strain	recipient strain	Parent strain genotype	Construction method	Method reference
A PhoQ/P-Regulated small RNA Regulates Sensitivity of *Escherichia coli* to Antimicrobial Peptides[194]	EG12976 *phoP*::Kan^R	Dissecting the PhoP regulatory network of *Escherichia coli* and *Salmonella enterica*	MG1655	F$^-$, lambda$^-$, rph-1	P1*vir* transduction	No reference given
B1500, a small membrane protein, connects the two-component systems EvgS/EvgA and PhoQ/PhoP in *Escherichia coli*[99]	MG1622 *phoP*::Cam^R	Molecular Characterization of the PhoP-PhoQ Two-Component System in *Escherichia coli* K-12: Identification of Extracellular Mg^{2+} Responsive Promoters	MG1601	*mgtA*::λplacMu55	tn10 mutagenesis	-
Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection[160]	JW1116-1 *phoP*::Kan^R	Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection	BW25113	*rrnB*3 Δ*lacZ*4787 *hsdR*514 Δ(araBAD)567 Δ(rhaBAD)568 rph-1.	Transformation	-
Dissecting the PhoP regulatory network of *Escherichia coli* and *Salmonella enterica*[96]	EG12976 *phoP*::Kan^R	Dissecting the PhoP regulatory network of *Escherichia coli* and *Salmonella enterica*	MG1655	F$^-$, lambda$^-$, rph-1	P1*vir* transduction	No reference given
Genomic SELEX Search for Target Promoters under the Control of the PhoQP-RstBA Signal Relay Cascade[263]	JD22184 *phoP*::Kan^R	Genomic SELEX Search for Target Promoters under the Control of the PhoQP-RstBA Signal Relay Cascade	KP7600	W3110 typeA *lacI^{q}lacZΔM15 galK2 galK22*	tn10 mutagenesis	-
Identification and Molecular Characterization of the	MP4022 *phoP*::Cam^R	Molecular characterization of the PhoP-PhoQ two-component system in	MC4100	F− Δ(argF-lac)U169 *araD*139 *rpsL*150 ptsF25 fibB5301 *rbsR deoC relA*1	tn10 mutagenesis	-

Mg^{2+} Stimulon of *Escherichia coli*[42]		*Escherichia coli* K-12: identification of extracellular Mg^{2+} responsive promoters				
Identification and Molecular Characterization of the Mg^{2+} Stimulon of *Escherichia coli*[42]	WP3022 *phoP*::Cam^R	Identification and Molecular Characterization of the Mg^{2+} Stimulon of *Escherichia coli*	W3110	F- λ- rph-1 INV(*rrnD*, *rrnE*)	P1*vir* transduction	No reference given
Lipid Trafficking Controls Endotoxin Acylation in Outer Membranes of *Escherichia coli*[264]	FS1000 *phoP*::Kan^R	Molecular Genetic Analysis of the *Escherichia coli phoP* Locus	MC1061	*araD*139 Del(araA-leu)7697 Del(lac)X74 *galK*16 *galE*15(GalS) lambda- e14- mcrA0 *relA*1 *rpsL*150(strR) *spoT*1 *mcrB*1 *hsdR*2	P1*vir* transduction	Experiments in molecular genetics Jeffrey H. Miller
MicA sRNA links the PhoP regulon to cell envelope stress[112]	MG1423 *phoP*::Cam^R	MicA sRNA links the PhoP regulon to cell envelope stress	MG1188		tn10 mutagenesis	-
MicA sRNA links the PhoP regulon to cell envelope stress[112]	MG1446 *phoP*::Kan^R	MicA sRNA links the PhoP regulon to cell envelope stress	MG1173	MG1655 Δ*lac*X174 λRS*ompT-lacZ*	P1*vir* transduction	No reference given
Molecular Characterization of the PhoP-PhoQ Two-Component System in *Escherichia coli* K-12: Identification of Extracellular Mg^{2+} Responsive Promoters[88]	MP4022 *phoP*::Cam^R	Molecular Characterization of the PhoP-PhoQ Two-Component System in *Escherichia coli* K-12: Identification of Extracellular Mg^{2+} Responsive Promoters	MC4100	F−Δ(argF-lac)U169 *araD*139 *rpsL*150 *ptsF*25 *fibB*5301 *rbsR deoC relA*1	P1*vir* transduction	No reference given
Molecular Characterization of the	MG1622 *phoP*::Cam^R	Molecular Characterization of the PhoP-PhoQ Two-	MG1601	*mgtA*::λplacMu55	tn10 mutagenesis	-

PhoP-PhoQ Two-Component System in *Escherichia coli* K-12: Identification of Extracellular Mg^{2+} Responsive Promoters[88]		Component System in *Escherichia coli* K-12: Identification of Extracellular Mg^{2+} Responsive Promoters				
Molecular Genetic Analysis of the *Escherichia coli phoP* Locus[102]	FS1000 *phoP*::Kan^R	Molecular Genetic Analysis of the *Escherichia coli* phoP Locus	MC1061	F- *araD*139 Del(araA-leu)7697 Del(lac)X74 *galK*16 *galE*15(GalS) lambda- e14- *mcrA*0 *relA*1 *rpsL*150(strR) *spoT*1 *mcrB*1 *hsdR*2	P1*vir* transduction	Experiments in molecular genetics Jeffrey H. Miller
Molecular Genetic Analysis of the *Escherichia coli phoP* Locus[102]	FS1002 *phoP*::Kan^R	Molecular Genetic Analysis of the *Escherichia coli* phoP Locus	MC4100	F- araDJ39 A(*lac*)U169 *rpsL*150 relAl thiflbBS301 deoCi ptsF25 rbsR	P1*vir* transduction	Experiments in molecular genetics Jeffrey H. Miller
Molecular Genetic Analysis of the *Escherichia coli phoP* Locus[102]	JC7623 Δ*phoP*	Genetic recombination in *Escherichia coli*: role of exonuclease I	JC7623	thr-l ara-14 leuB6 A(gpt-proA)62 lacYl sbcC201 tsx-33 supE44 galK2 lambdarac sbcBIS hisG4 rfbDl recB21 recC22 rpsL31 kdgK51 xyl-S mtl-l argE3 thi-i	Transformation	-
Multiple pathways for regulation of σS (RpoS) stability in *Escherichia coli* via the action of multiple anti-adaptors[193]	EG12976 *phoP*::Kan^R	Dissecting the PhoP regulatory network of *Escherichia coli* and *Salmonella enterica*	MG1655	F^-, $lambda^-$, rph-1	P1*vir* transduction	No reference given
Novel Aspects of the Acid Response Network of *E. coli* K-12 Are Revealed by	Δ*phoP* MG1655 phoP::CmR	Novel Aspects of the Acid Response Network of *E. coli* K-12 Are Revealed by a	MG1655		transformation	-

a Study of Transcriptional[265]		Study of Transcriptional				
Phenotypic differences between *Salmonella* and *Escherichia coli* resulting from the disparate regulation of homologous genes[103]	EG13711 *phoP*::Cam^R	Phenotypic differences between *Salmonella* and *Escherichia coli* resulting from the disparate regulation of homologous genes	EG13709	*pmrD*$^+$-*lacZY*$^+$ Km^R	P1*vir* transduction	Experiments in molecular genetics Jeffrey H. Miller
Phenotypic differences between *Salmonella* and *Escherichia coli* resulting from the disparate regulation of homologous genes[103]	EG13729 *phoP*::Kan^R	Phenotypic differences between *Salmonella* and *Escherichia coli* resulting from the disparate regulation of homologous genes	EG13709	*pmrD*$^+$-*lacZY*$^+$ Km^R	P1*vir* transduction	Experiments in molecular genetics Jeffrey H. Miller
Post-Transcriptional Control of the *Escherichia coli* PhoQ-PhoP Two-Component System by Multiple sRNAs Involves a Novel Pairing Region of GcvB[111]	MG1446 *phoP*::Kan^R	MicA sRNA links the PhoP regulon to cell envelope stress	MG1173	MG1655 Δ*lac*X174 λRS*ompT*-*lacZ*	P1*vir* transduction	No reference given
Regulation of Acid Resistance by Connectors of Two-Component Signal Transduction Systems in *Escherichia coli*[266]	JW1116-1 *phoP*790(del):: Kan^R	Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection	BW25113	*rrnB*3 Δ*lacZ*4787 *hsdR*514 Δ(araBAD)567 Δ(rhaBAD)568 rph-1.	Transformation	-
Signal Transduction Cascade between EvgA/EvgS and PhoP/PhoQ Two-	KMP1 *phoP*::Cam^R	Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*	KMY1	MK12 lRS45[Φ(emrK9-lacZ)]	P1*vir* transduction	No reference given

Component Systems of *Escherichia coli*[267]						
Signal Transduction Cascade between EvgA/EvgS and PhoP/PhoQ Two-Component Systems of *Escherichia coli*[267]	KMP2001 *phoP*::Cam^R	Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*	KMY2001	MK12 lRS45[W(emrK9-lacZ)], *evgS*	P1*vir* transduction	No reference given
Stimulus-dependent differential regulation in the *Escherichia coli* PhoQ–PhoP system[246]	TIM8 *phoPQ*::Cam^R	Stimulus-dependent differential regulation in the *Escherichia coli* PhoQ–PhoP system	TIM80	MG1655 Δ*phoQ attHK*::[pTM 27 Δ*Kan*]	P1*vir* transduction	Experiments in molecular genetics Jeffrey H. Miller
Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*[247]	KMP1 *phoP*::Cam^R	Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*	KMY1	MK12 lRS45[Φ(emrK9-lacZ)]	P1*vir* transduction	No reference given
Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*[247]	KMP2001 *phoP*::Cam^R	Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*	KMY2001	MK12 lRS45[W(emrK9-lacZ)], *evgS*	P1*vir* transduction	No reference given
Transcriptome analysis of all two-component regulatory system mutants of *Escherichia coli* K-12[95]	BW27558 (DE*(phoPQ)1244)*	Transcriptome analysis of all two-component regulatory system mutants of *Escherichia coli* K-12	BW25113	*rrnB*3 ΔlacZ4787 *hsdR*514 Δ(araBAD)567 Δ(rhaBAD)568 rph-1.	lambda red recombineering	-
Virulence, resistance to magainin II, and expression of pectate lyase are controlled by	FS1000 *phoP*::Kan^R	Molecular Genetic Analysis of the *Escherichia coli phoP* Locus	MC1061	araD139 Del(araA-leu)7697 Del(lac)X74 *galK*16 *galE*15(GalS) lambda- e14- mcrA0 relA1 rpsL150(strR)	P1*vir* transduction	Experiments in molecular genetics Jeffrey H. Miller

the PhoP-PhoQ two-component regulatory system responding to pH and magnesium in *Erwinia chrysanthemi* 3937[268]				*spoT*1 *mcrB*1 *hsdR*2		

Table E.1. *phoP* mutants generated in published articles, showing the method and recipient strain used to generate them and, for P1*vir* transduction, the original methodology reference (Experiments in molecular genetics Jeffrey H. Miller[183]), if cited in the text. Strain name and genotype are transcribed directly from the citing literature.

Appendix F

Four sequence logos were created, using the aligned sequences generated from the HMMER search using the PhoP and PhoQ sequence from *E. coli* K12-MG1655. HMMER is a tool that searches for protein homologues based on a chosen threshold. Thresholds can be set using either E-values or scores, and in this instance, a threshold of E< 1E-30 was chosen so that all sequences would have at least 67% similarity to the *E. coli* K12-MG1655 sequence. Moreover, this was done to ensure true homologues were found, as several domains and regions in both PhoP and PhoQ, such as the receiver domain of PhoP and the HAMP domain, which are parts of the sensor kinase domain in PhoQ, are similar to other sensor kinases[75].

The search yielded a total of 5,315 and 6,036 homologues for PhoQ and PhoP, respectively, suggesting that PhoQ may be less conserved than PhoP. The majority of these were from the order enterobacteriales, but a minority were from other members of the Gammaproteobacteria class, such as the orders vibrionales and Pseudomadales, specifically *Vibrio navarrensis* and *Pseudomonas stutzeri* (Data not shown). In total, 2,022 and 2,012 of the search results for PhoQ and PhoP, respectively, were from the *Escherichia* genus, the majority of which were *E. coli* isolates.

PhoP (Figure F.2) is more conserved than PhoQ (Figure F.4), with a greater degree of variation present in the HAMP domain of PhoQ homologues. However, like PhoP (Figure F.3), PhoQ (Figure F.5) is quite conserved among *E. coli* isolates.

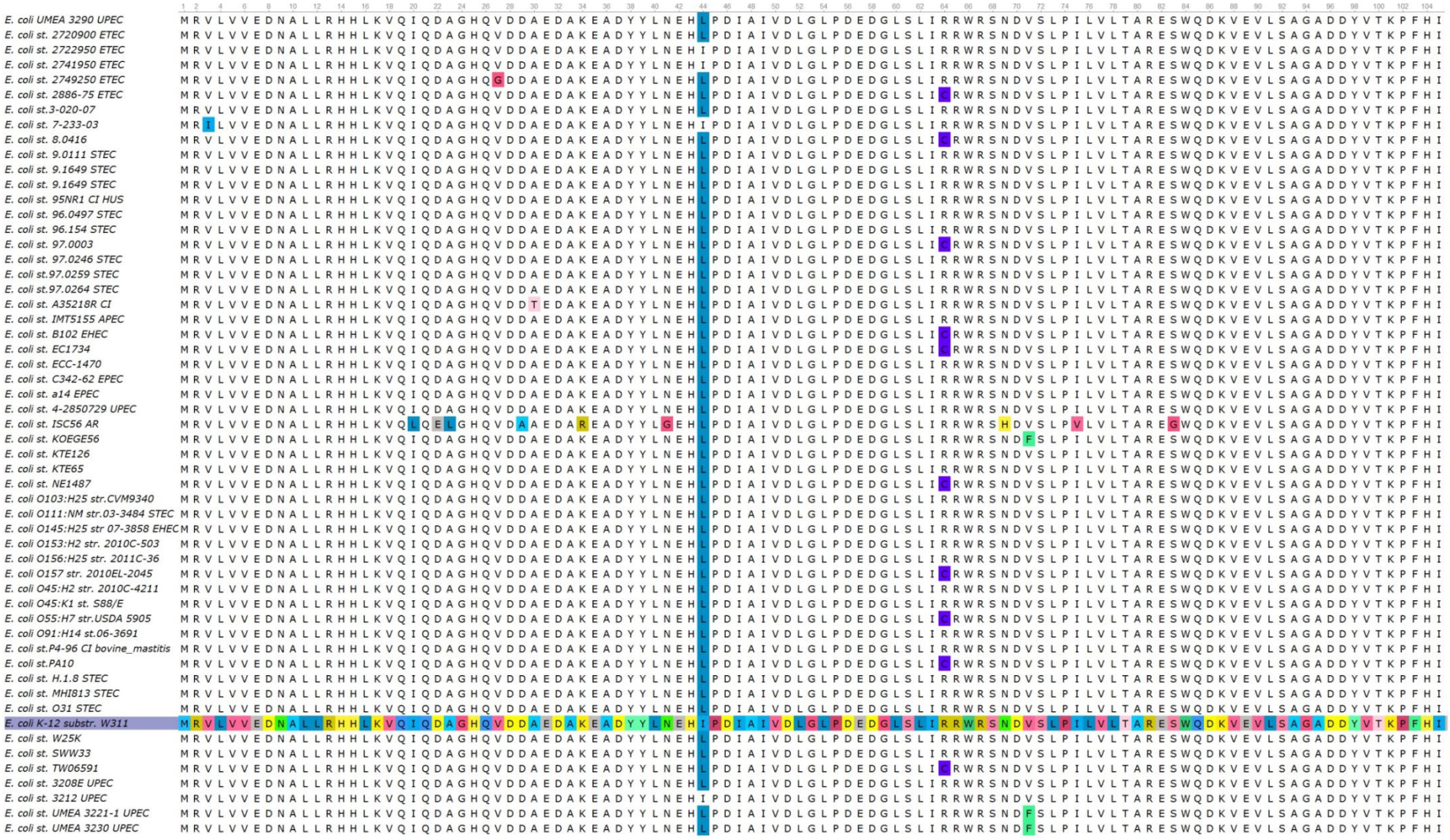


Figure F.1. Curated list of PhoP homologs found in different *E. coli* isolates(1aa-105aa). Aligned sequences generated from a HMMER analysis of PhoP sequences from MG1655, representing the sequence conservation of the 223 residue PhoP sequence for *E. coli* isolates with an E-value <1E-30 in the Ensembl bacteria database.

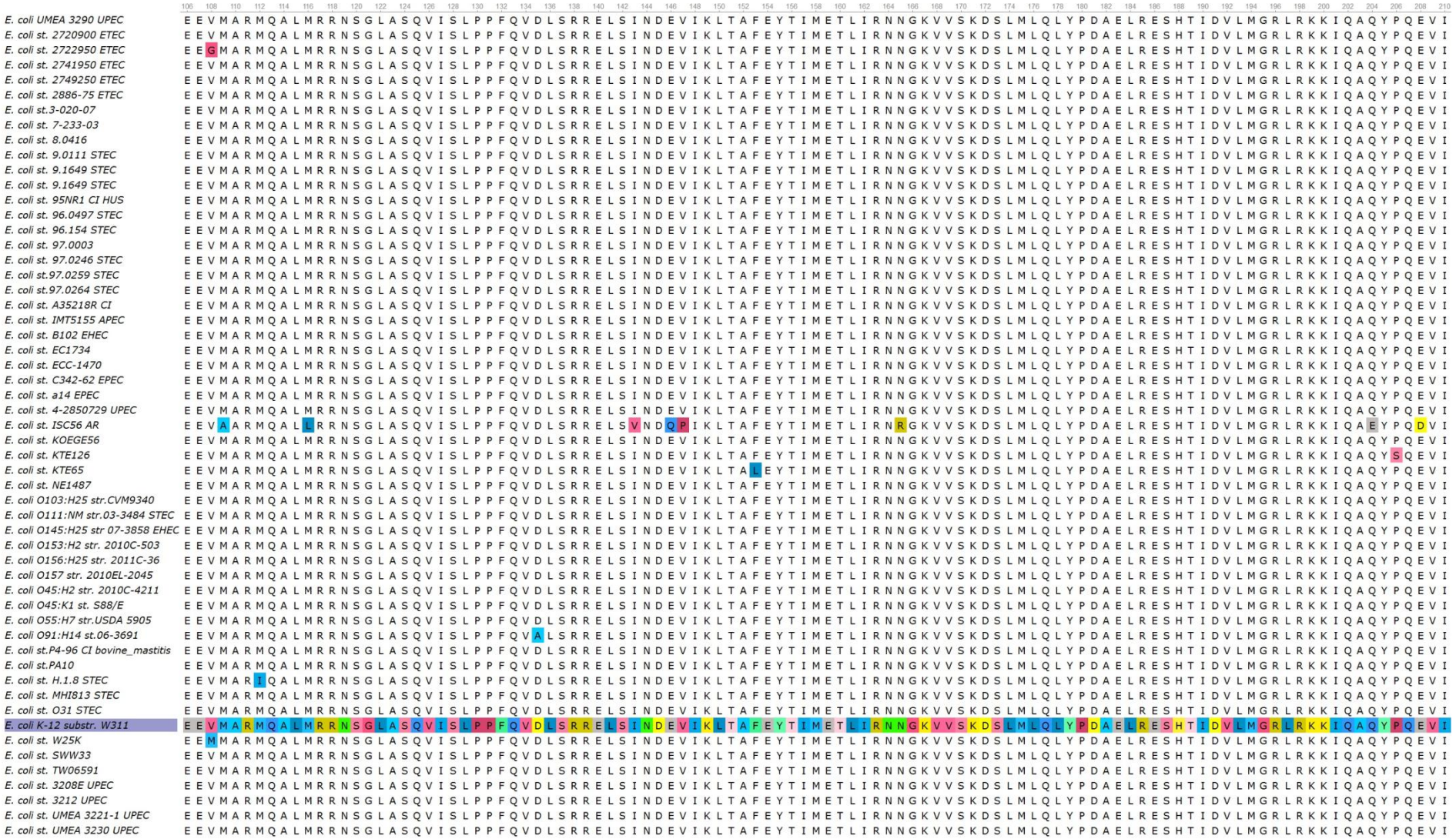


Figure F.1 cont (106aa-223aa). Curated list of PhoP homologs found in different *E. coli* isolates. Aligned sequences generated from a HMMER analysis of PhoP sequences from MG1655, representing the sequence conservation of the 223 residue PhoP sequence for *E. coli* isolates with an E-value <1E-30 in the Ensembl bacteria database.

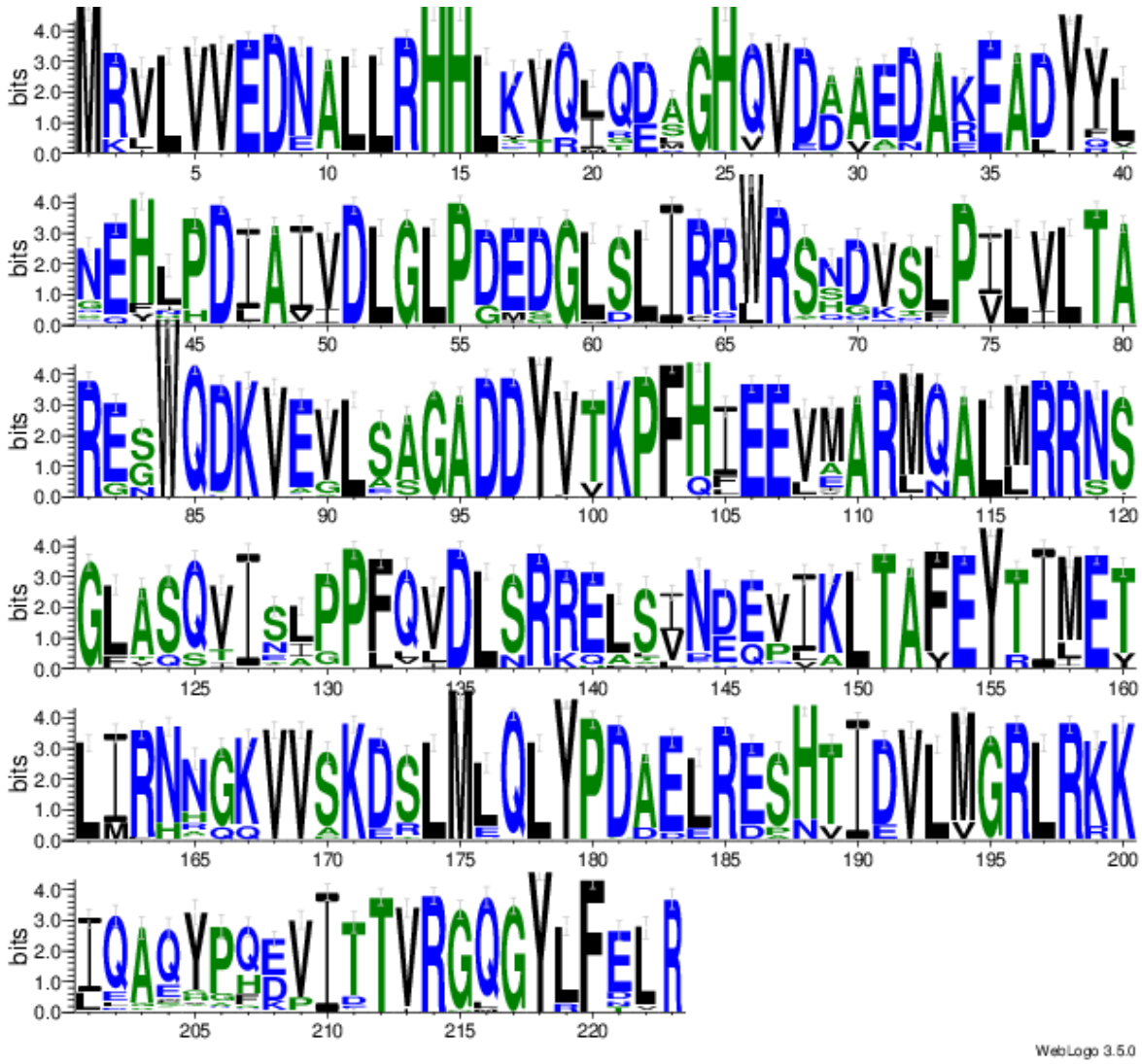


Figure F.2. Sequence logo created from aligned sequences generated from a HMMER analysis of PhoP sequences from MG1655, representing the sequence conservation of the 223 residue PhoP sequence for all PhoP homologues with an E-value <1E-30 in the Ensembl bacteria database. HMMER analysis was performed using tools available at ebi.ac.uk/Tools/hmmer, using an E-value threshold of 1E-30.

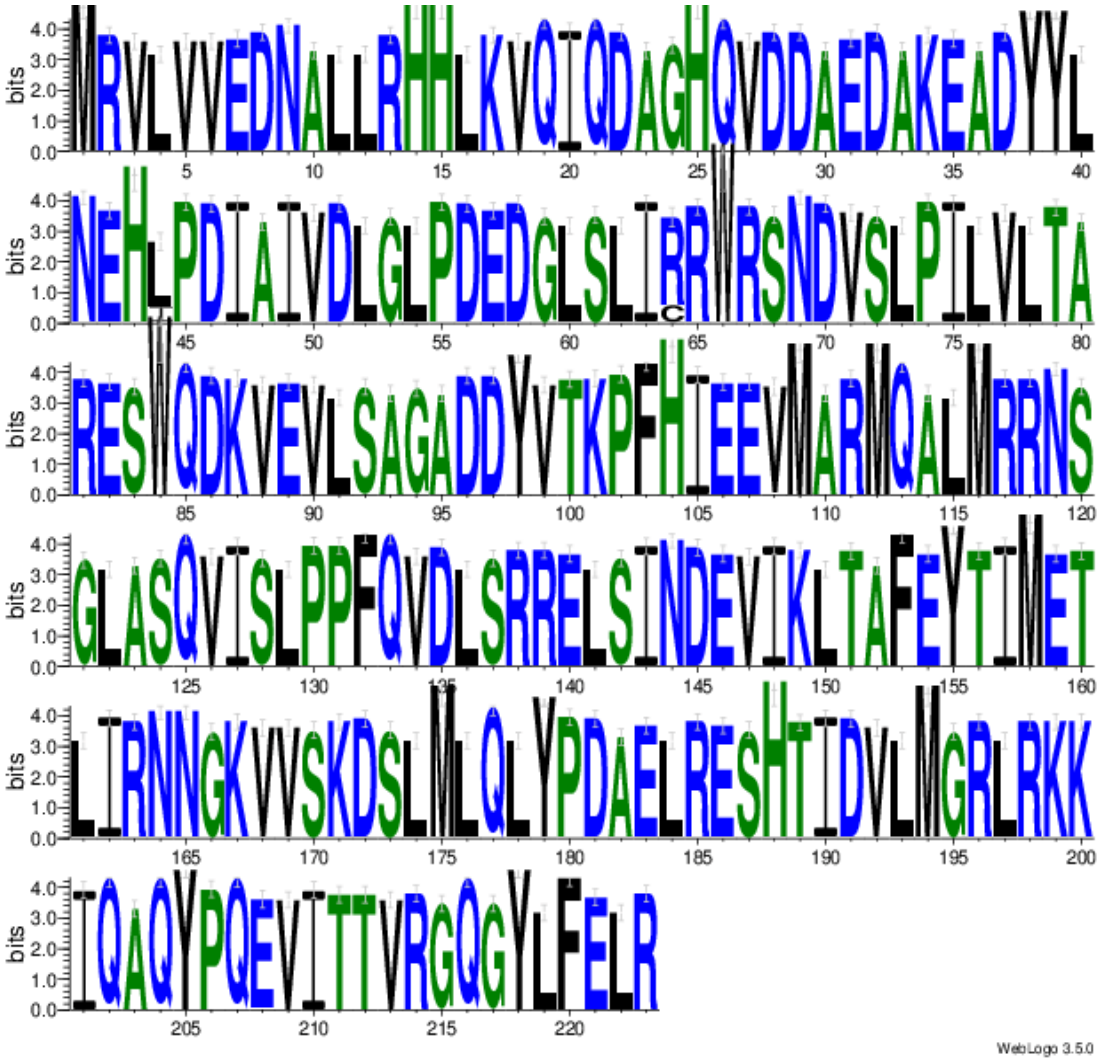


Figure F.3. Sequence logo created from aligned sequences generated from a HMMER analysis of PhoP sequences from MG1655, representing the sequence conservation of the 223 residue PhoP sequence for *E. coli* isolates with an E-value <1E-30 in the Ensembl bacteria database. HMMER analysis was performed using tools available at ebi.ac.uk/Tools/hmmer, using an E-value threshold of 1E-30.

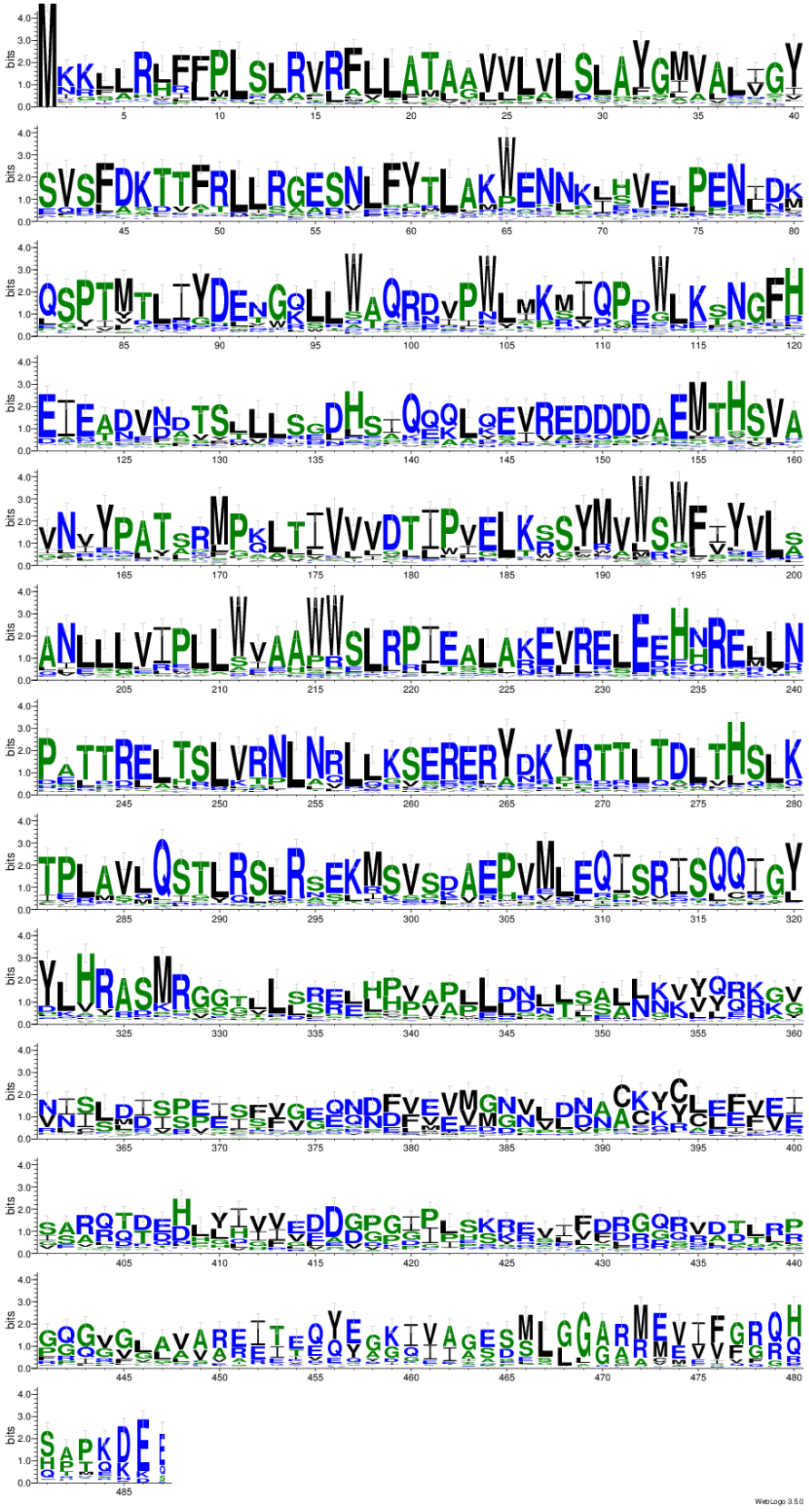


Figure F.4. Sequence logo created from aligned sequences generated from a HMMER analysis of PhoQ sequences from MG1655, representing the sequence conservation of the 486 residue PhoQ sequence for all PhoQ homologues with an E-value <1E-30 in the Ensembl bacteria database. HMMER analysis was performed using tools available at ebi.ac.uk/Tools/hmmer, using an E-value threshold of 1E-30.

Figure F.5. Sequence logo created from aligned sequences generated from a HMMER analysis of PhoQ sequences from MG1655, representing the sequence conservation of the 486 residue PhoQ sequence for *E. coli* isolates with an E-value <1E-30 in the Ensembl bacteria database. HMMER analysis was performed using tools available at ebi.ac.uk/Tools/hmmer, using an E-value threshold of 1E-30.

References

[1] S. Bouzari, A. Jafari, and M. M. Aslani, “Escherichia coli: A brief review of diarrheagenic pathotypes and their role in diarrheal diseases in Iran,” *Iranian Journal of Microbiology*, vol. 4, no. 3. pp. 102–117, 2012.

[2] E. A. Groisman and M. D. Winfiel, “Role of Nonhost Enviroments in the Lifestyles of *Salmonella* and *E. coli*,” *Appl. Environ. Mcrobiology*, vol. 69, no. 7, pp. 3687–3694, 2003.

[3] M. S. Donnenberg, *Escherichia coli: Pathotypes and Principles of Pathogenesis: Second Edition*. 2013.

[4] S. Bekal, R. Brousseau, L. Masson, G. Prefontaine, J. Fairbrother, and J. Harel, “Rapid identification of Escherichia coli pathotypes by virulence gene detection with DNA microarrays,” *J. Clin. Microbiol.*, vol. 41, no. 5, pp. 2113–2125, 2003.

[5] T. D. Ho and M. K. Waldor, “Enterohemorrhagic Escherichia coli O157:H7 gal mutants are sensitive to bacteriophage P1 and defective in intestinal colonization,” *Infect. Immun.*, vol. 75, no. 4, pp. 1661–1666, 2007.

[6] W. Deng *et al.*, “Assembly, structure, function and regulation of type III secretion systems,” *Nature Reviews Microbiology*, vol. 15, no. 6. pp. 323–337, 2017.

[7] Z. D. Blount, “The unexhausted potential of *E. coli*,” *Elife*, vol. 4, pp. 1–12, 2015.

[8] T. Ferenci *et al.*, “Genomic sequencing reveals regulatory mutations and recombinational events in the widely used MC4100 lineage of Escherichia coli K-12,” *J. Bacteriol.*, vol. 191, no. 12, pp. 4025–4029, 2009.

[9] L. Brown, J. M. Wolf, R. Prados-Rosales, and A. Casadevall, “Through the wall: extracellular vesicles in Gram-positive bacteria, mycobacteria and fungi.,” *Nat. Rev. Microbiol.*, vol. 13, no. 10, pp. 620–30, 2015.

[10] A. Typas, M. Banzhaf, C. a. Gross, and W. Vollmer, “From the regulation of peptidoglycan synthesis to bacterial growth and morphology,” *Nat. Rev. Microbiol.*, vol. 10, no. 2, pp. 123–136, 2012.

[11] H. Remaut and R. Fronzes, *Bacterial membranes : structural and molecular biology*. .

[12] K. C. Huang, R. Mukhopadhyay, B. Wen, Z. Gitai, and N. S. Wingreen, “Cell shape and cell-wall organization in Gram-negative bacteria.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 49, pp. 19282–7, 2008.

[13] A. Delcour, “Outer membrane permeability and antibiotic resistance.,” *Biochim. Biophys. Acta*, vol. 1794, no. 5, pp. 808–816, 2009.

[14] J. B. Parsons and C. O. Rock, “Bacterial lipids: Metabolism and membrane homeostasis,” *Progress in Lipid Research*, vol. 52, no. 3. pp. 249–276, 2013.

[15] J. E. Cronan, “Bacterial Membrane Lipids: Where Do We Stand?,” *Annu. Rev. Microbiol.*, vol. 57, no. 1, pp. 203–224, 2003.

[16] A. Van Dalen and B. De Kruijff, “The role of lipids in membrane insertion and translocation of bacterial proteins,” *Biochimica et Biophysica Acta - Molecular Cell Research*, vol. 1694, no. 1–3 SPEC.ISS. pp. 97–109, 2004.

[17] S. Vanounou, A. H. Parola, and I. Fishov, “Phosphatidylethanolamine and phosphatidylglycerol are segregated into different domains in bacterial membrane. A study with pyrene-labelled phospholipids,” *Mol. Microbiol.*, vol. 49, no. 4, pp. 1067–1079, 2003.

[18] C. Xu, X. Lin, H. Ren, Y. Zhang, S. Wang, and X. Peng, “Analysis of outer membrane proteome of Escherichia coli related to resistance to ampicillin and tetracycline,” in *Proteomics*, 2006, vol. 6, no. 2, pp. 462–473.

[19] J. H. Weiner and L. Li, “Proteome of the Escherichia coli envelope and technological challenges in membrane proteome analysis,” *Biochimica et Biophysica Acta - Biomembranes*, vol. 1778, no. 9. pp. 1698–1713, 2008.

[20] J. H. Weiner and R. A. Rothery, “Bacterial Cytoplasmic Membrane,” in *Encyclopedia of Life Sciences*, 2007.

[21] A. Bouhss, A. E. Trunkfield, T. D. H. Bugg, and D. Mengin-Lecreulx, “The biosynthesis of peptidoglycan lipid-linked intermediates,” *FEMS Microbiology Reviews*, vol. 32, no. 2. pp. 208–233, 2008.

[22] J. Mecsas, P. E. Rouviere, J. W. Erickson, T. J. Donohue, and C. A. Gross, “The activity of sigma E, an Escherichia coli heat-inducible sigma-factor, is modulated by expression of outer membrane proteins.,” *Genes Dev.*, vol. 7, no. 12B, pp. 2618–28, Dec. 1993.

[23] A. H. Delcour, “Outer membrane permeability and antibiotic resistance.,” *Biochim. Biophys. Acta*, vol. 1794, no. 5, pp. 808–16, 2009.

[24] M. P. Bos, V. Robert, and J. Tommassen, “Biogenesis of the gram-negative bacterial outer membrane.,” *Annu. Rev. Microbiol.*, vol. 61, pp. 191–214, 2007.

[25] H. Nikaido, “Molecular basis of bacterial outer membrane permeability revisited.,” *Microbiol. Mol. Biol. Rev.*, vol. 67, no. 4, pp. 593–656, 2003.

[26] H. Killmann, R. Benz, and V. Braun, “Properties of the FhuA channel in the Escherichia coli outer membrane after deletion of FhuA portions within and outside the predicted gating loop,” *J. Bacteriol.*, vol. 178, no. 23, pp. 6913–6920, 1996.

[27] J. L. Thomassin, J. R. Brannon, B. F. Gibbs, S. Gruenheid, and H. Le Moual, “OmpT outer membrane proteases of enterohemorrhagic and enteropathogenic Escherichia coli contribute differently to the degradation of human LL-37,” *Infect. Immun.*, vol. 80, no. 2, pp. 483–492, 2012.

[28] G. B. Melikyan, S. A. Brener, D. C. Ok, and F. S. Cohen, “Inner but not outer membrane leaflets control the transition from glycosylphosphatidylinositol-anchored influenza hemagglutinin-induced hemifusion to full fusion,” *J. Cell Biol.*, vol. 136, no. 5, pp. 995–1005, 1997.

[29] H. Mori and K. Ito, “The Sec protein-translocation pathway,” *Trends in Microbiology*, vol. 9, no. 10. pp. 494–500, 2001.

[30] Y. Liu, X. Fu, J. Shen, H. Zhang, W. Hong, and Z. Chang, “Periplasmic proteins of Escherichia coli are highly resistant to aggregation: Reappraisal for roles of molecular chaperones in periplasm,” *Biochem. Biophys. Res. Commun.*, vol. 316, no. 3, pp. 795–801, 2004.

[31] S. Behrens-Kneip, “The role of SurA factor in outer membrane protein transport and virulence,” *International Journal of Medical Microbiology*, vol. 300, no. 7. pp. 421–428, 2010.

[32] J. Schwalm, T. F. Mahoney, G. R. Soltes, and T. J. Silhavy, “Role for Skp in LptD assembly in Escherichia coli,” *J. Bacteriol.*, vol. 195, no. 16, pp. 3734–3742, 2013.

[33] D. Huber and B. Bukau, “DegP: a Protein ‘Death Star,’” *Structure*, vol. 16, no. 7. pp. 989–990, 2008.

[34] D. P. Ricci and T. J. Silhavy, “The Bam machine: A molecular cooper,” *Biochimica et Biophysica Acta - Biomembranes*, vol. 1818, no. 4. pp. 1067–1084, 2012.

[35] T.-W. Chang, Y.-M. Lin, C.-F. Wang, and Y.-D. Liao, “Outer membrane lipoprotein Lpp is Gram-negative bacterial cell surface receptor for cationic antimicrobial peptides.,” *J. Biol. Chem.*, vol. 287, no. 1, pp. 418–28, 2012.

[36] R. A. Kramer *et al.*, “Lipopolysaccharide regions involved in the activation of Escherichia coli outer membrane protease OmpT.,” *Eur. J. Biochem.*, vol. 269, no. 6, pp. 1746–52, 2002.

[37] X. Wang and P. J. Quinn, “Lipopolysaccharide: Biosynthetic pathway and structure modification,” *Progress in Lipid Research*, vol. 49, no. 2. pp. 97–107, 2010.

[38] O. Holst, “The structures of core regions from enterobacterial lipopolysaccharides - An update,” *FEMS Microbiology Letters*, vol. 271, no. 1. pp. 3–11, 2007.

[39] K. Moon, D. A. Six, H. J. Lee, C. R. H. Raetz, and S. Gottesman, “Complex transcriptional and post-transcriptional regulation of an enzyme for lipopolysaccharide modification,” *Mol. Microbiol.*, vol. 89, no. 1, pp. 52–64, 2013.

[40] N. Kayagaki *et al.*, “Noncanonical Inflammasome Activation by Intracellular LPS Independent of TLR4,” *Science (80-.).*, vol. 341, no. 6151, pp. 1246–1249, 2013.

[41] J. Van Heijenoort, “Biosynthesis of Bacterial Peptidoglycan,” in *Microbial Glycobiology*, 2010, pp. 285–304.

[42] S. Minagawa *et al.*, “Identification and molecular characterization of the Mg2+ stimulon of Escherichia coli.,” *J. Bacteriol.*, vol. 185, pp. 3696–3702, 2003.

[43] P. Romby, F. Vandenesch, and E. G. H. Wagner, “The role of RNAs in the regulation of virulence-gene expression,” *Current Opinion in Microbiology*, vol. 9, no. 2. pp. 229–236, 2006.

[44] K. Chen and N. Rajewsky, “The evolution of gene regulation by transcription factors and microRNAs.,” *Nat. Rev. Genet.*, vol. 8, no. 2, pp. 93–103, 2007.

[45] B. M. Emerson, “Specificity of gene regulation,” *Cell*, vol. 109, no. 3. pp. 267–270, 2002.

[46] M. Vital, B. Chai, B. Ostman, J. Cole, K. T. Konstantinidis, and J. M. Tiedje, “Gene

expression analysis of E. coli strains provides insights into the role of gene regulation in diversification.," *ISME J.*, pp. 1–11, 2014.

[47] J. L. Keck, D. D. Roche, a S. Lynch, and J. M. Berger, "Structure of the RNA polymerase domain of E. coli primase.," *Science*, vol. 287, no. 5462, pp. 2482–2486, 2000.

[48] J. E. Mitchell, D. Zheng, S. J. W. Busby, and S. D. Minchin, "Identification and analysis of 'extended -10' promoters in Escherichia coli," *Nucleic Acids Res.*, vol. 31, no. 16, pp. 4689–4695, 2003.

[49] A. S. N. Seshasayee, K. Sivaraman, and N. M. Luscombe, "An Overview of Prokaryotic Transcription Factors," 2011, pp. 7–23.

[50] H. Bremer, P. Dennis, and M. Ehrenberg, "Free RNA polymerase and modeling global transcription in Escherichia coli," *Biochimie*, vol. 85, no. 6. pp. 597–609, 2003.

[51] H. H. Hsu, K. M. Chung, T. C. Chen, and B. Y. Chang, "Role of the Sigma Factor in Transcription Initiation in the Absence of Core RNA Polymerase," *Cell*, vol. 127, no. 2, pp. 317–327, 2006.

[52] D. G. Vassylyev, M. N. Vassylyeva, A. Perederina, T. H. Tahirov, and I. Artsimovitch, "Structural basis for transcription elongation by bacterial RNA polymerase.," *Nature*, vol. 448, no. 7150, pp. 157–162, 2007.

[53] S. A Darst, "Bacterial RNA polymerase," *Curr. Opin. Struct. Biol.*, vol. 11, no. 2, pp. 155–162, 2001.

[54] B. Z. Ring, W. S. Yarnell, and J. W. Roberts, "Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing.," *Cell*, vol. 86, no. 3, pp. 485–493, 1996.

[55] K. Severinov, "RNA polymerase structure-function: Insights into points of transcriptional regulation," *Current Opinion in Microbiology*, vol. 3, no. 2. pp. 118–125, 2000.

[56] H. Cook and D. W. Ussery, "Sigma factors in a thousand E. coli genomes," *Environ. Microbiol.*, vol. 15, no. 12, pp. 3121–3129, 2013.

[57] S. E. Ades, 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.*, vol. 185, no. 8, pp. 2512–2519, 2003.

[58] J. Vogel and K. Papenfort, "Small non-coding RNAs and the bacterial outer membrane," *Current Opinion in Microbiology*, vol. 9, no. 6. pp. 605–611, 2006.

[59] N. Windbichler, F. von Pelchrzim, O. Mayer, E. Csaszar, and R. Schroeder, "Isolation of small RNA-binding proteins from E. coli: evidence for frequent interaction of RNAs with RNA polymerase," *RNA Biol.*, vol. 5, no. 1, pp. 30–40, 2008.

[60] F. Mika and R. Hengge, "Small regulatory RNAs in the control of motility and biofilm formation in E. Coli and Salmonella," *International Journal of Molecular Sciences*, vol. 14, no. 3. pp. 4560–4579, 2013.

[61] A. Filloux, *Bacterial regulatory networks*. Caister Academic Press, 2012.

[62] N. Sedlyarova *et al.*, "sRNA-Mediated Control of Transcription Termination in E. coli," *Cell*, vol. 167, no. 1, p. 111–121.e13, 2016.

[63] L. S. Waters and G. Storz, "Regulatory RNAs in Bacteria," *Cell*, vol. 136, no. 4. pp. 615–628, 2009.

[64] M. Mandal and R. R. Breaker, "Gene regulation by riboswitches.," *Nat. Rev. Mol. Cell Biol.*, vol. 5, no. 6, pp. 451–463, 2004.

[65] M. M. Babu and S. A. Teichmann, "Evolution of transcription factors and the gene regulatory network in Escherichia coli," *Nucleic Acids Research*, vol. 31, no. 4. pp. 1234–1244, 2003.

[66] G. Karp, *Cell and molecular biology : concepts and experiments*. John Wiley, 2010.

[67] A. Ishihama, "A Revolutionary Paradigm of Bacterial Genome Regulation," in *Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria*, vol. 1, 2016, pp. 23–36.

[68] S. A. F. T. van Hijum, M. H. Medema, and O. P. Kuipers, "Mechanisms and evolution of control logic in prokaryotic transcriptional regulation.," *Microbiol. Mol. Biol. Rev.*, vol. 73, no. 3, p. 481–509, Table of Contents, 2009.

[69] M. M. Babu, *Bacterial gene regulation and transcriptional networks*. .

[70] A. Martínez-Antonio, S. C. Janga, H. Salgado, and J. Collado-Vides, "Internal-sensing machinery directs the activity of the regulatory network in Escherichia coli," *Trends in Microbiology*, vol. 14, no. 1. pp. 22–27, 2006.

[71] E. Balleza *et al.*, "Regulation by transcription factors in bacteria: Beyond description," *FEMS Microbiology Reviews*, vol. 33, no. 1. pp. 133–151, 2009.

[72] T. R. Hughes, "Introduction to 'a handbook of transcription factors,'" *Subcell. Biochem.*, vol. 52, pp. 1–6, 2011.

[73] T. R. Sorrells and A. D. Johnson, "Making sense of transcription networks," *Cell*, vol. 161, no. 4, pp. 714–723, 2015.

[74] A. Y. A. Mitrophanov and E. E. a Groisman, "Signal integration in bacterial two-component regulatory systems.," *Genes Dev.*, vol. 22, no. 19, pp. 2601–2611, 2008.

[75] S. I. Miller, A. M. Kukral, and J. J. Mekalanos, "A two-component regulatory system (phoPphoQ) controls Salmonella typhimurium virulence," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 86, no. 13, pp. 5054–5058, 1989.

[76] E. A. Groisman, "The pleiotropic two-component regulatory system PhoP-PhoQ," *Journal of Bacteriology*, vol. 183, no. 6. pp. 1835–1842, 2001.

[77] A. M. Stock, V. L. Robinson, and P. N. Goudreau, "Two-component signal transduction," *Annu. Rev. Biochem.*, vol. 69, pp. 183–215, 2000.

[78] M. Y. Galperin, A. N. Nikolskaya, and E. V. Koonin, "Novel domains of the prokaryotic two-component signal transduction systems," *FEMS Microbiology Letters*, vol. 203, no. 1. pp. 11–21, 2001.

[79] M. T. Laub, E. G. Biondi, and J. M. Skerker, “Phosphotransfer Profiling: Systematic Mapping of Two-Component Signal Transduction Pathways and Phosphorelays,” *Methods in Enzymology*, vol. 423. pp. 531–548, 2007.

[80] G. E. Schaller, S. H. Shiu, and J. P. Armitage, “Two-component systems and their co-option for eukaryotic signal transduction,” *Current Biology*, vol. 21, no. 9. 2011.

[81] J. J. E. Bijlsma and E. A. Groisman, “Making informed decisions: Regulatory interactions between two-component systems,” *Trends in Microbiology*, vol. 11, no. 8. pp. 359–366, 2003.

[82] S. Wang, “Bacterial Two-Component Systems : Structures and Signaling Mechanisms,” *Protein Phosphorylation Hum. Heal. Dr. Cai Huang*, no. Figure 1, pp. 439–466, 2012.

[83] R. B. Bourret, “Receiver domain structure and function in response regulator proteins,” *Current Opinion in Microbiology*, vol. 13, no. 2. pp. 142–149, 2010.

[84] M. Y. Galperin, “Structural classification of bacterial response regulators: Diversity of output domains and domain combinations,” *J. Bacteriol.*, vol. 188, no. 12, pp. 4169–4182, 2006.

[85] H. Itou and I. Tanaka, “The OmpR-family of proteins: insight into the tertiary structure and functions of two-component regulator proteins.,” *J. Biochem.*, vol. 129, no. 3, pp. 343–350, 2001.

[86] U. Romling, M. Y. Galperin, and M. Gomelsky, “Cyclic di-GMP: the first 25 years of a universal bacterial second messenger.,” *Microbiol. Mol. Biol. Rev.*, vol. 77, no. 1, pp. 1–52, 2013.

[87] P. Bachhawat and A. M. Stock, “Crystal structures of the receiver domain of the response regulator PhoP from Escherichia coli in the absence and presence of the phosphoryl analog beryllofiuoride,” *J. Bacteriol.*, vol. 189, no. 16, pp. 5987–5995, 2007.

[88] A. Kato, H. Tanabe, and R. Utsumi, “Molecular characterization of the PhoP-PhoQ two-component system in Escherichia coli K-12: Identification of extracellular Mg2+-responsive promoters,” *J. Bacteriol.*, vol. 181, no. 17, pp. 5516–5520, 1999.

[89] T. Lemmin, C. S. Soto, G. Clinthorne, W. F. DeGrado, and M. Dal Peraro, “Assembly of the Transmembrane Domain of E. coli PhoQ Histidine Kinase: Implications for Signal Transduction from Molecular Simulations,” *PLoS Comput. Biol.*, vol. 9, no. 1, 2013.

[90] C. J. Alteri, J. R. Lindner, D. J. Reiss, S. N. Smith, and H. L. T. Mobley, “The broadly conserved regulator PhoP links pathogen virulence and membrane potential in Escherichia coli,” *Mol. Microbiol.*, vol. 82, no. 1, pp. 145–163, 2011.

[91] H. Mulcahy, L. Charron-Mazenod, and S. Lewenza, “Extracellular DNA chelates cations and induces antibiotic resistance in Pseudomonas aeruginosa biofilms,” *PLoS Pathog.*, vol. 4, no. 11, 2008.

[92] S. B. Walters, E. Dubnau, I. Kolesnikova, F. Laval, M. Daffe, and I. Smith, “The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis,” *Mol. Microbiol.*, vol. 60, no.

2, pp. 312–330, 2006.

[93] D. Destoumieux-Garzón, M. Duperthuy, A. Vanhove, P. Schmitt, and S. Wai, “Resistance to Antimicrobial Peptides in Vibrios,” *Antibiotics*, vol. 3, no. 4, pp. 540–563, 2014.

[94] P. Monsieurs *et al.*, “Comparison of the PhoPQ regulon in Escherichia coli and Salmonella typhimurium.,” *J. Mol. Evol.*, vol. 60, pp. 462–474, 2005.

[95] T. Oshima *et al.*, “Transcriptome analysis of all two-component regulatory system mutants of Escherichia coli K-12,” *Mol Microbiol*, vol. 46, no. 1, pp. 281–291, 2002.

[96] I. Zwir *et al.*, “Dissecting the PhoP regulatory network of Escherichia coli and Salmonella enterica,” *Proc Natl Acad Sci U S A*, vol. 102, no. 8, pp. 2862–2867, 2005.

[97] Z. Ma, N. Masuda, and J. W. Foster, “Characterization of EvgAS-YdeO-GadE branched regulatory circuit governing glutamate-dependent acid resistance in Escherichia coli,” *J. Bacteriol.*, vol. 186, no. 21, pp. 7378–7389, 2004.

[98] Y. Eguchi *et al.*, “Signal transduction cascade between EvgA/EvgS and PhoP/PhoQ two-component systems of Escherichia coli.,” *J. Bacteriol.*, vol. 186, pp. 3006–3014, 2004.

[99] Y. Eguchi *et al.*, “B1500, a small membrane protein, connects the two-component systems EvgS/EvgA and PhoQ/PhoP in Escherichia coli.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 47, pp. 18712–18717, 2007.

[100] L. Guo *et al.*, “Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides,” *Cell*, vol. 95, no. 2, pp. 189–198, 1998.

[101] E. J. Rubin, C. M. Herrera, A. A. Crofts, and M. S. Trent, “PmrD is required for modifications to escherichia coli endotoxin that promote antimicrobial resistance,” *Antimicrob. Agents Chemother.*, vol. 59, no. 4, pp. 2051–2061, 2015.

[102] E. A. Groisman, F. Heffron, and F. Solomon, “Molecular genetic analysis of the Escherichia coli phoP locus,” *J. Bacteriol.*, vol. 174, no. 2, pp. 486–491, 1992.

[103] M. D. Winfield and E. a Groisman, “Phenotypic differences between Salmonella and Escherichia coli resulting from the disparate regulation of homologous genes.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 49, pp. 17162–7, 2004.

[104] R. C. Charles *et al.*, “Comparative proteomic analysis of the PhoP regulon in Salmonella enterica Serovar Typhi versus Typhimurium,” *PLoS One*, vol. 4, no. 9, 2009.

[105] R. T. Gill, M. P. DeLisa, M. Shiloach, T. R. Holoman, and W. E. Bentley, “OmpT expression and activity increase in response to recombinant chloramphenicol acetyltransferase overexpression and heat shock in E. coli.,” *J. Mol. Microbiol. Biotechnol.*, vol. 2, no. 3, pp. 283–9, 2000.

[106] N. Masuda and G. M. Church, “Escherichia coli gene expression responsive to levels of the response regulator EvgA,” *J Bacteriol*, vol. 184, no. 22, p. 6225–34., 2002.

[107] Y. Jin, R. M. Watt, A. Danchin, and J. Huang, “Small noncoding RNA GcvB is a novel regulator of acid resistance in Escherichia coli.,” *BMC Genomics*, vol. 10, p. 165,

2009.

[108] M. S. Pittman, H. C. Robinson, and R. K. Poole, “A bacterial glutathione transporter (Escherichia coli CydDC) exports reductant to the periplasm,” *J. Biol. Chem.*, vol. 280, no. 37, pp. 32254–32261, 2005.

[109] D. Niegowski and S. Eshaghi, “The CorA family: Structure and function revisited,” *Cellular and Molecular Life Sciences*, vol. 64, no. 19–20. pp. 2564–2574, 2007.

[110] E. A. Groisman *et al.*, “Bacterial Mg2+ homeostasis, transport, and virulence.,” *Annu. Rev. Genet.*, vol. 47, no. 102, pp. 625–46, 2013.

[111] A. Coornaert, C. Chiaruttini, M. Springer, and M. Guillier, “Post-Transcriptional Control of the Escherichia coli PhoQ-PhoP Two-Component System by Multiple sRNAs Involves a Novel Pairing Region of GcvB,” *PLoS Genet.*, vol. 9, no. 1, 2013.

[112] A. Coornaert, A. Lu, P. Mandin, M. Springer, S. Gottesman, and M. Guillier, “MicA sRNA links the PhoP regulon to cell envelope stress.,” *Mol. Microbiol.*, vol. 76, pp. 467–479, 2010.

[113] M. E. Salazar, A. I. Podgornaia, and M. T. Laub, “The small membrane protein MgrB regulates PhoQ bifunctionality to control PhoP target gene expression dynamics,” *Mol. Microbiol.*, vol. 102, no. 3, pp. 430–445, 2016.

[114] P. E. Rouvière, a De Las Peñas, J. Mecsas, C. Z. Lu, K. E. Rudd, and C. a Gross, “rpoE, the gene encoding the second heat-shock sigma factor, sigma E, in Escherichia coli.,” *The EMBO journal*, vol. 14, no. 5. pp. 1032–1042, 1995.

[115] H. J. Chung, W. Bang, and M. A. Drake, “Stress response of Escherichia coli,” *Comprehensive Reviews in Food Science and Food Safety*, vol. 5, no. 3. pp. 52–64, 2006.

[116] E. Klauck, A. Typas, and R. Hengge, “The sigmaS subunit of RNA polymerase as a signal integrator and network master regulator in the general stress response in Escherichia coli.,” *Sci. Prog.*, vol. 90, no. Pt 2-3, pp. 103–27, 2007.

[117] R. G. Groat, J. E. Schultz, E. Zychlinsky, A. Bockman, and A. Matin, “Starvation proteins in Escherichia coli: Kinetics of synthesis and role in starvation survival,” *J. Bacteriol.*, vol. 168, no. 2, pp. 486–493, 1986.

[118] F. Arsène, T. Tomoyasu, and B. Bukau, “The heat shock response of Escherichia coli.,” *Int. J. Food Microbiol.*, vol. 55, no. 1–3, pp. 3–9, 2000.

[119] D. A. Parsell and R. T. Sauer, “Induction of a heat shock-like response by unfolded protein in Escherichia coli: dependence on protein level not protein degradation.,” *Genes Dev.*, vol. 3, no. 8, pp. 1226–1232, 1989.

[120] R. M. Vabulas, S. Raychaudhuri, M. Hayer-Hartl, and F. U. Hartl, “Protein folding in the cytoplasm and the heat shock response.,” *Cold Spring Harbor perspectives in biology*, vol. 2, no. 12. 2010.

[121] D. B. Straus, W. A. Walter, and C. A. Gross, “The heat shock response of E. coli is regulated by changes in the concentration of σ32,” *Nature*, vol. 329, no. 6137, pp. 348–351, 1987.

[122] D. E. Jenkins, E. A. Auger, and A. Matin, “Role of RpoH, a heat shock regulator protein, in Escherichia coli carbon starvation protein synthesis and survival,” *J. Bacteriol.*, vol. 173, no. 6, pp. 1992–1996, 1991.

[123] P. K. Sorger, “Heat shock factor and the heat shock response,” *Cell*, vol. 65, no. 1990, pp. 363–366, 1991.

[124] K. Yamanaka, “Cold shock response in Escherichia coli.,” *J. Mol. Microbiol. Biotechnol.*, vol. 1, no. 2, pp. 193–202, 1999.

[125] C. J. Dorman, “DNA supercoiling and bacterial gene expression.,” *Sci. Prog.*, vol. 89, no. Pt 3-4, pp. 151–66, 2006.

[126] M. H. Weber and M. A. Marahiel, “Bacterial cold shock responses.,” *Science progress*, vol. 86, no. Pt 1-2. pp. 9–75, 2003.

[127] K. Yamanaka and M. Inouye, “Induction of CspA, an E. coli major cold-shock protein, upon nutritional upshift at 37°C,” *Genes to Cells*, vol. 6, no. 4, pp. 279–290, 2001.

[128] A. M. Giuliodori, A. Brandi, C. O. Gualerzi, and C. L. Pon, “Preferential translation of cold-shock mRNAs during cold adaptation.,” *RNA*, vol. 10, no. 2, pp. 265–276, 2004.

[129] T. L. Raivio, “Envelope stress responses and Gram-negative bacterial pathogenesis,” *Molecular Microbiology*, vol. 56, no. 5. pp. 1119–1128, 2005.

[130] R. G. Raffa and T. L. Raivio, “A third envelope stress signal transduction pathway in Escherichia coli,” *Mol. Microbiol.*, vol. 45, no. 6, pp. 1599–1611, 2002.

[131] A. J. Darwin, “The phage-shock-protein response,” *Molecular Microbiology*, vol. 57, no. 3. pp. 621–628, 2005.

[132] T. L. Raivio and T. J. Silhavy, “The σE and Cpx regulatory pathways: Overlapping but distinct envelope stress responses,” *Curr. Opin. Microbiol.*, vol. 2, no. 2, pp. 159–165, Apr. 1999.

[133] P. N. Danese and T. J. Silhavy, “The σ(E) and the Cpx signal transduction systems control the synthesis of periplasmic protein-folding enzymes in Escherichia coli,” *Genes Dev.*, vol. 11, no. 9, pp. 1183–1193, 1997.

[134] A. De Las Peñas, L. Connolly, and C. A. Gross, “SigmaE is an essential sigma factor in Escherichia coli.,” *J. Bacteriol.*, vol. 179, no. 21, pp. 6862–4, 1997.

[135] J. E. Button, T. J. Silhavy, and N. Ruiz, “A suppressor of cell death caused by the loss of sigmaE downregulates extracytoplasmic stress responses and outer membrane vesicle production in Escherichia coli.,” *J. Bacteriol.*, vol. 189, no. 5, pp. 1523–30, 2007.

[136] B. M. Alba and C. A. Gross, “Regulation of the Escherichia coli sigma-dependent envelope stress response.,” *Mol. Microbiol.*, vol. 52, no. 3, pp. 613–619, 2004.

[137] D. Missiakas, M. P. Mayer, M. Lemaire, C. Georgopoulos, and S. Raina, “Modulation of the Escherichia coli sigmaE (RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins.,” *Mol. Microbiol.*, vol. 24, no. 2, pp. 355–371, 1997.

[138] J. D. Hayden and S. E. Ades, “The extracytoplasmic stress factor, σE, is required to maintain cell envelope integrity in Escherichia coli,” *PLoS One*, vol. 3, no. 2, 2008.

[139] J. T. Wade *et al.*, “Extensive functional overlap between σ factors in Escherichia coli,” *Nat. Struct. Mol. Biol.*, vol. 13, no. 9, pp. 806–814, 2006.

[140] H. H. Weber, T. T. Polen, J. J. Heuveling, V. F. V. F. Wendisch, and R. Hengge-Aronis, “Genome-wide analysis of the general stress response network in Escherichia coli: sigmaS-dependent genes, promoters, and sigma factor selectivity.,” *J. Bacteriol.*, vol. 187, no. 5, pp. 1591–1603, 2005.

[141] A. Farewell, K. Kvint, and T. Nyström, “Negative regulation by RpoS: A case of sigma factor competition,” *Mol. Microbiol.*, vol. 29, no. 4, pp. 1039–1051, 1998.

[142] A. Battesti, N. Majdalani, and S. Gottesman, “The RpoS-mediated general stress response in Escherichia coli.,” *Annu. Rev. Microbiol.*, vol. 65, pp. 189–213, 2011.

[143] A. G. Franchini, J. Ihssen, and T. Egli, “Effect of global regulators RpoS and cyclic-AMP/CRP on the catabolome and transcriptome of Escherichia coli K12 during carbon- and energy-limited growth,” *PLoS One*, vol. 10, no. 7, 2015.

[144] S. Mukhopadhyay, J. P. Audia, R. N. Roy, and H. E. Schellhorn, “Transcriptional induction of the conserved alternative sigma factor RpoS in Escherichia coli is dependent on BarA, a probable two-component regulator,” *Mol. Microbiol.*, vol. 37, no. 2, pp. 371–381, 2000.

[145] F. Mika and R. Hengge, “A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of ??S (RpoS) in E. coli,” *Genes Dev.*, vol. 19, no. 22, pp. 2770–2781, 2005.

[146] R. Hengge, “Proteolysis of σS (RpoS) and the general stress response in Escherichia coli,” *Res. Microbiol.*, vol. 160, pp. 667–676, 2009.

[147] X. Tu, T. Latifi, A. Bougdour, S. Gottesman, and E. A. Groisman, “The PhoP/PhoQ two-component system stabilizes the alternative sigma factor RpoS in Salmonella enterica.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 36, pp. 13503–8, 2006.

[148] L. Brown, D. Gentry, T. Elliott, and M. Cashel, “DksA affects ppGpp induction of RpoS at a translational level,” *J. Bacteriol.*, vol. 184, no. 16, pp. 4455–4465, 2002.

[149] S. M. Soucy, J. Huang, and J. P. Gogarten, “Horizontal gene transfer: building the web of life,” *Nat. Rev. Genet.*, vol. 16, no. 8, pp. 472–482, 2015.

[150] A. Iyer, “Bacteriophages in *Escherichia coli* antimicrobial resistance,” *Adv. Biosci. Biotechnol.*, vol. 4, no. 3, pp. 469–476, 2013.

[151] E. Kutter and A. Sulakvelidze, *Bacteriophages: biology and applications*. 2004.

[152] B. Guttman, R. Raya, and E. Kutter, “Basic Phage Biology,” *Bacteriophages Biol. Appl.*, pp. 29–63, 2005.

[153] N. D. Zinder and J. Lederberg, “GENETIC EXCHANGE IN SALMONELLA 1,” *J. Bacteriol.*, vol. 64, no. 5, p. 679, 1952.

[154] L. C. Thomason, N. Costantino, and D. L. Court, “*E. coli* Genome Manipulation by P1 Transduction,” in *Current Protocols in Molecular Biology*, 2007, p. 1.17.1-1.17.8.

[155] R. K. Chan and D. Botstein, “Specialized transduction by bacteriophage P22 in Salmonella typhimurium: genetic and physical structure of the transducing genomes

and the prophage attachment site," *Genetics*, vol. 83, no. 3 II, pp. 433–458, 1976.

[156] M. B. Łobocka *et al.*, "Genome of bacteriophage P1," *J. Bacteriol.*, vol. 186, no. 21, pp. 7032–7068, 2004.

[157] S. K. Amundsen, A. F. Taylor, and G. R. Smith, "The RecD subunit of the Escherichia coli RecBCD enzyme inhibits RecA loading, homologous recombination, and DNA repair.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 97, no. 13, pp. 7399–404, 2000.

[158] A. F. Taylor, D. W. Schultz, A. S. Ponticelli, and G. R. Smith, "RecBC enzyme nicking at chi sites during DNA unwinding: Location and orientation-dependence of the cutting," *Cell*, vol. 41, no. 1, pp. 153–163, 1985.

[159] D. Dermić, "Functions of multiple exonucleases are essential for cell viability, DNA repair and homologous recombination in recD mutants of Escherichia coli," *Genetics*, vol. 172, no. 4, pp. 2057–2069, 2006.

[160] T. Baba *et al.*, "Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection.," *Mol. Syst. Biol.*, vol. 2, p. 2006.0008, 2006.

[161] E. Guisbert, V. A. Rhodius, N. Ahuja, E. Witkin, and C. A. Gross, "Hfq modulates the sigmaE-mediated envelope stress response and the sigma32-mediated cytoplasmic stress response in Escherichia coli.," *J. Bacteriol.*, vol. 189, no. 5, pp. 1963–73, Mar. 2007.

[162] P. P. Cherepanov and W. Wackernagel, "Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant.," *Gene*, vol. 158, no. 1, pp. 9–14, 1995.

[163] M. S. McClain, I. C. Blomfield, and B. I. Eisenstein, "Roles of fimB and fimE in site-specific DNA inversion associated with phase variation of type 1 fimbriae in Escherichia coli.," *J. Bacteriol.*, vol. 173, no. 17, pp. 5308–14, Sep. 1991.

[164] M. S. McClain, I. C. Blomfield, K. J. Eberhardt, and B. I. Eisenstein, "Inversion-independent phase variation of type 1 fimbriae in Escherichia coli," *J. Bacteriol.*, vol. 175, no. 14, pp. 4335–4344, 1993.

[165] S. El-Labany, B. K. Sohanpal, M. Lahooti, R. Akerman, and I. C. Blomfield, "Distant cis-active sequences and sialic acid control the expression of fimB in Escherichia coli K-12," *Mol. Microbiol.*, vol. 49, no. 4, pp. 1109–1118, 2003.

[166] V. A. Rhodius, W. C. Suh, G. Nonaka, J. West, and C. A. Gross, "Conserved and variable functions of the ??E stress response in related genomes," *PLoS Biol.*, vol. 4, no. 1, pp. 0043–0059, 2006.

[167] E. B. Gogol, V. A. Rhodius, K. Papenfort, J. Vogel, and C. A. Gross, "Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon," *Proc. Natl. Acad. Sci.*, vol. 108, no. 31, pp. 12875–12880, 2011.

[168] A. Moores *et al.*, "RfaH suppresses small RNA MicA inhibition of fimB expression in Escherichia coli K-12," *J. Bacteriol.*, vol. 196, no. 1, pp. 148–156, 2014.

[169] Qiagen and Qiagen, "QIAGEN® Plasmid Purification Handbook," *Plasmid*, no. November, pp. 1–52, 2005.

[170] Qiagen, “QIAamp DNA Mini Kit - QIAGEN,” *Qiagene Protoc.*, no. June, p. 72, 2012.

[171] 2, “QIAGEN DNA extraction,” *DNA Extr. Protoc.*, no. August, 2001.

[172] Quiagen, “QIAquick ® Spin Handbook QIAGEN Sample and Assay Technologies,” *Qiagen*, vol. Volume 20, no. March, pp. 9–10, 2008.

[173] J. Sambrook, E. F. Fritsch, and T. Maniatis, “Molecular cloning: a laboratory manual,” *Cold Spring Harbor Laboratory Press*. p. 626, 1989.

[174] B. Müller-Hill, “Experiments with gene fusions,” *Trends Genet.*, vol. 1, p. 61, 1985.

[175] NEB, “NEBuffer Performance Chart with Restriction Enzymes,” *Ipswich, Massachusetts, USA*, 2015.

[176] QIAGEN, “QIAquick Gel Extraction Kit Protocol,” *QIAquick Spin Handb.*, vol. 03/2001, pp. 23–24, 2001.

[177] C. C. Richardson, “Enzymes in DNA Metabolism,” *Annu. Rev. Biochem.*, vol. 38, no. 1, pp. 795–840, Jun. 1969.

[178] D. Hanahan, J. Jessee, and F. R. Bloom, “Plasmid transformation of Escherichia coli and other bacteria,” *Methods in Enzymology*, vol. 204. pp. 63–114, 1991.

[179] D. Yu, H. M. Ellis, E. C. Lee, N. a Jenkins, N. G. Copeland, and D. L. Court, “An efficient recombination system for chromosome engineering in Escherichia coli,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 97, no. 11, pp. 5978–5983, 2000.

[180] L. C. Thomason, N. Costantino, D. V. Shaw, and D. L. Court, “Multicopy plasmid modification with phage ?? Red recombineering,” *Plasmid*, vol. 58, no. 2, pp. 148–158, 2007.

[181] R. K. Saiki *et al.*, “Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia.,” *Science*, vol. 230, no. 4732, pp. 1350–1354, 1985.

[182] I. C. Blomfield, V. Vaughn, R. F. Rest, and B. I. Eisenstein, “Allelic exchange in Escherichia coli using the Bacillus subtilis sacB gene and a temperature-sensitive pSC101 replicon,” *Mol. Microbiol.*, vol. 5, no. 6, pp. 1447–1457, 1991.

[183] J. H. Miller, “Experiments in molecular genetics,” *Cold Spring Harb. Lab. Press. Cold Spring Harb. NY*, vol. 433, pp. 352–355, 1972.

[184] M. B. Lees and S. Paxman, “Modification of the lowry procedure for the analysis of proteolipid protein,” *Anal. Biochem.*, vol. 47, no. 1, pp. 184–192, 1972.

[185] L. Technologies, “WesternBreeze ® Chemiluminescent Western Blot Immunodetection Kit Limited Product Warranty and Disclaimer Details.”

[186] Thermofisher, “iBlot ® Gel Transfer System.” Thermofisher, 2011.

[187] T. Bergmiller, M. Ackermann, and O. K. Silander, “Patterns of evolutionary conservation of essential genes correlate with their compensability,” *PLoS Genet.*, vol. 8, no. 6, 2012.

[188] A. Devaraj and K. Fredrick, “Short spacing between the Shine-Dalgarno sequence

and P codon destabilizes codon-anticodon pairing in the P site to promote +1 programmed frameshifting," *Mol. Microbiol.*, vol. 78, no. 6, pp. 1500–1509, 2010.

[189] P. Perron-Savard, G. De Crescenzo, and H. Le Moual, "Dimerization and DNA binding of the Salmonella enterica PhoP response regulator are phosphorylation independent," *Microbiology*, vol. 151, no. 12, pp. 3979–3987, Dec. 2005.

[190] K. Yamamoto, K. Hirao, T. Oshima, H. Aiba, R. Utsumi, and A. Ishihama, "Functional characterization in vitro of all two-component signal transduction systems from Escherichia coli," *J. Biol. Chem.*, vol. 280, no. 2, pp. 1448–1456, 2005.

[191] N. Casali and A. Preston, "E. coli Plasmid Vectors : Methods and Applications," *Methods in Molecular Biology*, vol. 235. p. 316, 2003.

[192] G. Aidelberg, B. D. Towbin, D. Rothschild, E. Dekel, A. Bren, and U. Alon, "Hierarchy of non-glucose sugars in Escherichia coli," *BMC Syst. Biol.*, vol. 8, no. 1, p. 133, 2014.

[193] A. Bougdour, C. Cunning, P. J. Baptiste, T. Elliott, and S. Gottesman, "Multiple pathways for regulation of σS (RpoS) stability in Escherichia coli via the action of multiple anti-adaptors," *Mol. Microbiol.*, vol. 68, no. 2, pp. 298–313, 2008.

[194] K. Moon and S. Gottesman, "A PhoQ/P-regulated small RNA regulates sensitivity of Escherichia coli to antimicrobial peptides," *Mol. Microbiol.*, vol. 74, no. 6, pp. 1314–1330, 2009.

[195] I. N. Wang, "Lysis timing and bacteriophage fitness," *Genetics*, vol. 172, no. 1, pp. 17–26, 2006.

[196] T. E. Waddell, K. Franklin, A. Mazzocco, A. M. Kropinski, and R. P. Johnson, "Generalized transduction by lytic bacteriophages.," *Methods Mol. Biol.*, vol. 501, pp. 293–303, 2009.

[197] M. Masters, "The frequency of P1 transduction of the genes of Escherichia coli as a function of chromosomal position: preferential transduction of the origin of replication.," *Mol. Gen. Genet.*, vol. 155, no. 2, pp. 197–202, Oct. 1977.

[198] L. H. Caporale, *The implicit genome*. Oxford University Press, 2006.

[199] J. M. Peregrin-Alvarez, S. Tsoka, and C. A. Ouzounis, "The phylogenetic extent of metabolic enzymes and pathways," *Genome Res.*, vol. 13, no. 3, pp. 422–427, 2003.

[200] K. E. Kram *et al.*, "Adaptation of Escherichia coli to Long-Term Serial Passage in Complex Medium: Evidence of Parallel Evolution.," *mSystems*, vol. 2, no. 2, 2017.

[201] C. Martelli, A. De Martino, E. Marinari, M. Marsili, and I. Pérez Castillo, "Identifying essential genes in Escherichia coli from a metabolic optimization principle.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 8, pp. 2607–11, 2009.

[202] A. R. Joyce *et al.*, "Experimental and computational assessment of conditionally essential genes in Escherichia coli," *J. Bacteriol.*, vol. 188, no. 23, pp. 8259–8271, 2006.

[203] B. Schiffrin *et al.*, "Skp is a multivalent chaperone of outer-membrane proteins," *Nat. Struct. & Mol. Biol.*, no. July, pp. 1–11, 2016.

[204] R. Freitag, "Biotechnology Applications," *Encycl. Anal. Sci.*, no. 1, pp. 258–267, 2005.

[205] A. Kato, H. Tanabe, and R. Utsumi, “Molecular characterization of the PhoP-PhoQ two-component system in Escherichia coli K-12: identification of extracellular Mg2+-responsive promoters.,” *J. Bacteriol.*, vol. 181, pp. 5516–5520, 1999.

[206] L. A. Clifton *et al.*, “Effect of divalent cation removal on the structure of gram-negative bacterial outer membrane models.,” *Langmuir*, vol. 31, no. 1, pp. 404–12, 2015.

[207] H. L. Alakomi, A. Paananen, M. L. Suihko, I. M. Helander, and M. Saarela, “Weakening effect of cell permeabilizers on gram-negative bacteria causing biodeterioration,” *Appl. Environ. Microbiol.*, vol. 72, no. 7, pp. 4695–4703, 2006.

[208] K. Iida and M. Koike, “Cell wall alterations of gram-negative bacteria by aminoglycoside antibiotics.,” *Antimicrob. Agents Chemother.*, vol. 5, no. 1, pp. 95–7, Jan. 1974.

[209] M. R. Hemm, B. J. Paul, J. Miranda-Ríos, A. Zhang, N. Soltanzad, and G. Storz, “Small stress response proteins in Escherichia coli: Proteins missed by classical proteomic studies,” *J. Bacteriol.*, vol. 192, no. 1, pp. 46–58, 2010.

[210] H. K. Alexander and L. M. Wahl, “Fixation probabilities depend on life history: Fecundity, generation time and survival in a burst-death model,” *Evolution (N. Y).*, vol. 62, no. 7, pp. 1600–1609, 2008.

[211] Z. Patwa *et al.*, “The fixation probability of beneficial mutations.,” *J. R. Soc. Interface*, vol. 5, no. 28, pp. 1279–89, 2008.

[212] V. Bacun-Druzina, Z. Cagalj, and K. Gjuracic, “The growth advantage in stationary-phase (GASP) phenomenon in mixed cultures of enterobacteria,” *FEMS Microbiol. Lett.*, vol. 266, no. 1, pp. 119–127, 2007.

[213] L. Boe, M. Danielsen, S. Knudsen, J. B. Petersen, J. Maymann, and P. R. Jensen, “The frequency of mutators in populations of Escherichia coli,” *Mutat. Res. - Fundam. Mol. Mech. Mutagen.*, vol. 448, no. 1, pp. 47–55, 2000.

[214] J. L. Gibson *et al.*, “The ??e stress response is required for stress-induced mutation and amplification in Escherichia coli,” *Mol. Microbiol.*, vol. 77, no. 2, pp. 415–430, 2010.

[215] C. D. Herring and F. R. Blattner, “Conditional Lethal Amber Mutations in Essential Escherichia coli Genes,” *J. Bacteriol.*, vol. 186, no. 9, pp. 2673–2681, 2004.

[216] M. Obrist, S. Milek, E. Klauck, R. Hengge, and F. Narberhaus, “Region 2.1 of the Escherichia coli heat-shock sigma factor RpoH (??32) is necessary but not sufficient for degradation by the FtsH protease,” *Microbiology*, vol. 153, no. 8, pp. 2560–2571, 2007.

[217] J. Mecsas, P. E. Rouviere, J. W. Erickson, T. J. Donohue, and C. A. Gross, “The activity of sigma E, an Escherichia coli heat-inducible sigma-factor, is modulated by expression of outer membrane proteins,” *Genes Dev*, vol. 7, no. 12B, pp. 2618–2628, 1993.

[218] S. E. Goelz and J. E. Cronan, “The positional distribution of fatty acids in Escherichia coli phospholipids is not regulated by sn-glycerol 3-phosphate levels,” *J. Bacteriol.*, vol. 144, no. 1, pp. 462–464, 1980.

[219] C. Dartigalongue, D. Missiakas, and S. Raina, “Characterization of the Escherichia coli σE Regulon,” *J. Biol. Chem.*, vol. 276, no. 24, pp. 20866–20875, 2001.

[220] U. K. Sharma and D. Chatterji, “Transcriptional switching in Escherichia coli during stress and starvation by modulation of σ70 activity,” *FEMS Microbiology Reviews*, vol. 34, no. 5. pp. 646–657, 2010.

[221] M. Shepherd, “The CydDC ABC transporter of Escherichia coli: new roles for a reductant efflux pump,” *Biochem. Soc. Trans.*, vol. 43, no. 5, pp. 908–912, 2015.

[222] A. M. Lippa and M. Goulian, “Perturbation of the oxidizing environment of the periplasm stimulates the PhoQ/PhoP system in Escherichia coli.,” *J. Bacteriol.*, vol. 194, no. 6, pp. 1457–63, Mar. 2012.

[223] R. Horlacher and W. Boos, “Characterization of TreR, the major regulator of the Escherichia coli trehalose system,” *J. Biol. Chem.*, vol. 272, no. 20, pp. 13026–13032, 1997.

[224] E. A. Groisman, J. Kayser, and F. C. Soncini, “Regulation of polymyxin resistance and adaptation to low-Mg2+ environments,” *J. Bacteriol.*, vol. 179, no. 22, pp. 7040–7045, 1997.

[225] C. A. Henderson *et al.*, “Characterization of MicA interactions suggests a potential novel means of gene regulation by small non-coding RNAs,” *Nucleic Acids Res.*, vol. 41, no. 5, pp. 3386–3397, 2013.

[226] K. I. Udekwu and E. G. H. Wagner, “Sigma E controls biogenesis of the antisense RNA MicA.,” *Nucleic Acids Res.*, vol. 35, pp. 1279–1288, 2007.

[227] S. Raina, D. Missiakas, and C. Georgopoulos, “The rpoE gene encoding the sigma E (sigma 24) heat shock sigma factor of Escherichia coli.,” *The EMBO journal*, vol. 14, no. 5. pp. 1043–1055, 1995.

[228] P. Soriano, “Generalized lacZ expression with the ROSA26 Cre reporter strain.,” *Nat. Genet.*, vol. 21, no. 1, pp. 70–71, 1999.

[229] G. Klein, A. Stupak, D. Biernacka, P. Wojtkiewicz, B. Lindner, and S. Raina, “Multiple transcriptional factors regulate transcription of the rpoE Gene in Escherichia coli under different growth conditions and when the lipopolysaccharide biosynthesis is defective,” *J. Biol. Chem.*, vol. 291, no. 44, pp. 22999–23019, 2016.

[230] M. D. Abràmoff, P. J. Magalhães, and S. J. Ram, “Image processing with imageJ,” *Biophotonics International*, vol. 11, no. 7. pp. 36–41, 2004.

[231] S. Sternberg, “Biomedical Image Processing,” *IEEE Comput.*, vol. 16, pp. 22–34, 1983.

[232] X. Xia, “Position Weight Matrix, Gibbs Sampler, and the Associated Significance Tests in Motif Characterization and Prediction,” *Scientifica (Cairo).*, vol. 2012, p. e917540, 2012.

[233] J. E. Reid and L. Wernisch, “STEME: Efficient em to find motifs in large data sets,” *Nucleic Acids Res.*, vol. 39, no. 18, 2011.

[234] K. Okonechnikov *et al.*, “Unipro UGENE: A unified bioinformatics toolkit,” *Bioinformatics*, vol. 28, no. 8. pp. 1166–1167, 2012.

[235] N. Jayaram, D. Usvyat, and A. C. R. Martin, “Evaluating tools for transcription factor binding site prediction,” *BMC Bioinformatics*, 2016.

[236] C. E. Grant, T. L. Bailey, and W. S. Noble, “FIMO: Scanning for occurrences of a given motif,” *Bioinformatics*, vol. 27, no. 7, pp. 1017–1018, 2011.

[237] K. Yamamoto, H. Ogasawara, N. Fujita, R. Utsumi, and A. Ishihama, “Novel mode of transcription regulation of divergently overlapping promoters by PhoP, the regulator of two-component system sensing external magnesium availability,” *Mol. Microbiol.*, vol. 45, no. 2, pp. 423–438, 2002.

[238] A. Forbes, “Classification-algorithm evaluation: five performance measures based on confusion matrices,” *J. Clin. Monit. Comput.*, vol. 11, no. 3, pp. 189–206, 1995.

[239] E. C. Rouchka and C. T. Hardin, “rMotifGen: random motif generator for DNA and protein sequences.,” *BMC Bioinformatics*, vol. 8, p. 292, 2007.

[240] M. D. Johnson, N. A. Burton, B. Gutiérrez, K. Painter, and P. A. Lund, “RcsB is required for inducible acid resistance in Escherichia coli and acts at gadE-dependent and -independent promoters,” *J. Bacteriol.*, vol. 193, no. 14, pp. 3653–3656, 2011.

[241] J. J. Bass *et al.*, “An overview of technical considerations for Western blotting applications to physiological research,” *Scandinavian Journal of Medicine and Science in Sports*, vol. 27, no. 1. pp. 4–25, 2017.

[242] N. S. Lipman, L. R. Jackson, L. J. Trudel, and F. Weis-Garcia, “Monoclonal Versus Polyclonal Antibodies: Distinguishing Characteristics, Applications, and Information Resources,” *ILAR J.*, vol. 46, no. 3, pp. 258–268, 2005.

[243] O. Bannach *et al.*, “Detection of prion protein particles in blood plasma of scrapie infected sheep,” *PLoS One*, vol. 7, no. 5, 2012.

[244] S. J. Labrie, J. E. Samson, and S. Moineau, “Bacteriophage resistance mechanisms,” *Nat. Rev. Microbiol.*, vol. 8, no. 5, pp. 317–327, 2010.

[245] B. J. Newman and M. Masters, “The variation in frequency with which markers are transduced by phage P1 is primarily a result of discrimination during recombination,” *MGG Mol. Gen. Genet.*, vol. 180, no. 3, pp. 585–589, 1980.

[246] T. Miyashiro and M. Goulian, “Stimulus-dependent differential regulation in the Escherichia coli PhoQ PhoP system,” *Proc. Natl. Acad. Sci.*, vol. 104, no. 41, pp. 16305–16310, 2007.

[247] Y. Eguchi *et al.*, “Transcriptional regulation of drug efflux genes by EvgAS, two-component system in Escherichia coli,” *Microbiology*, vol. 149, no. 10, pp. 2819–2828, 2003.

[248] “Rao Research Group.” [Online]. Available: http://www.scs.illinois.edu/rao/protocol-ecoli-p1.php. [Accessed: 22-Sep-2017].

[249] B. San Francisco and K. Lab, “P1 TRANSDUCTION (adapted from initial protocol by Dr. Brad Weart-Levin lab) Step 2: Grow phage on donor strain,” 2008.

[250] “Methods:P1 transduction - EcoliWiki.” [Online]. Available: http://ecoliwiki.net/colipedia/index.php/Methods:P1_transduction. [Accessed: 22-

Sep-2017].

[251] Z. D. Dalebroux, S. Matamouros, D. Whittington, R. E. Bishop, and S. I. Miller, “PhoPQ regulates acidic glycerophospholipid content of the Salmonella Typhimurium outer membrane.,” *Proc. Natl. Acad. Sci. U. S. A.*, pp. 6–11, 2014.

[252] C. Pelletier, P. Bourlioux, and J. Van Heijenoort, “Effects of sub-minimal inhibitory concentrations of EDTA on growth of Escherichia coli and the release of lipopolysaccharide,” *FEMS Microbiology Letters*, vol. 117, no. 2. pp. 203–206, 1994.

[253] R. Ernst, C. S. Ejsing, and B. Antonny, “Homeoviscous Adaptation and the Regulation of Membrane Lipids,” *Journal of Molecular Biology*, vol. 428, no. 24. pp. 4776–4791, 2016.

[254] A. Emiola, S. S. Andrews, C. Heller, and J. George, “Crosstalk between the lipopolysaccharide and phospholipid pathways during outer membrane biogenesis in *Escherichia coli*,” *Proc. Natl. Acad. Sci.*, vol. 113, no. 11, p. 201521168, 2016.

[255] E. D. Bergmann and S. Sicher, “Mode of action of chloramphenicol.,” *Nature*, vol. 170, no. 4335, pp. 931–932, 1952.

[256] E. Scholar, “Kanamycin,” in *xPharm: The Comprehensive Pharmacology Reference*, 2007, pp. 1–5.

[257] S. Lejona, M. E. Castelli, M. L. Cabeza, L. J. Kenney, E. García Véscovi, and F. C. Soncini, “PhoP Can Activate Its Target Genes in a PhoQ-Independent Manner,” *J. Bacteriol.*, vol. 186, no. 8, pp. 2476–2480, 2004.

[258] H. Nicoloff, S. Gopalkrishnan, and S. E. Ades, “Appropriate Regulation of the σ(E)-Dependent Envelope Stress Response Is Necessary To Maintain Cell Envelope Integrity and Stationary-Phase Survival in Escherichia coli.,” *J. Bacteriol.*, vol. 199, no. 12, p. JB.00089-17, Jun. 2017.

[259] D. Y. Kim, “Two stress sensor proteins for the expression of sigmaE regulon: DegS and RseB,” *Journal of Microbiology*, vol. 53, no. 5. pp. 306–310, 2015.

[260] B. O. Cezairliyan and R. T. Sauer, “Inhibition of regulated proteolysis by RseB.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 10, pp. 3771–3776, 2007.

[261] A. A. M. Al Mamun *et al.*, “Identity and function of a large gene network underlying mutagenic repair of DNA breaks.,” *Science*, vol. 338, no. 6112, pp. 1344–8, 2012.

[262] O. Fayet, T. Ziegelhoffer, and C. Georgopoulos, “The groES and groEL heat shock gene products of Escherichia coli are essential for bacterial growth at all temperatures,” *J. Bacteriol.*, vol. 171, no. 3, pp. 1379–1385, 1989.

[263] H. Ogasawara, A. Hasegawa, E. Kanda, T. Miki, K. Yamamoto, and A. Ishihama, “Genomic SELEX search for target promoters under the control of the PhoQP-RstBA signal relay cascade,” *J. Bacteriol.*, vol. 189, no. 13, pp. 4791–4799, 2007.

[264] W. Jia, A. El Zoeiby, T. N. Petruzziello, B. Jayabalasingham, S. Seyedirashti, and R. E. Bishop, “Lipid trafficking controls endotoxin acylation in outer membranes of Escherichia coli.,” *J. Biol. Chem.*, vol. 279, pp. 44966–44975, 2004.

[265] N. A. Burton, M. D. Johnson, P. Antczak, A. Robinson, and P. A. Lund, “Novel Aspects

of the Acid Response Network of E. coli K-12 Are Revealed by a Study of Transcriptional Dynamics," *J. Mol. Biol.*, vol. 401, no. 5, pp. 726–742, 2010.

[266] Y. Eguchi, E. Ishii, K. Hata, and R. Utsumi, "Regulation of acid resistance by connectors of two-component signal transduction systems in Escherichia coli," *J. Bacteriol.*, vol. 193, no. 5, pp. 1222–1228, 2011.

[267] Y. Eguchi *et al.*, "Signal Transduction Cascade between EvgA/EvgS and PhoP/PhoQ Two-Component Systems of Escherichia coli," *J. Bacteriol.*, vol. 186, no. 10, pp. 3006–3014, 2004.

[268] M. M. Haque and S. Tsuyumu, "Virulence, resistance to magainin II, and expression of pectate lyase are controlled by the PhoP-PhoQ two-component regulatory system responding to pH and magnesium in Erwinia chrysanthemi 3937," *J. Gen. Plant Pathol.*, vol. 71, no. 1, pp. 47–53, 2005.

[269] B. Aman and G. Ciobanu, "Mutual Mobile Membranes Systems with Surface Objects," in *Seventh Brainstorming Week on Membrane Computing*, 2009, vol. 1, pp. 29–39.